



Balkan Journal of Electrical & Computer Engineering

An International Peer Reviewed, Referred, Indexed and Open Access Journal

www.bajece.com

Vol : 10
No : 2
Year : 2022
ISSN : 2147 - 284X



It is abstracted and indexed in, Index Google Scholarship, the PSCR, Cross ref, DOAJ, Research Bible, Indian Open Access Journals (OAJ), Institutional Repositories (IR), J-Gate (Informatics India), Ulrich's, International Society of Universal Research in Sciences, DRJI, EyeSource, Cosmos Impact Factor, Cite Factor, SIS Scientific Indexing Service, IJIF, iifactor. ULAKBİM-TR Dizin.

General Publication Director & Editor-in-Chief
Musa Yılmaz, Batman University, Turkey.

Vice Editor
Hamidreza Nazarpouya, University of California Riverside, USA

Scientific Committee
Abhishek Shukla (India)
Abraham Lomi (Indonesia)
Aleksandar Georgiev (Bulgaria)
Arunas Lipnickas (Lithuania)
Audrius Senulis (Lithuania)
Belle R. Upadhyaya (USA)
Brijender Kahanwal (India)
Chandar Kumar Chanda (India)
Daniela Dzhonova-Atanasova (Bulgaria)
Deris Stiawan (Indonesia)
Emel Onal (Turkey)
Emine Ayaz (Turkey)
Enver Hatimi (Kosovo)
Ferhat Sahin (USA)
Gursel Alici (Australia)
Hakan Temeltaş (Turkey)
Ibrahim Akduman (Turkey)
Jan Izykowski (Poland)
Javier Bilbao Landatxe (Spain)
Jelena Dikun (Lithuania)
Karol Kyslan (Slovakia)
Kunihiko Nabeshima (Japan)
Lambros Ekonomou (Greece)
Lazhar Rahmani (Algerie)
Marcel Istrate (Romania)
Marija Eidukeviciute (Lithuania)
Milena Lazarova (Bulgaria)
Muhammad Hadi (Australia)
Muhamed Turkanović (Slovenia)
Mourad Houabes (Algerie)
Murari Mohan Saha (Sweden)
Nick Papanikolaou (Greece)
Okyay Kaynak (Turkey)
Osman Nuri Ucan (Turkey)
Ozgun E. Mustecaplioglu (Turkey)
Padmanaban Sanjeevikumar (India)
Ramazan Caglar (Turkey)
Rumen Popov (Bulgaria)
Tarek Bouktir (Algeria)
Sead Berberovic (Croatia)
Seta Bogosyan (USA)
Savvas G. Vassiliadis (Greece)
Suwarno (Indonesia)
Tulay Adali (USA)
Yogeshwarsing Calleecharan (Mauritius)
YangQuan Chen (USA)
Youcef Soufi (Algeria)

Aim & Scope

The journal publishes original papers in the extensive field of Electrical-Electronics and Computer engineering. It accepts contributions which are fundamental for the development of electrical engineering, computer engineering and its applications, including overlaps to physics. Manuscripts on both theoretical and experimental work are welcome. Review articles and letters to the editors are also included.

Application areas include (but are not limited to): Electrical & Electronics Engineering, Computer Engineering, Software Engineering, Biomedical Engineering, Electrical Power Engineering, Control Engineering, Signal and Image Processing, Communications & Networking, Sensors, Actuators, Remote Sensing, Consumer Electronics, Fiber-Optics, Radar and Sonar Systems, Artificial Intelligence and its applications, Expert Systems, Medical Imaging, Biomedical Analysis and its applications, Computer Vision, Pattern Recognition, Robotics, Industrial Automation.



ISSN: 2147- 284X
Vol: 10
No : 2
Year: April 2022

CONTENTS

- S. Buyrukoğlu, A. Akbaş;** Machine Learning based Early Prediction of Type 2 Diabetes: A New Hybrid Feature Selection Approach using Correlation Matrix with Heatmap and SFS,.....**110 - 117**
- S. İkizoğlu;** Distributed Formation Control Algorithm for Improved String Stability in Heterogenic Vehicle Platoons,.....**118 - 124**
- H. Çelebi;** A Dual-Radio Hybrid Mesh Topology for Multi-Hop Industrial IoT Networks in Harsh Environments,.....**125 - 131**
- İ. E. Saçu;** Fractional Integration Based Feature Extractor for EMG Signals,.....**132 - 138**
- S. Yılmaz;** Stabilization of Chaos in a Cancer Model: The Effect of Oncotripsy,.....**139 - 149**
- Ö. F. Ertuğrul, Y. Sönmez, N. Sezgin, E. Akil;** Assessment of Epileptic Seizures and Non-Epileptic Seizures via Wearable Sensors and Priori Detection of Epileptic Seizures,.....**150 - 155**
- L. Kouhalvandi;** A Prospective Look on Optimization Methods For RFID Systems: Requirements, Challenges and Implementation Aspects,.....**156 - 169**
- D. Çelik;** Performance Analysis of Three Levels Three Switches Vienna-Type Rectifier Based on Direct Power Control,.....**170 - 177**
- D. Özcelik, Ö. Taştan;** A Weakly Supervised Clustering Method for Cancer Subgroup Identification,.....**178 - 186**
- Z. O. Davarci, M. Şahin, O. Akar;** Estimation by ANN of Luminous Efficacy of Lamps Used for Lighting,.....**187 - 197**
- F. N. Deniz, M. Günay;** Coefficient Diagram Method Based Decentralized Controller for Fractional Order TITO Systems,....**198 - 208**
- T. Sönmezocak;** The Use of Kalman Filter in Control The PanTilt Two-Axis Robot With Wearable System,.....**209 - 213**
- Z. Yılmaz, L. Bayındır;** Lidar-based Robot Detection and Positioning using Machine Learning Methods,.....**214 - 223**
- M. Atasoyu;** A 0.6-V 11-uW PVT tolerant DTMOS inverter based OTA,.....**224 - 229**

**BALKAN
JOURNAL OF
ELECTRICAL & COMPUTER ENGINEERING**
(An International Peer Reviewed, Indexed and Open Access Journal)

Contact
Batman University
Department of Electrical-Electronics Engineering
Bati Raman Campus Batman-Turkey

Web: <http://dergipark.gov.tr/bajece>
<https://www.bajece.com>
e-mail: bajece@hotmail.com

Machine Learning based Early Prediction of Type 2 Diabetes: A New Hybrid Feature Selection Approach using Correlation Matrix with Heatmap and SFS

Selim Buyrukoglu and Ayhan Akbas

Abstract—A new hybrid machine learning method for the prediction of type 2 diabetes is introduced and explained in detail. Also outcomes are compared with the similar researches. Early prediction of diabetes is crucial to take necessary measures (i.e. changing eating habits, patient weight control etc.), to defer the emergence of diabetes and to reduce the death rate to some extent and ease medical care professionals' decision making in preventing and managing diabetes mellitus. The purpose of this study is the creation of a new hybrid feature selection approach combination of Correlation Matrix with Heatmap and Sequential forward selection (SFS) to reveal the most effective features in the detection of diabetes. A diabetes data set with 520 instances and seven features were studied with the application of the proposed hybrid feature selection approach. The evaluation of the selected optimal features was measured by applying Support Vector Machines(SVM), Random Forest(RF), and Artificial Neural Networks(ANN) classifiers. Five evaluation metrics, namely, Accuracy, F-measure, Precision, Recall, and AUC showed the best performance with ANN (99.1%), F-measure (99.1%), Precision (99.3%), Recall (99.1%), and AUC (99.8%). Our proposed hybrid feature selection model provided a more promising performance with ANN compared to other machine learning algorithms.

Index Terms—Artificial Neural Network, Correlation Matrix, Sequential Forward Selection, Diabetes Mellitus, Hybrid Feature Selection

I. INTRODUCTION


DIABETES Mellitus (DM), commonly known as diabetes, is a metabolic disease that causes higher-than-normal blood sugar level. The insulin hormone secreted by the body carries excess sugar in the blood to your cells to be stored or used for energy. In diabetes, the body either does not produce enough insulin or cannot use insulin effectively to reduce the glucose level in the blood. As a result, uncontrolled

high blood sugar can damage the nerve, eyes, kidneys, and other organs [1]. The general symptoms of diabetes include increased hunger, increased thirst, weight loss, frequent urination, blurry vision, extreme fatigue and sores that don't heal. Pre-diabetes occurs when your blood sugar is higher than normal, but usually it's not high enough for a diagnosis of type 2 diabetes. Therefore, early prediction of diabetes is crucial to take necessary measures (i.e. changing eating habit, weight control etc.), to defer the emergence of diabetes and to reduce the death rate to some extent.


Machine Learning (ML), a domain of artificial intelligence (AI), makes use of capabilities of computers to learn automatically without explicitly programmed and improve from experience. Machine learning in healthcare is becoming more widely used and is helping clinicians in many different ways [2]. One of the most common use cases in medicine for machine learning is the diagnosis decision support [3]. The power of machine learning models is its ability to process huge datasets beyond the scope of human capability [4], and then reliably convert analysis of that data into clinical insights to provide better information to doctors at the point of decision making. There are plenty of studies on Diabetes that employed a machine learning (ML) based models in the literature as systematically reviewed by De Silva et. al. (2020) for various databases[5]. In our research, we developed a hybrid model for selecting most dominating features so as to maximize the output accuracy. The data analyzed in this paper, has been collected from the patients of Sylhet Diabetes Hospital in Sylhet, Bangladesh (2020) [6]. Features in this dataset are listed and described in Table I.

In the literature, there are numerous studies on the diagnosis of diabetes mellitus. With advances in computer sciences, machine learning techniques have become common in health-care. Overview of studies employing ML models in their researches has been published by Chaki et. al. (2020) [7] and Kavakiotis et. al.(2017) [8] who conducted systematic reviews on the studies on machine learning and artificial intelligence based Diabetes Mellitus detection and provided a detailed overview of DM detection. Another group of researches, Jashwanth et. al.(2020) [9] studied the performance of six machine learning classifiers (namely, support vector

SELİM BUYRUKOĞLU is with the Department of Computer Engineering, Engineering Faculty, Cankiri Karatekin University, Cankiri, Turkey e-mail: (sbuyrukoglu@karatekin.edu.tr)

 <https://orcid.org/0000-0001-7844-3168>

AYHAN AKBAS is with the Department of Computer Engineering, Engineering Faculty, Abdullah Gul University, Kayseri, Turkey e-mail: (ayhan.akbas@agu.edu.tr)

 <https://orcid.org/0000-0002-6425-104X>

machine, K-nearest neighbors, logistic regression, naive bayes, gradient boosting and random forest) on Pima Indian Diabetes Database which is a common dataset for diabetes studies. The performance of all the six classifiers were compared using Accuracy score, Receiver Operating Curve (ROC), Precision, Recall, F-measure evaluated from each model. Random Forest classifier has shown the highest performance compared to remaining classifiers used in their proposed methodology. Lai et. al [10] also developed a predictive model to predict the risk for developing DM using Logistic Regression and Gradient Boosting Machine (GBM) techniques in Canadian patients. Fasting blood glucose, body mass index, high-density lipoprotein, and triglycerides were the most important predictors in their model. Accuracy for the proposed GBM model is 84.7% and for the proposed Logistic Regression model is 84.0%. In another study [11], classifiers (Decision Tree (DT), Artificial Neural Networks (ANN), Logistic Regression (LR) and Naive Bayes (NB)) were compared for the risk of diabetes prediction, then bagging and boosting techniques were investigated for improving the robustness. Random Forest (RF) algorithm was suggested as the best performance of disease risk classification. To the best of our knowledge, our approach is novel in terms of the selection of features and employment of hybrid combination of classifier to achieve the most accurate outcome. Most of the studies in the literature utilized single classifier models contrary to our approach. Moreover, we used a new dataset consisting of all diabetes symptoms unlike the great majority of studies on early prediction of diabetes that employed pima indian diabetes dataset[12] where most related symptoms are missing and features in the dataset are very limited.

Other research using deep learning techniques has been carried out by Swapna et al. [13]. They utilized the RR-interval signals known as heart rate variability (HRV) signals (derived from electrocardiogram (ECG) signals) and used for the non-invasive detection of diabetes. In their research, they employed long short-term memory (LSTM), convolutional neural network (CNN) and their combinations for extracting complex temporal dynamic features of the input HRV data which was then, passed to SVM (Support Vector Machine) for classification. The system proposed is able to diagnose diabetes using ECG signals with an accuracy of 95.7%. A similar study that employed machine learning approach was carried out by Zou et. al. [14]. In their study, they used decision tree, random forest and neural network to predict diabetes mellitus in people. The dataset they used is obtained from a hospital physical examinations and contains 14 attributes. Principal component analysis (PCA) and minimum redundancy maximum relevance (mRMR) are employed to reduce the dimensionality. The results showed that prediction with random forest could reach the highest accuracy (ACC = 0.8084) when all the attributes were used.

In the paper, we introduced the area of the study, mentioning the similar researches in section I. In the section II, the proposed approach is given together with the data collection and preprocessing steps. Here, classifiers used in the algorithms are summarized briefly. Experimental results and implementation are given and discussed in section III. Finally, section IV includes conclusions and future directions of the study.

II. PROPOSED APPROACH

In this study, we have implemented a new hybrid feature selection approach. Fig. 1 illustrates the proposed architecture which consists of several phases of data processing steps.

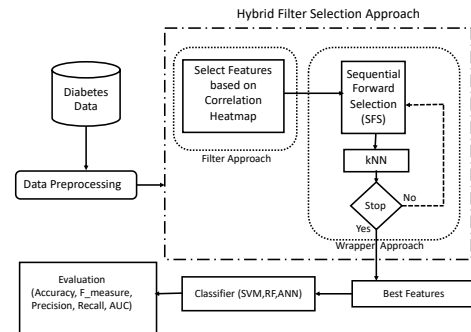


Fig. 1: Architecture of the proposed approach

A. Data Collection and Preprocessing

The diabetes data set is collected from patients in Sylhet Diabetes Hospital in Sylhet, Bangladesh [6]. It has 520 instances (342 confirmed Diabetes positive and 178 Diabetes negative) and also contains seventeen features and two classes. Description of datasets is presented in Table I. The type of these features are categorical except only age feature (numeric).

The aim of data collection here is to predict diabetes disease applying a new hybrid feature selection approach. The importance of the using features in the prediction of diabetes disease was already specified in Section I. The most important features in the prediction of diabetes disease will be highlighted through the proposed hybrid feature selection approach. There is on missing value and all observations are complete.

B. Correlation Matrix with Heatmap

Features relation to each other or the target feature is provided through correlation matrix as illustrated in Fig. 2. Correlation matrix is a graphical representation of data to identify which features are most related to the target feature. Each feature in a dataset is represented as colors which mean that the colors notice information to researchers in terms of the correlation between features. We created our heatmap of a correlation matrix which is presented in 5 steps in Algorithm 1.

C. Sequential Forward Selection (SFS)

Feature selection algorithms are used to eliminate unusable and redundant features. [15] highlighted that feature selection is an efficient way in terms of dimensional reduction of data with high dimensionality. Sequential forward selection (SFS) algorithm was applied to select an optimal subset of data from the 17 features. The reason behind the selection of SFS in this study is that wrapper feature selection approaches are widely

Feature	Description	Value	Ratio
Age	Age of patient	Numeric	Range:16-90 Mean:48, Std.Dev:12.15
Gender	Sex	Male/Female	63.1% / 36.9%
Polyuria	Frequent urination	Yes/No	50.4% / 49.6%
Polydipsia	Excessive thirst	Yes/No	55.2% / 44.8%
Sudden Weight Loss	Uncontrolled weight loss	Yes/No	58.3% / 41.7%
Weakness	Condition of being weak	Yes/No	58.7% / 41.3%
Polyphagia	Abnormal desire to consume excessive amounts of food	Yes/No	54.4% / 45.6%
Genital thrush	Genital fungal infections	Yes/No	77.7% / 22.3%
Visual Blurring	Unclear vision	Yes/No	55.2% / 44.8%
Itching	Irritating sensation of the skin	Yes/No	51.3% / 48.7%
Irritability	Become easily anger	Yes/No	75.8% / 24.2%
Delayed healing	Impaired healing	Yes/No	54.0% / 46.0%
Partial paresis	Slight/Partial Paralysis	Yes/No	56.9% / 43.1%
Muscle stiffness	Status that muscles feel tight and more difficult to move	Yes/No	62.5% / 37.5%
Alopecia	Hair Loss	Yes/No	65.6% / 34.4%
Obesity	Obesity	Yes/No	83.1% / 16.9%
Class (Target)	Diabetes Status	Positive/Negative	61.5% / 38.5%

TABLE I: Description of Features in the dataset

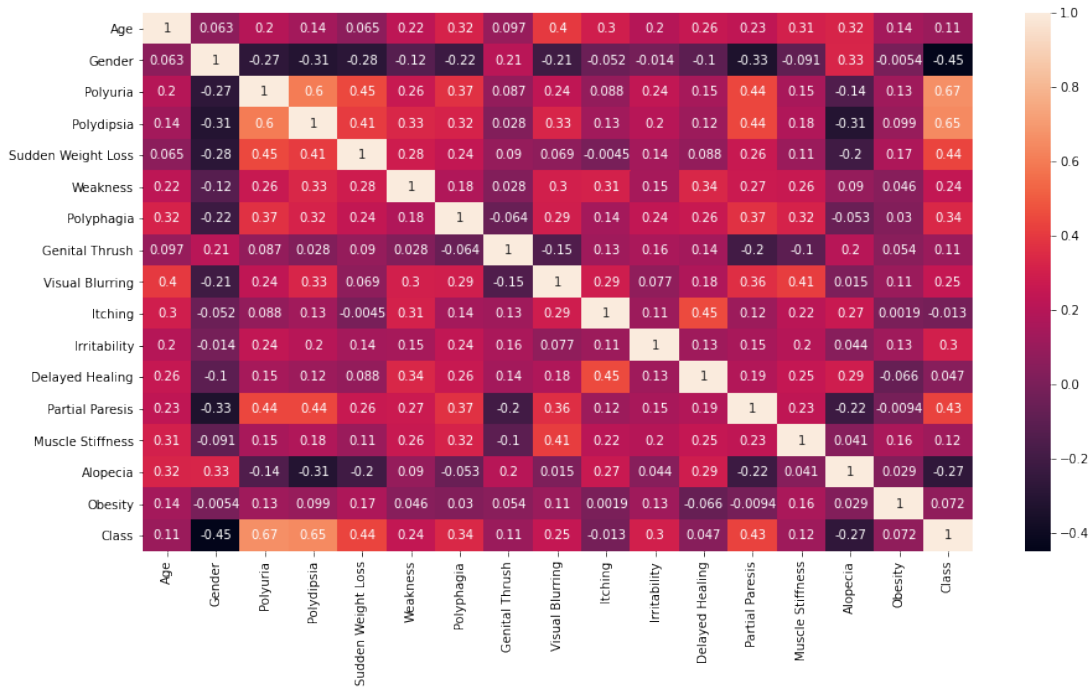


Fig. 2: Heatmap of the Correlation Matrix for the Diabetes Data

Algorithm 1: Steps of Creation of Correlation Matrix with Heatmap

1. Import Data;
2. Identify Independent Columns (X) and Target Column (Y);
3. Create Correlation Matrix (get correlations of each features in dataset);
4. Create Heatmap with size (17, 17);
5. Plot and Export Heatmap;

is created by the SFS algorithm initially. The feature has the highest feature importance score is added to the empty set. The features are ranked based on feature importance to extract the top features using extra tree classifier. Then, the feature has the second-highest score is added to the set. The feature adding process to the set continues until the classifier can no longer improve the performance. The k-nearest neighbors (KNN) classifier is implemented in this study since it is one of the most effective machine learning algorithm in terms of classification [16]. Once the best classification performance is achieved, the added features to the set are utilized for classification since they work best together. Algorithm 2 presents pseudo-code for the SFS algorithm.

used to detect fundamental interaction between features and also SFS is a user-friendly wrapper approach. An empty set

Algorithm 2: Sequential Forward Selection Pseudo Code

Input: $Y = y_1, y_2, \dots, y_d$ whole d -dimensional feature set are taken as input

Output: $X_k = x_j \mid j = 1, 2, \dots, k; x_j \in Y$, where $k = (0, 1, 2, \dots, d)$

SFS returns a subset of features; the number of selected features k , where $k < d$, has to be specified a priority.

1. Create an empty set: $X_0 = \phi, k = 0$

We initialize the algorithm with an empty set ϕ ("null set") so that $k = 0$ (where k is the size of the subset).

2. Select best remaining feature:

$$x^+ = \arg \max J(X_k + x), \text{ where } x \in Y - X_k$$

$$X_{k+1} = X_k + x^+$$

$$k = k + 1$$

Goto Step 2

3. Termination: $k = p$

D. Classifiers

Support Vector Machines (SVM): Support Vector Machine (SVM) is an associated machine learning algorithm mostly used for classification and regression analysis [17]. Given a set of training examples, An SVM maps training examples to points in space so as to maximize the width of the gap between the two categories. In other words, a support vector machine can be defined as a vector space-based machine learning method that finds a decision boundary between two classes that are the furthest from any point in the training data. Support vector machines are mostly used to separate binary classification data, for example separating each data in a data set into female or male.

Random Forest (RF): Random forests (RF) or random decision forests is a collective learning method that estimates class (classification) or number (regression) according to the type of the problem by creating a large number of decision trees during the training phase for classification, regression and other tasks [18]. Random decision forests address the problem of decision trees over fitting training sets. Different number of trees (50, 100, 250, 500) was tested in the prediction of type 2 diabetes. The best performance was obtained with five hundred trees.

Artificial Neural Network (ANN): Artificial neural networks (ANN) is a computing technology inspired by the information processing technique of the human brain [19]. With ANN, the way the simple biological nervous system works is imitated. That is to say, it is digital modeling of biological neuron cells and the synaptic bond that these cells establish with each other. Neurons connect to each other in various ways to form networks. These networks have the capacity to learn, memorize and reveal the relationship between data. In other words, ANNs provide solutions to problems that normally require a person's natural abilities to think and observe. The main reason why a person can produce solutions to problems that require thinking and observing skills is the ability of the

human brain, and therefore the human being, to learn by living or experimenting. In this study, one input layer with 7 nodes and one hidden layer with 14 neurons were tested.

E. Evaluation Metrics

The evaluation of the machine learning algorithm is always a vital part of every project. When the model is evaluated using one criterion, it may yield satisfactory results, but poor results may be obtained when compared with other criteria or metric. In our study, We mostly use classification accuracy to measure the performance of a model, but this is not enough to reach a solid conclusion. Mostly used classifiers are given here.

Accuracy: It is the ratio of number of correct predictions to the total number of input samples. Accuracy is a good measure when the target classes in the data are nearly balanced.

F-Measure: F-measure or F1 Score, also known as the Sørensen–Dice coefficient or Dice similarity coefficient (DSC), is the Harmonic Mean between precision and recall. The range for F1 Score is [0, 1]. It tells you how precise your classifier is (how many instances it classifies correctly), as well as how robust it is (it does not miss a significant number of instances).

10-fold Cross validation: Cross-validation is a technique to evaluate predictive models by partitioning the original sample into a training set to train the model, and a test set to evaluate it.

Precision and Recall: Precision (also called positive predictive value) is the fraction of relevant instances among the retrieved instances, while recall (also known as sensitivity) is the fraction of relevant instances that were retrieved. Both precision and recall are therefore based on relevance [20].

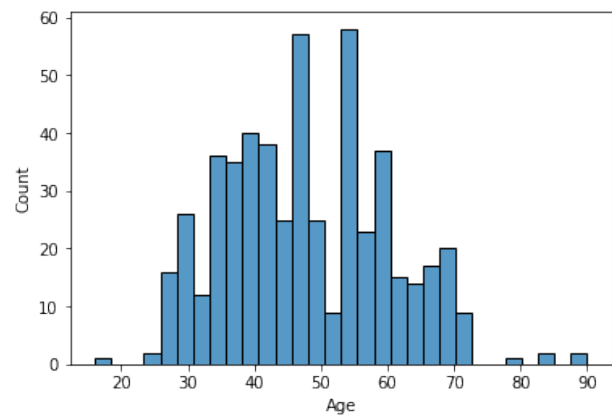


Fig. 3: Patient Age Distribution

III. ANALYSIS AND DISCUSSION

Filter and wrapper feature selection methods were embedded to increase the efficiency and accuracy of the classification models. Diabetes data (see Section 2.1) were used for the purposes of this study.

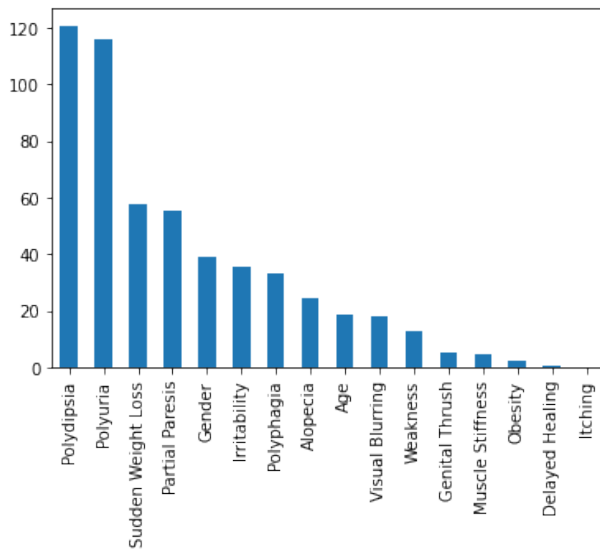


Fig. 4: Feature Score with respect to their contribution towards Class

A. Implementation

Two phases (filter and wrapper) were proposed in the proposed hybrid feature selection approach. In the first phase, correlation matrix with heatmap filter method was used to rank features between the features and target class. A threshold value of correlation heatmap scores was determined to select k-best features. Ideally, features are selected if the correlation score between two features and target class is more than 0.5. However, there are only the features Polyuria and Polydipsia have more than 0.5 correlation score with the target class. Also, most of the features have correlation scores between ~ 0.1 and 0.5. Thus, the threshold value of correlation heatmap scores was determined as ~ 0.1 . Then, 11 features were selected at the end of the first phase and these features are used in the second phase as input. A heatmap of the correlation matrix for the diabetes data is presented in Fig. 2. The correlation between features and target class higher than 0.5 are displayed in dark green which are Polyuria and Polydipsia. These two features play an important role in the detection of diabetes and that the two features show correlation is therefore not surprising.

In the second phase, sequential forward selection (SFS) approach was implemented for selecting the k-best features. The optimal number of features was selected based on feature importance method. It ranks features and gives scores for each feature of the data. Feature importance method is a three-based classifier extracting the k-best number of feature. Finally, SFS wrapper feature selection method was applied to select the optimal feature set. At the end, 7 features were selected over 11 features. The evaluation of the selected optimal features were measured by applying Support Vector Machine, Random Forest and Artificial Neural Network classifiers. Five evaluation matrices were used in the evaluation of the proposed approach such as Accuracy, F-measure, Precision, Recall and AUC. Table II presents the classification results for the proposed

hybrid feature selection approach (both CM + SFS), original feature set, correlation matrix with heatmap (only correlation matrix with heatmap feature selection was applied), SFS (only SFS was applied)

The results obtained from the data in Table II that ANN has the best accuracy rate (99.1%) using 7 features selected based on the proposed hybrid approach while SVM has the lowest accuracy rate (95.8%). Overall, ANN has the best accuracy rate based on not only hybrid method but also Sequential Forward Selection (SFS), Confusion Matrix (CM) and Original (96.5%, 94.6%, 95.6% respectively). As highlighted in Section III, 10-fold cross-validation was used to test the performance of the classification algorithms. As can be seen from Table II that 7 features were selected based on the proposed hybrid approach. These features are set out in Table III.

B. Results

Fig. 5 illustrates the AUC results for SVM, RF and ANN classifiers based on different feature sets. From Fig. 5, it is observed that ANN has the highest AUC rates for each feature set (purple color). Also, The AUC rates of ANN for hybrid method (99.8%) when compared with original features (98.8%), correlation heatmap (99.0%) and SFS (99.6%) methods. Additionally, SVM and RF classifiers achieved to obtain highest AUC rates with hybrid method (98.5%, 99.7% respectively).

C. Comparison of previous studies and the proposed model

Table IV compares the results obtained from the previous studies and the proposed model for prediction of diabetes. The purpose of the studies is the same with our study. Anshuman Guha [21], Ahmed Kareem et al [22], Kezban Alpan [23] and Jingyu Xue [24] used the same data set with the data set used in this study consisting of 520 instances. Ahmed Kareem et al. achieved to obtain 98.8% accuracy applying RF value without applying any feature selection technique. Kezban Alpan (2020) also accomplished 98% accuracy applying kNN with nine selected features. These features were selected based on Gain Ratio feature selection technique. However, there is no detailed information about how these models were fitted in these studies. On the other hand, various diabetes data were used in different studies and there are four studies achieved to reach more than 90% accuracy. For instance, Tapak et al.[25] applied the RF using 2000 diabetes samples and achieved 98.6% accuracy. Also, the study of D. Jashwanth Reddy [26] presented the successful result applying SVM using 6500 samples (98.6%). Furthermore, the studies of Abbas et. al. 2019 and Aishwarya Mujumdar [27] provided higher accuracy value (96.8% and 96% respectively).

IV. CONCLUSION AND FUTURE DIRECTIONS

Early detection of diabetes plays an important role in treatment. Datasets consist of many features and they can be chosen using significant feature selection methods to build a

Classifier	Method	Selected Features	Accuracy	F-measure	Precision	Recall
SVM	Original	16	0.903	0.823	0.823	0.823
	Heatmap	11	0.911	0.856	0.857	0.865
	SFS	9	0.948	0.885	0.886	0.890
	Hybrid	7	0.958	0.921	0.922	0.924
RF	Original	16	0.906	0.906	0.906	0.906
	Heatmap	11	0.925	0.926	0.928	0.925
	SFS	9	0.963	0.963	0.964	0.963
	Hybrid	7	0.988	0.988	0.988	0.988
ANN	Original	16	0.956	0.956	0.957	0.956
	Heatmap	11	0.946	0.946	0.947	0.946
	SFS	9	0.965	0.965	0.965	0.965
	Hybrid	7	0.991	0.991	0.993	0.991

TABLE II: Comparison of Performances of Classifiers

Feature ID	Feature
4	Polydipsia
3	Polyuria
5	Sudden Weight Loss
13	Partial Paresis
11	Irritability
7	Polyphagia
1	Age

TABLE III: Selected features with hybrid approach

Studies	Data Size	Classifier	Accuracy
Anshuman Guha [21]	520	RF	94.8%
Ahmed Kareem et al. [22]	520	RBF	98.8%
Kezban Alpan [23]	520	kNN	98.0%
Jingyu Xue [24]	520	SVM	96.5%
D. Jashwanth Reddy [26]	2000	RF	98.48%
Maniruzzaman et al.[28]	768	LDA,QDA,NB,GPC	81.9%
Deng and Kasabo [29]	768	ESOM	78.4%
Christobel and Sivaprakasam [16]	768	KNN	78.1%
Farahmandian et al. [30]	768	SVM,KNN,NB,ID3,CART,c5.0	81.0%
Khashei et al.[31]	68	LDA,QDA,KNN,SVM,ANN,HPM	80.0%
Nongyao Nai-arun [32]	30112	RF	85.5%
Aishwarya Mujumdar [27]	800	Logistic Regression	96.0%
Tapak et al. [25]	6500	SVM	98.6%
Abbas et al. [33]	1492	SVM-RBF	96.8%
Proposed Approach I	520	Hybrid + ANN	99.1%
Proposed Approach II	520	Hybrid + RF	98.8%

TABLE IV: Results from Previous Studies

well predictive model. The reason for this is that all feature selection methods cannot take into account important features to enhance the predicting process. Note that significant features are used to build well predictive models that help health-care professionals to treat patients. A reasonable approach to tackle this issue could be to describe different feature selection approaches. In this case, a predictive model has been built based on a new hybrid feature selection approach. The proposed hybrid feature selection approach combines filter (correlation matrix with heatmap) and wrapper (sequential forward selection) methods. Initially, 11 features were selected applying correlation matrix with heatmap over 16 features. Then, sequential forward selection (SFS) approach was implemented for selecting the k-best features using these 11 features, and so 7 features were selected through SFS. Three different machine learning algorithms applied to the selected features (feature subsets) and then their performances were compared to reveal the efficiency of the proposed new hybrid feature selection approach on the diabetes detection. This study revealed that the best classification accuracy (99.1%) is

obtained by applying the Artificial Neural Network algorithm with feature set generating from the proposed hybrid feature selection approach. Finally, we can see that machine learning algorithms and the proposed hybrid feature selection approach have made outstanding contributions in the diabetes data. Our future research will be about the combination of feature selection approach and deep learning algorithms to improve classification accuracy based on the image dataset.

REFERENCES

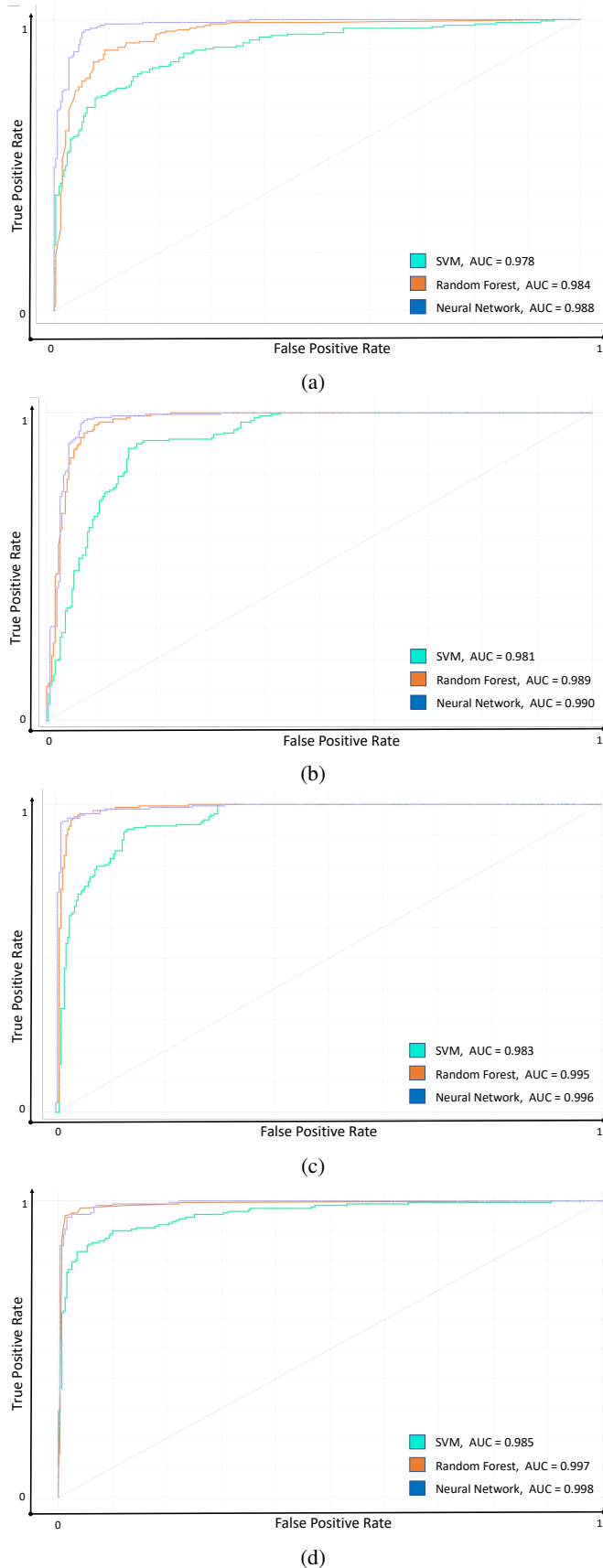


Fig. 5: ROC curve of (a) Original, (b) Heatmap, (c) SFS and (d) Hybrid

- [1] Stephanie Watson, "Everything You Need to Know About Diabetes," 2020. [Online]. Available: <https://www.healthline.com/health/diabetes>
- [2] K. Shailaja, B. Seetharamulu, and M. A. Jabbar, "Machine learning in healthcare: A review," in *2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, 2018, pp. 910–914.
- [3] N. Peiffer-Smadja, T. Rawson, R. Ahmad, A. Buchard, G. Pantelis, F.-X. Lescure, G. Birgand, and A. Holmes, "Machine learning for clinical decision support in infectious diseases: A narrative review of current applications," *Clinical Microbiology and Infection*, vol. 26, 09 2019.
- [4] E. Sevinc, "A novel evolutionary algorithm for data classification problem with extreme learning machines," *IEEE Access*, vol. 7, pp. 122419–122427, 2019.
- [5] K. D. Silva, W. K. Lee, A. Forbes, R. T. Demmer, C. Barton, and J. Enticott, "Use and performance of machine learning models for type 2 diabetes prediction in community settings: A systematic review and meta-analysis," *International Journal of Medical Informatics*, vol. 143, no. August, p. 104268, 2020. [Online]. Available: <https://doi.org/10.1016/j.ijmedinf.2020.104268>
- [6] H. Zheng, H. W. Park, D. Li, K. H. Park, K. H. Ryu, J. Xue, F. Min, F. Ma, N. P. Tigga, S. Garg, Stephanie Watson, I. Kavakiotis, O. Tsave, A. Salifoglou, N. Maglaveras, I. Vlahavas, I. Chouvarda, D. Jashwanth Reddy, B. Mounika, S. Sindhu, T. Pranayteja Reddy, N. Sagar Reddy, G. Jyothsna Sri, K. Swaraja, K. Meenakshi, P. Kora, M. M. F. Islam, R. Ferdousi, S. Rahman, H. Y. Bushra, S. Gupta, A. Guha, D. Jain, V. Singh, A. K. Farahat, A. Ghodsi, M. S. Kamel, J. Chaki, S. Thillai Ganesh, S. K. Cidham, S. Ananda Theertan, V. Bolón-Canedo, N. Sánchez-Maróño, A. Alonso-Betanzos, K. Alpan, G. S. Ilgi, K. Akyol, and B. Şen, "Likelihood Prediction of Diabetes at Early Stage Using Data Mining Techniques," *Knowledge and Information Systems*, vol. 15, no. 3, pp. 113–125, 2020.
- [7] J. Chaki, S. Thillai Ganesh, S. K. Cidham, and S. Ananda Theertan, "Machine learning and artificial intelligence based Diabetes Mellitus detection and self-management: A systematic review," *Journal of King Saud University - Computer and Information Sciences*, 2020. [Online]. Available: <https://doi.org/10.1016/j.jksuci.2020.06.013>
- [8] I. Kavakiotis, O. Tsave, A. Salifoglou, N. Maglaveras, I. Vlahavas, and I. Chouvarda, "Machine Learning and Data Mining Methods in Diabetes Research," *Computational and Structural Biotechnology Journal*, vol. 15, pp. 104–116, 2017. [Online]. Available: <https://doi.org/10.1016/j.csbj.2016.12.005>
- [9] D. Jashwanth Reddy, B. Mounika, S. Sindhu, T. Pranayteja Reddy, N. Sagar Reddy, G. Jyothsna Sri, K. Swaraja, K. Meenakshi, and P. Kora, "Predictive machine learning model for early detection and analysis of diabetes," *Materials Today: Proceedings*, 2020. [Online]. Available: <https://doi.org/10.1016/j.matpr.2020.09.522>
- [10] H. Lai, H. Huang, K. Keshavjee, A. Guergachi, and X. Gao, "Predictive models for diabetes mellitus using machine learning techniques," *BMC Endocrine Disorders*, vol. 19, no. 1, pp. 1–9, 2019.
- [11] N. Nai-Arun and R. Mounghmai, "Comparison of Classifiers for the Risk of Diabetes Prediction," *Procedia Computer Science*, vol. 69, pp. 132–142, 2015. [Online]. Available: <http://dx.doi.org/10.1016/j.procs.2015.10.014>
- [12] Kaggle, "Pima Indians Diabetes Dataset," 2021. [Online]. Available: <https://www.kaggle.com/uciml/pima-indians-diabetes-database>
- [13] G. Swapna, R. Vinayakumar, and K. P. Soman, "Diabetes detection using deep learning algorithms," *ICT Express*, vol. 4, no. 4, pp. 243–246, 2018. [Online]. Available: <https://doi.org/10.1016/j.icte.2018.10.005>
- [14] Q. Zou, K. Qu, Y. Luo, D. Yin, Y. Ju, and H. Tang, "Predicting diabetes mellitus with machine learning techniques," *Frontiers in Genetics*, vol. 9, Nov. 2018. [Online]. Available: <https://doi.org/10.3389/fgene.2018.00515>
- [15] S. Pratama, A. Muda, Y.-H. Choo, and N. Muda, "Computationally inexpensive sequential forward floating selection for acquiring significant features for authorship invarianceness in writer identification," *International Journal of New Computer Architectures and their Applications (IJNCAA)*, vol. 1, pp. 581–598, 01 2011.
- [16] Y. A. Christobel and P. Sivaprakasam, "A New Classwise k Nearest Neighbor (CKNN) Method for the Classification of Diabetes Dataset," *International Journal of Engineering and Advanced Technology*, vol. 2, no. 3, pp. 396–400, 2013.
- [17] Wikipedia, "Support vector machine," 2021. [Online]. Available: https://en.wikipedia.org/wiki/Support_vector_machine
- [18] —, "Random Forest," 2021. [Online]. Available: https://en.wikipedia.org/wiki/Random_forest

- [19] —, “Artificial Neural Network,” 2021. [Online]. Available: https://en.wikipedia.org/wiki/Artificial_neural_network
- [20] —, “Precision and Recall,” 2021. [Online]. Available: https://en.wikipedia.org/wiki/Precision_and_recall
- [21] A. Guha, “Building Explainable and Interpretable model for Diabetes Risk Prediction,” *International Journal of Engineering Research and Technology*, vol. 9, no. 09, pp. 1037–1042, 2020.
- [22] A. Kareem, L. Shi, L. Wei, and Y. Tao, “A Comparative Analysis and Risk Prediction of Diabetes at Early Stage using Machine Learning Approach A Comparative Analysis and Risk Prediction of Diabetes at Early Stage using Machine Learning Approach,” *International Journal of Future Generation Communication and Networking*, vol. 13, no. 3, pp. 4151–4163, 2020.
- [23] K. Alpan and G. S. Ilgi, “Classification of Diabetes Dataset with Data Mining Techniques by Using WEKA Approach,” in *2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*. IEEE, oct 2020, pp. 1–7.
- [24] J. Xue, F. Min, and F. Ma, “Research on diabetes prediction method based on machine learning,” *Journal of Physics: Conference Series*, vol. 1684, no. 1, 2020.
- [25] L. Tapak, H. Mahjub, O. Hamidi, and J. Poorolajal, “Real-data comparison of data mining methods in prediction of diabetes in iran,” *Healthcare Informatics Research*, vol. 19, no. 3, p. 177, 2013.
- [26] D. Reddy, B. Mounika, S. Sindhu, T. Reddy, N. Reddy, G. Sri, K. Swaraja, M. Kollati, and P. Kora, “Predictive machine learning model for early detection and analysis of diabetes,” *Materials Today: Proceedings*, 10 2020.
- [27] A. Mujumdar and V. Vaidehi, “Diabetes Prediction using Machine Learning Algorithms,” *Procedia Computer Science*, vol. 165, pp. 292–299, 2019. [Online]. Available: <https://doi.org/10.1016/j.procs.2020.01.047>
- [28] M. Maniruzzaman, M. J. Rahman, B. Ahammed, and M. M. Abedin, “Classification and prediction of diabetes disease using machine learning paradigm,” *Health Information Science and Systems*, vol. 8, no. 1, Jan. 2020.
- [29] D. Deng and N. Kasabov, “On-line pattern analysis by evolving self-organizing maps,” *Neurocomputing*, vol. 51, pp. 87–103, apr 2003.
- [30] M. Farahmandian, Y. Lotfi, and I. Maleki, “Data Mining Algorithms Application in Diabetes Diseases Diagnosis : A Case Study,” *MAGNT Research Report*, vol. 3, no. 1, pp. 989–997, 2015.
- [31] M. Khashei, S. Eftekhari, and J. Parvzian, “Diagnosing diabetes type ii using a soft intelligent binary classification model,” *Review of Bioinformatics and Biometrics*, vol. 1, no. 1, pp. 9–23, 2012.
- [32] N. Nai-arun and R. Mounghmai, “Comparison of classifiers for the risk of diabetes prediction,” *Procedia Computer Science*, vol. 69, pp. 132–142, 2015.
- [33] H. T. Abbas, L. Alic, M. Erraguntla, J. X. Ji, M. Abdul-Ghani, Q. H. Abbasi, and M. K. Qaraqe, “Predicting long-term type 2 diabetes with support vector machine using oral glucose tolerance test,” *PLOS ONE*, vol. 14, no. 12, p. e0219636, Dec. 2019.



Ayhan Akbas received his B.S. and M.S. degrees from Middle East Technical University in Electrical and Electronics Engineering in 1991 and 1995, Ph.D degree from Computer Engineering from TOBB ETU in 2016, respectively. He is currently working as assistant professor at Computer Engineering Dept. in Abdullah Gul University. His areas of interest are Wireless Sensor Networks, IoT, Wireless Communication, and machine learning.

BIOGRAPHIES



Selim Buyrukoglu is an Asst. Prof. in Computer Engineering Department at Cankiri Karatekin University. He completed his PhD at Loughborough University, UK in 2019, He received an MSc degree in Advance Computer Science in 2014 from Leicester University, UK and also a BSc degree in Computer Engineering in 2010 from European University of Lefke, TRNC. He is particularly focusing on the application of Machine Learning and Deep Learning methods on interdisciplinary subjects.

Distributed Formation Control Algorithm for Improved String Stability in Heterogenic Vehicle Platoons

Hakan Sert and Serhat İkizoğlu

Abstract— Achieving speed, safety and energy optimization in vehicle platoons is among the important topics of recent years. In this context, this work focuses on improving the string stability (SS) of heterogeneous platoons. Better SS allows for smaller gap between vehicles, which means shorter time headway for Cooperative Adaptive Cruise Control (CACC) or inter-vehicle following-distance for platooning. Shorter time headway or inter-vehicle distance results in better road use and less fuel consumption. Rather than compensating for dynamic differences by means of low level control schemes by implementing pre-compensators at vehicle level, this work takes a different approach, where a higher level of control is preferred to improve SS: The platoon formation. To achieve this, it is important to ensure that each vehicle is in its optimum position in the platoon based on its dynamics for highest string stability. This generally leads to the vehicle with highest inertia to lead the platoon and the lowest one as the last follower. In this study, a platoon formation algorithm is proposed to run on each vehicle, leading to an optimum overall string structure. The efficiency of the algorithm is demonstrated by simulations in Matlab(R) for a four-vehicle platoon.

Index Terms— Cruise control, distributed formation, heterogenic vehicle platoon, string stability.

I. INTRODUCTION

IN RECENT years, traffic congestion and energy saving have been the utmost major issues in road transportation. Vehicle platooning has already proven to have potential to improve both. The basic idea in platooning is to have a string of vehicles with closer gap than normal driving by applying automation. In this sense, automation in its simplest form is a cooperative adaptive cruise control (CACC) in which a following-strategy is applied for a constant time gap. While

adaptive cruise control is based on measuring distance with the vehicle ahead using a forward-facing radar, CACC also requires vehicle to vehicle (V2V) communication, which is done via wireless radio frequency (RF) communication, the IEEE 802.11p protocol over 5.9GHz base frequency.


Benefits of platoons come up from its vehicle-following strategy. Platooning differs from CACC in that it has a constant-distance following strategy rather than a constant time gap. Less inter-vehicle gap means more vehicles per kilometer, which leads to better use of the road and less traffic jams. It also improves aerodynamics of the whole platoon reducing fuel cost and emission in parallel. It can also be said that the platoon aids safety since the reaction time of automation is much less than that of the human driver.

Besides the opportunities described above, there are some challenges with vehicle platoons. One of the major challenges is to ensure the string stability (SS). There are various definitions for SS in the literature [1]. In this study we define SS as the fact that the error for the distance between two consecutive vehicles remains positive as we move along the string. Here, the error is defined as the difference between the actual distance and the value calculated according to the predefined equation. SS ensures that there will be no collision between the members of the platoon or no split under acceleration. Thus, not only each vehicle needs to exhibit stability individually, but also the entire platoon needs to be stable as well.

Several studies are carried out in the literature regarding the string stability. In [2], a model reference adaptive control augmentation is proposed in which the control objective is to augment a baseline CACC, proven to be string stable in the homogeneous scenario, with an adaptive control term that compensates for each vehicle's unknown driveline dynamics. In [3] a constant time headway policy is employed for a platoon-stable controller design based on feedback linearization. In [4], a controller design method is developed that allows the string stability requirement to be explicitly included in the controller synthesis specifications. [5] presents a distributed finite-time adaptive integral-sliding-mode (ISM) control approach. In [6], novel criteria for string stability are proposed for mixed traffic platoons that consists of both automated and manual driving cars. A mixed traffic string stability definition is proposed to guarantee the boundedness of the motion states fluctuation upstream as well as the safety

HAKAN SERT, is with Department of Control and Automation Engineering, Istanbul Technical University, Istanbul, Turkey, (e-mail: hsert@itu.edu.tr).

SERHAT İKİZOĞLU, is with Department of Control and Automation Engineering, Istanbul Technical University, Istanbul, Turkey (e-mail: ikizoglus@itu.edu.tr).

 <https://orcid.org/0000-0003-2394-7988>

Manuscript received September 13, 2021; accepted Feb 16, 2022.
DOI: 10.17694/bajece.1009213

of the entire platoon. [7] presents a periodic switching control method for an automated vehicle platoon to minimize the overall fuel consumption. Considering the nonlinearity of switching operation, the concept of bounded stability for a platoon of vehicles is defined to replace conventional internal stability and string stability. The distributed servo loop controller based on dual-pulse-and-glide operation is designed for each vehicle, wherein a sectionalized switching map is adopted for proper mode selection. [8] investigates the nonlinear vehicle platoon control problems with external disturbances. The authors apply the quadratic spacing policy (QSP) into the platoon control, in which the desired inter-vehicle distance is a quadratic function in terms of the vehicle's velocities. They propose a platoon control scheme based on the distributed integrated sliding mode (DISM) and they also perform a string stability analysis accordingly. In [9] the authors propose a centralized collision-free solution on the basis of model predictive control that guarantees asymptotic platoon tracking of speed changes and satisfaction of system constraints during the transient process. In [10] the authors investigate the impact of applying a constant distance gap (CDG) policy for starting platoons at traffic lights, where they claim that the applicability of CDG in real traffic is limited, due to its demand on complex communication topologies in order to achieve string stability. However, they explain that they were able to show its capability to increase the capacity of traffic light controlled intersections.

This work focuses on improving SS of a heterogeneous platoon where each member of the platoon has different dynamics in contrast to a homogenous platoon where each vehicle is considered to present the same dynamic behavior. Improving SS allows for smaller gap between the vehicles which means shorter time headway for CACC. Shorter time headway or inter-vehicle distance results in better road use and less fuel consumption. Ideally, this leads to road trains (touch condition) which is practically not possible due to marginal SS behavior under physical constraints.

Rather than compensating for dynamic differences by means of low level control schemes by implementing pre-compensators, this study takes a different approach to improve SS at a higher level control: The platoon formation. We try to ensure that each vehicle in the platoon is in its optimum position based on its dynamics for highest string stability. Considering the vehicle dynamics to be proportional to engine dynamics, it results that the vehicle with the highest inertia takes the leadership of the platoon and the one with the lowest is the last follower. A platoon formation algorithm is developed and run on each vehicle leading to an optimum overall string structure.

In line with the explanations above, the rest of this paper is arranged as follows: Section 2 focuses on vehicle model with CACC in a platoon and drive cycle following a brief information about SS. Section 3 covers the proposed algorithm and Section 4 provides the simulation results. In Section 5 we draw conclusions along with future steps of the study.

II. VEHICLE MODEL WITH CACC

II.1 String stability

If we consider the changes in vehicle dynamics (acceleration, speed or position) as noise, the string stability can be described as the condition that this noise attenuates as it propagates along the vehicle string [11]. String stability is concerned with the interactions in a platoon such that vehicles must not collide with each other [12].

II.2 Mathematical background

Figure 1 illustrates a heterogeneous vehicle string, where inter-vehicle communication is considered to be established over wireless network. Each vehicle receives data from the vehicle in front and sends its own to the one behind over the wireless communication channel.

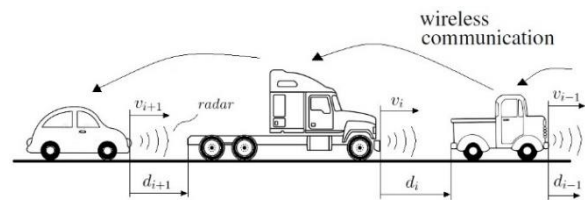


Fig. 1. CACC-equipped heterogeneous vehicle platoon [1]

In a vehicle string, the desired following-distance is given as:

$$d_{r,i}(t) = r_i + h v_i(t), \quad i \in S_m \quad (1)$$

where h is the following-time gap in seconds and r_i is the desired distance in meters between the vehicles at full stop condition; v_i is the speed of the i th vehicle in m/s. $S_m = \{i \in \mathbb{N} \mid 1 \leq i \leq m\}$ is the set of all vehicles in a platoon of length (number of vehicles) $m \in \mathbb{N}$. $i = 0$ is reserved for the platoon leader. The inter-vehicle distance decreases as the platoon speed decreases and it increases with an increase in the platoon speed [13]

The following-distance error between two vehicles travelling consecutively is given as:

$$\begin{aligned} e_i(t) &= d_i(t) - d_{r,i}(t) \\ &= (q_{i-1}(t) - q_i(t) - L_i) - (r_i + h v_i(t)) \end{aligned} \quad (2)$$

Here, q_i and q_{i-1} are the rear-bumper positions of vehicles in the rear and front, respectively; L_i stands for the vehicle length. Nonlinear model for the i th vehicle in a platoon is given as [13]:

$$\begin{aligned} \dot{d}_i &= v_{i-1} - v_i \\ \dot{v}_i &= a_i \\ \dot{a}_i &= f_i(v_i, a_i) + g_i(v_i) \eta_i \end{aligned} \quad (3)$$

where v_i and a_i stand for the velocity and acceleration of the i th vehicle, respectively and η_i represents the engine input. The functions f_i and g_i are given by the following equations:

$$f_i(v_i, a_i) = -\frac{2K_{di}}{m_i} v_i a_i - \frac{1}{\tau_i(v_i)} \left[a_i + \frac{K_{di}}{m_i} v_i^2 + \frac{d_{mi}}{m_i} \right] \quad (4)$$

$$g_i(v_i) = \frac{1}{m_i \tau_i(v_i)}$$

where m_i is the vehicle mass, τ_i is the time constant of its engine, K_{di} is the aerodynamic drag coefficient and d_{mi} is the mechanical drag force [13]. In order to obtain a control system structure, we define

$$\eta_i = m_i u_i + K_{di} v_i^2 + d_{mi} + 2\tau_i K_{di} v_i a_i \quad (5)$$

where u_i represents the external input signal of the i th vehicle. Substituting (5) and (4) in (3) and assuming that τ_i is constant, the resulting linearized vehicle model will be

$$\begin{pmatrix} \dot{a}_i \\ \dot{v}_i \\ \dot{u}_i \end{pmatrix} = \begin{pmatrix} v_{i-1} - v_i \\ a_i \\ -\frac{1}{\tau_i} a_i + \frac{1}{\tau_i} u_i \end{pmatrix}, i \in S_m \quad (6)$$

This model is actually derived by applying pre-compensator to a more complex model by input-output linearization with state feedback [11]. To reach a state space model of the followers in the platoon, a new control input ξ_i is defined as follows [2], [14]:

$$h\dot{u}_i = -u_i + \xi_i \quad (7)$$

$$\xi_i = \begin{cases} K_p e_i + K_d \dot{e}_i + u_{i-1}, & \forall i \in S_m \\ u_r, & i = 0 \end{cases} \quad (8)$$

Here, u_r stands for the input corresponding to the desired acceleration of the leader; the input u_{i-1} is received over the wireless communication between the vehicles i and $i-1$. Finally, string stable reference dynamics for the heterogeneous platoon is obtained as [2], [14]

$$\begin{pmatrix} \dot{e}_i \\ \dot{v}_i \\ \dot{a}_i \\ \dot{u}_i \end{pmatrix} = \begin{pmatrix} 0 & -1 & -h & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1/\tau_i & 1/\tau_i \\ K_p/h & -K_d/h & -K_d & -1/h \end{pmatrix} \begin{pmatrix} e_i \\ v_i \\ a_i \\ u_i \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ K_d/h & 1/h \end{pmatrix} \begin{pmatrix} v_{i-1} \\ u_{i-1} \end{pmatrix} \quad (9)$$

Here, τ_i is the time constant representing the driveline dynamics of the i th vehicle in seconds. Equation (9) can be generalized as;

$$\dot{x}_i = A_i x_i + B w_i, \forall i \in S_m \quad (10)$$

where x_i is the i th vehicle's state vector and w_i represents the external input vector. Since the leading vehicle follows the reference signal only, its state space model will be:

$$\begin{pmatrix} \dot{e}_0 \\ \dot{v}_0 \\ \dot{a}_0 \\ \dot{u}_0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1/\tau_0 & 1/\tau_0 \\ 0 & 0 & 0 & -1/h \end{pmatrix} \begin{pmatrix} e_0 \\ v_0 \\ a_0 \\ u_0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1/h \end{pmatrix} u_r \quad (11)$$

II.3 Simulation example

Based on Eq. (10), different definitions are proposed in the literature for SS and stability analyses are performed accordingly [2, 14]. However, the common attitude is that the stability is studied by observing the change in distance error, velocity and/or acceleration. At this point we recall that we define SS as the condition that the absolute change in the abovementioned dynamics attenuates as it propagates along the vehicle string and relate the heterogeneity of the string to different time constants of vehicle engines in the platoon. As a novelty of this study, we do not search for controllers to compensate for dynamic differences between the vehicles for SS; instead, we arrange the platoon formation for optimum stability. That is, the vehicle with the highest inertia leads the platoon, whereas the one with the lowest inertia is the last follower.

We formed a sample heterogeneous 4-vehicle platoon where we determined the time constants of the members as given in Table I. The number of the vehicles in the string was limited to 4 for the sake of simplicity and the fact that longer strings would not be practical [12].

TABLE I
TIME CONSTANTS (τ_i [s]) CHOSEN FOR A SAMPLE 4-VEHICLE HETEROGENEOUS PLATOON

i	0	1	2	3
τ_i	0,1	0,2	0,4	0,5

A 60-seconds drive cycle was created from stand-still, acceleration, deceleration to full-stop as shown in Figure 2.

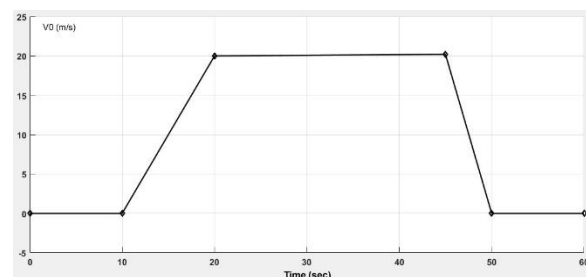


Fig. 2. Sample drive cycle

III. THE PLATOON FORMATION ALGORITHM

We developed a formation algorithm which takes the dynamics of all vehicles in the string into consideration. For this purpose, the state-flow method was used in Matlab(R) environment. The algorithm includes the operation modes, interfaces, transition rules and actions described below. The corresponding flowchart for a single vehicle is presented in Fig. 3.

III.1 Operation Modes

In a 4-vehicle platoon, a vehicle can only be in one of the following operation modes at a time:

1. Manual Drive (MD)
2. Free Agent (FA)
3. Platoon Leader (PL)
4. Follower-1 (F1)
5. Follower-2 (F2)
6. Follower-3 (F3)

Possible transitions between modes of operation are listed in Table II, where rows and columns indicate the modes before and after the transition, respectively.

TABLE II
OPERATION MODES AND TRANSITIONS

	MD	FA	PL	F1	F2	F3
MD		X				
FA	X		X	X		X
PL		X		X		
F1		X	X		X	
F2		X		X		X
F3		X			X	

III.2 Interfaces

Event triggers are:

1. Brake pedal
2. Accelerator pedal
3. Manual steering
4. Vehicle failure

System inputs are:

1. Enable switch
2. Join acknowledge
3. Split acknowledge

System outputs are:

1. Join request
2. Split request

III.3 Transition rules

The transition rules are determined as follows:

From manual drive to free agent mode:

1. No failure
2. Vehicle speed above 30km/h
3. Driver should press enable switch

From free agent mode to leader position:

1. Join request should be accepted (either sent or received over V2V)
2. Should have the smallest time constant among the platoon members

From free agent to the *i*th follower:

1. Join request should be accepted (either sent or received over V2V)
2. Must have a time constant greater than that of the (i-1)th vehicle and smaller than that of the (i+1)th vehicle.

III.4 Actions

Manual drive:

1. Stop V2V broadcasting,
2. Switch to free agent when conditions are fulfilled.

Free agent:

1. Start V2V broadcasting,
2. Switch to manual drive when conditions are fulfilled,
3. Switch to leader when conditions are fulfilled,
4. Switch to follower when conditions are fulfilled.

Follower vehicle:

1. Keep the following distance,
 2. Split when needed,
 3. Switch to leader when conditions are fulfilled,
- Switch to manual drive when conditions are fulfilled.

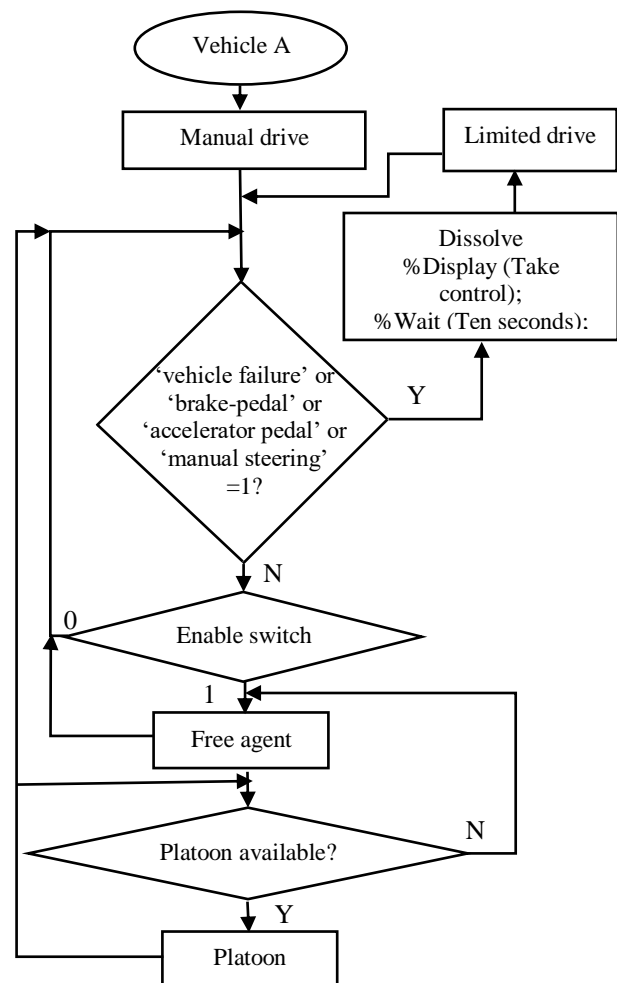


Fig. 3. Flow chart for a single vehicle (e.g. vehicle A)

Leader:

1. Broadcast platoon ID,
 2. Broadcast target gap,
 3. Switch to follower when conditions are fulfilled,
- Dissolve the platoon when conditions are fulfilled.

Fig. 4 illustrates the join to platoon based on vehicle dynamics.

IV. SIMULATION RESULTS

In this section, we demonstrate how the proposed algorithm improves the string stability by considering the vehicle dynamics in contrast to a platoon disregarding this fact. Four vehicles with different dynamics listed in Table 1 are run in the simulation environment, each time arranging them in a different order to observe the difference in the string stability behavior of the platoon. The error signal is considered as the difference between the reference following-distance (determined as 5 meters in this study) and the actual one.

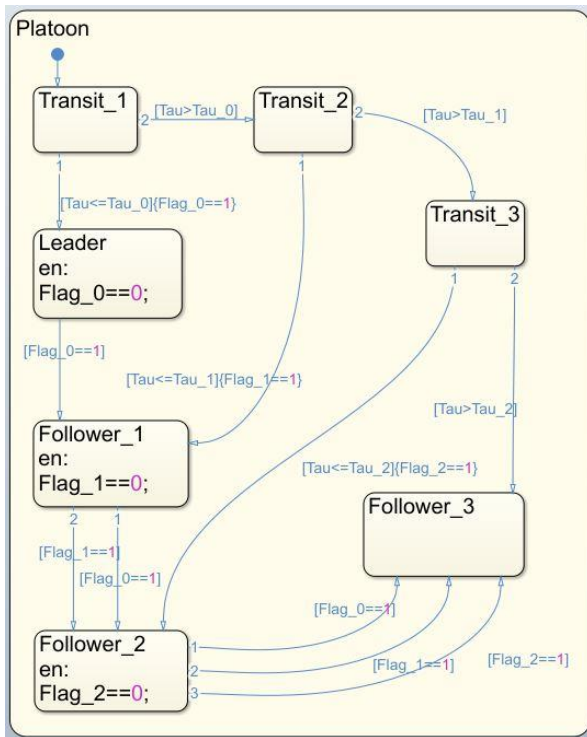


Fig. 4. Join to platoon based on vehicle dynamics (Tau: τ)

According to the definition of the string stability, the error signal should attenuate as it propagates along the vehicle string. We note that only the positive values of the error signal have to be taken into consideration, since this corresponds to the condition that vehicles are getting closer than allowed. Thus, these are the only values that have potential to reduce the driving safety of the vehicles in the platoon.

The simulation results are summarized in Figure 5. The stability of the string is analyzed based on the error propagation results. Here, the inspection is focused on how the following-distance changes as we move through the platoon. As an example, for scenario A, error signals remain positive as

Scenario	Vi	Min. Following Distance between V1-V2 (m)		Min. Following Distance between V2-V3 (m)		Min. Following Distance between V3-V4 (m)		String Stable?
		V1	V2	V2	V3	V3	V4	
A	1/τi	10	5	5	2,5	2,5	2	String Stable? YES
B	1/τi	10	2	2	2,5	2,5	5	String Stable? No
C	1/τi	2	2,5	2,5	5	5	10	String Stable? No
D	1/τi	2	10	10	5	5	2,5	String Stable? Yes

Fig. 5. Minimum following-distances and string stability for different scenarios

we move through the string. The minimum following-distances have approximately the same value (5.1m and 5.2m) being greater than the reference value set at 5m. Thus, we can conclude that the platoon in scenario A is string-stable. The vehicle following-distances for this scenario are depicted in Figure 6.

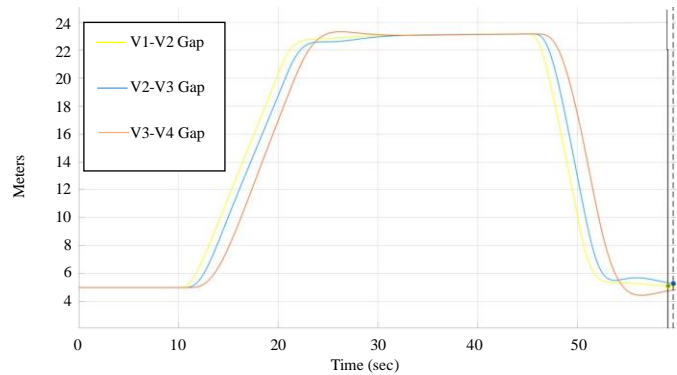


Fig. 6. Following-distances of the vehicles in scenario A

In Figure 7, vehicle following-distances are shown for the platoon in scenario B. The error signal goes to negative as we move through the string (e_i : 5.2-5=0.2m, 4.5-5=-0.5m, 4.4-5=-0.6m, respectively); that is, the following-distance decreases when we move towards the back of the platoon. Thus, we can say that the platoon in scenario B is not string-stable.

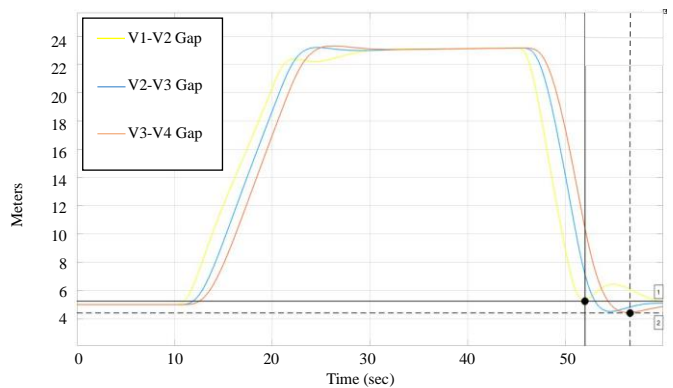


Fig. 7. Following-distances of the vehicles in scenario B

Figure 8 presents the graphics for the vehicle following-distances for the platoon in scenario C. Though the following-distance between the third couple of vehicles (V3&V4) is greater than that between the second couple (V2&V3), the error signal is negative along the whole string. Thus, we can disclose that the platoon in scenario C is not string stable.

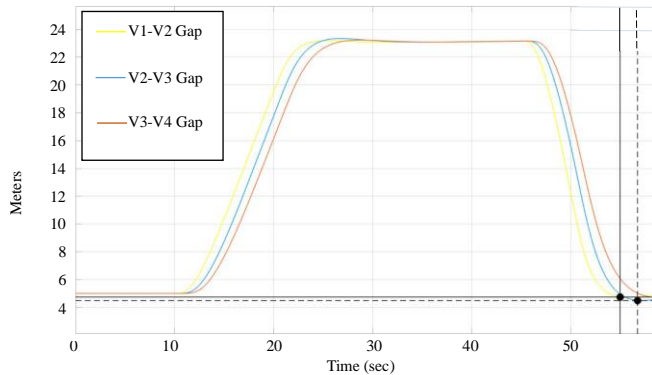


Fig. 8. Following-distances of the vehicles in scenario C

Finally, Figure 9 illustrates the vehicle following-distances for the platoon in scenario D. The error signal being negative between the leader and its follower ($3.9-5=-1.1\text{m}$) turns to positive as we move through the platoon. Thus, we can conclude that the platoon in scenario D is string stable.

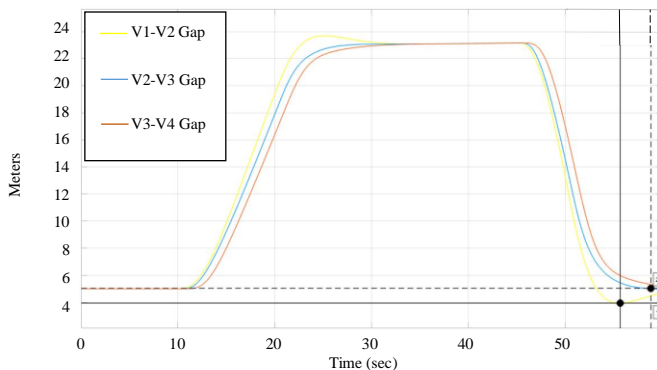


Fig. 9. Following-distances of the vehicles in scenario D

V. CONCLUSIONS

This work has focused on improving string stability of heterogeneous platoons by implementing a new algorithm: The platoon formation. This algorithm takes the dynamics of the platoon members into consideration and decides on the position of each member regarding its inertia in contrast to algorithms in the literature, which apply complex control schemes by implementing pre-compensators at vehicle level. The proposed algorithm is studied under different platoon scenarios. Based on the simulation results, we can conclude that when constructing a platoon, taking each member's dynamics into consideration helps to improve the string stability. Obviously, it is possible to compensate the differences of each vehicle's longitudinal dynamics by implementing pre-filters to longitudinal controllers of the

vehicles, but that would limit the overall performance of the platoon. This method is considered to be an alternative to low level adaptive controllers. Even for the string stable platoons, the proposed formation algorithm has potential to further decrease the following-distances which results in improvement of overall fuel economy and road usage. Exchanging vehicle information (load, power, brake performance etc.) between the vehicle members is a part of draft platoon standards, which makes the algorithm possible for real world applications.

The next step as the future work is that the algorithm will be extended to cover the determination of the correct time that the free agent will join the platoon.

REFERENCES

- [1] Shuo Fenga , Yi Zhanga , Shengbo Eben Li, Zhong Cao, Henry X. Liuc , Li Li, "String stability for vehicular platoon control: Definitions and analysis methods", *Annual Reviews in Control*, 47, pp 81-97, 2019
- [2] Harfouch, Y.A., Yuan, S., and Baldi, S., Adaptive Control of Interconnected Networked Systems with Application to Heterogeneous Platooning, 2017 13th IEEE International Conference on Control and Automation (ICCA) July 3-6, 2017. Ohrid, Macedonia
- [3] Sun, M., Lewis, F.L., and Ge, S.S. Platoon-Stable
- [4] Adaptive Controller Design, 43rd IEEE Conference on Decision and Control December 14-17, 2004 Atlantis, Paradise Island, Bahamas
- [5] Ploeg, J., Shukla, D.P., Wouw, N., and Nijmeijer, H. Controller Synthesis for String Stability of Vehicle Platoons, *IEEE Transactions On Intelligent Transportation Systems*, Vol. 15, N. 2, April 2014
- [6] Guo, X., Wang, J., Liao, F., and Teo, R.S.H. Distributed Adaptive Integrated-Sliding-Mode Controller Synthesis for String Stability of Vehicle Platoons, *IEEE Transactions On Intelligent Transportation Systems*, VOL. 17, NO. 9, September 2016
- [7] Li, F. and Wang, Y. Cooperative Adaptive Cruise Control for String Stable Mixed Traffic: Benchmark and Human-Centered Design, *IEEE Transactions On Intelligent Transportation Systems*, VOL. 18, NO. 12, December 2017
- [8] Li, S.E., Li, R., Wang, J., Hu, X., Cheng, B., and Li, K., Stabilizing Periodic Control of Automated Vehicle Platoon with Minimized Fuel Consumption, *IEEE Transactions On Transportation Electrification*, Vol. 3, No. 1, March 2017
- [9] Lei Zuo, Ye Zhang, Maode Yan and Wenrui Ma, Distributed Integrated Sliding Mode-Based Nonlinear Vehicle Platoon Control with Quadratic Spacing Policy, *Complexity*, 2020.
- [10] Huarong Zheng; Jun Wu; Weimin Wu; Rudy R. Negenborn, Cooperative distributed predictive control for collision-free vehicle platoons, *IET Intelligent Transport Systems*, Volume 13, Issue 5, 2018.
- [11] Kay Massow, Ilja Radusch, and Robert Shorten, On Constant Distance Spacing Policies for Cooperative Adaptive Cruise Control, arXiv:1909.10913v2, Cornell University, 2020.
- [12] Jeroen Ploeg, Nathan van de Wouw, and Henk Nijmeijer String Stability of Vehicle Platoons, *Book of Abstracts of the 30Th Benelux Meeting On Systems and Control*, 2014
- [13] Eyre J., Yanakiev, D. and Kanellakopoulos, I.A Simplified Framework for String Stability Analysis of Automated Vehicles, *Journal of Vehicle System Dynamics*, 30:5, 375-405, 1998
- [14] Stankovic, S.S., Stanojevic, M.J. and Siljak, D.D., Decentralized Overlapping Control of a Platoon of Vehicles, *IEEE Transactions On Control Systems Technology*, Vol.8, No.5, 2000
- [15] Ploeg, J., Wouw, N. and Nijmeijer, H. Lp String Stability of Cascaded Systems, *IEEE Transactions On Control Systems Technology*, Vol. 22, No. 2, March 2014

BIOGRAPHIES



HAKAN SERT received his B.Sc. degree in Electrical Engineering from Yıldız Technical University, Istanbul, in 2003 and the M.Sc. degree in Control and Automation Engineering from Istanbul Technical University in 2018. He is currently with Ford Otosan, Istanbul, Turkey. His research interests are

intelligent transport systems, vehicle power distribution and control systems, vehicle network systems.



SERHAT İKIZOĞLU graduated from Istanbul Technical University (ITU)-Control and Computer Engineering. He completed his doctorate at the same institute in 1992. He is currently Assoc. Prof. in Control and Automation Eng. Dept. at Istanbul Technical University. His areas of interest cover measurement and

instrumentation, control, mechatronics and biomedical.

A Dual-Radio Hybrid Mesh Topology for Multi-Hop Industrial IoT Networks in Harsh Environments

Hasari Celebi

Abstract—Industrial Internet of things (IIoT) is a paradigm that changes the way people interact with infrastructures by enabling ubiquitous connection to the Internet. It allows to design connected infrastructures in industrial environments, e.g., factories, in order to support innovative services and improve efficiency. While the advantages of IIoT are numerous, it has some fundamental issues such as propagation of signal through metallic obstacles, i.e., walls in industrial environments. To address this issue, we propose a hybrid wired-wireless mesh node and network topology called hybrid mesh network (HMN), enabling the signal to cross over the metallic obstacles in harsh environments. The proposed hybrid mesh node utilizes dual-radio feature of IEEE 802.11 Wi-Fi standard to improve the performance of the proposed HMN further considering multi-hop communications. The effects of the packet size and different fading channels on the performance of the proposed HMN are investigated through various simulations. In addition, the performance of the proposed HMN is compared with that of conventional wireless mesh network (WMN). The results reveal that the proposed HMN outperforms conventional WMN and the proposed topology is promising for the implementation of high performance wireless mesh networks in harsh environments.

Index Terms—Wireless mesh networks, harsh channel environments, multi-hop communications, industrial internet of things (IIoT).

I. INTRODUCTION

WITH the recent explosive interests in the Internet of Things (IoT), the researchers have started to focus on wireless connectivity. Various wireless connectivity options are available, especially for Industrial IoT (IIoT) applications [1]. These options include IEEE 802.15.4 based technologies (e.g., Zigbee, 6LoWPAN), Wi-Fi, and Bluetooth Low Energy (BLE) [2], [3]. Among these options, Zigbee [4], 6LoWPAN [5] and Wi-Fi are promising technologies due to their supports of wireless mesh network (WMN) topology.

IEEE 802.11s standard specify the WMN topology for Wi-Fi networks [6]. This technology enables clients to support multiple connections. In other words, it is not necessary that clients connect to only one of the other clients on the network, but also communicate with multiple clients simultaneously. Even if any of the nodes become inoperable, the other nodes can communicate directly with each other or indirectly through each other. Therefore, WMNs are more practical and robust than other types of wireless networks.

Note that the WMNs can consist of mesh points, mesh routers, mesh gateways and mesh clients depending on the type

of WMN topology. The characteristic features of WMNs are dynamic configuration and ability of self organization, quick detection of link errors, error-tolerant connection and backup option on the network, expandable coverage with multi-hop architecture [7]. Moreover, the performance degradation of network, use of management protocols, applicability of different types of wireless networks and mesh topology are corresponding to the scalability, ease of use, collaborative and accessibility parameters respectively as the critical design requirements for WMNs.

One of the main challenges of WMNs in the harsh environments, such as industrial areas (e.g. factories), is covering large areas using closely spaced mesh nodes, e.g. IIoT devices. Especially, the penetration of the signal through metallic walls and obstacles in industrial environments is critical for achieving ubiquitous network coverage [8]. Hence, we propose a hybrid mesh network (HMN) topology, which is a combination of wireless and wired mesh networking, and it allows the signals to cross-over the metallic walls through cables in harsh environments.

The proposed topology in this paper involves the usage of a hybrid mesh node, which encompasses both wired and wireless connections similar to [9]. The hybrid mesh node in simple term is a single node that can be potentially used at the walls or any other separation, but supporting both wireless and wired mesh networks. The hybrid mesh node has a wireless interface on one side and Ethernet cable interface on the other side. The use of two hybrid mesh nodes across the wall or any other harsh physical separation serve the purpose. This primarily happens by converting a wireless mesh network signal to an Ethernet signal at one hybrid node and reverse process at the other hybrid node.

Considering its easy and fast deployment with existing infrastructures [10] together with its high throughput and low latency advantages [11], WMN is a good candidate for harsh environmental conditions, such as post disaster communication [12], emergency crisis in smart cities [10], and industrial automation environment [13]. On the other hand, IoT is also one of the main technologies that comes into mind as a solution in harsh environmental conditions like high risky areas [14], energy optimization and data recording in this kind of environments [15], [16].

The rest of the paper is organized as follows: In Section II, previous works related to current topic is presented. Then, the proposed HMN topology is introduced in Section III. The performance of the proposed HMN topology is demonstrated

in Section IV. Finally, the conclusions and future works are concluded in Section V.

II. RELATED WORKS

WMN is a wireless system with the decentralized topology, generally to serve a large geographical area. Every node in the network can independently interact with the other nodes around it without any centralized system and distribute data among each other. Each node having access to the surrounding nodes, makes sure that no interruption of data flow happens from one point to other if any of the intermediate node fails. The high level of reliability, along with the flexibility, self healing and self configurable properties [17] make WMN an ideal choice to integrate into emerging technologies and it is expected to play an important role in the IoT applications [18].

A routing protocol associated with every topology determines the specific choice of the route with which the components involved in the network communicates. A WMN can be employed with three types of protocols, which are proactive, reactive and hybrid protocols [19]. In proactive protocol [20], the route information is exchanged between the hosts periodically so that each node has the knowledge about the overall network. This protocol reduces the route discovery delay of the network but it is generally used for small networks. For instance, the protocols like Destination-Sequenced Distance-Vector Routing (DSDV) [21], Optimized Link State Routing Protocol (OLSR) [22] are the examples for proactive routing. On the other hand, the reactive protocols are the routing protocols that each node does not keep the knowledge of the full network topology but just only on demand. In this case, the route request will be flooded only when there is a requirement. As a result, the latency will be higher when a new route need to be discovered upon request. But, the reactive protocols reduce the traffic overhead as the request will be flooded only upon a requirement. The Dynamic Source Routing (DSR) [23] and Adhoc On-Demand Vector (AODV) [24] [25] are typically used reactive protocols. Given that the hybrid protocols are the combination of both proactive and reactive protocols, it exploits the advantages of both approaches. Additionally, Hybrid Wireless Mesh Protocol (HWMP) [26] and its secure version Secure Hybrid Wireless Mesh Protocol (SHWMP) are the common hybrid routing protocols that can be cited.

The conventional WMN find itself in a difficult position to fit in topologies with harsh environments. This is attributed to the fact that the path loss is high across these harsh environments and would finally end up being in no signal for such situation [27]. This along with the packet loss, which can happen with every hop in the WMNs, further exacerbates the performance of a network [28]. At the present, to the best of our knowledge, there is not any practical solution which can provide a reliable connection between the clients across signal opaque separation walls using WMN. A conventional WMN can provide a good connection in a non-harsh environments (e.g., open air) generally with a hop count of not more than four and also in a topology without any hard core separation walls. A normal wall cause permissible signal attenuation but

with high opaque severity separation, large amount of signal attenuation can occur which lead the signal not to penetrate the obstacles. Using a wired connection across the wall in such scenario can minimize the signal attenuation and make the signal pass through the metallic walls. Hence, a combination of conventional wireless mesh networking along with a wired mesh networking (i.e., HMN) can serve the requirement of providing a reliable communication with high throughput and low latency in such harsh environments, which is the main contribution of this paper. Although there are protocols like Open Shortest Path First (OSPF) which is basically designed for wired links and has the possibility of extending to the wireless links [29], the major issue with these kind of protocols is that they are not currently deployed for the WMNs and the complexity can arise while trying to utilize its features over WMN.

III. THE DESIGN OF HYBRID MESH NETWORK TOPOLOGY

A. The Proposed Hybrid Mesh Node Structure

This section describes the structure of the proposed hybrid mesh node. The hybrid mesh node mainly consists of modules performing encapsulation and decapsulation, which is shown in Figure 1.

In this structure, any received wireless mesh signal is broadcasted to other wireless mesh nodes like in a conventional wireless mesh network. The same signal is also converted to an Ethernet frame by the encapsulation module in order to cross over the metallic obstacles. Then, the encapsulated signal is later put through the Ethernet cable, which is connected across the metallic obstacles to another hybrid mesh node on the other side of the metallic obstacles. The hybrid mesh node is best suited to be positioned in high density obstacle environment where wireless coverage is obstructed by the metallic surface and walls. This topology is specially developed to address harsh environments, such as factories. The Ethernet and mesh frames are converted into the network seamlessly by means of hybrid mesh node which combines wireless and wired interfaces.

B. The Implementation of Hybrid Mesh Node in OMNET++ Environments

We have implemented the proposed hybrid mesh node in OMNET++ environments [30]. Note that OMNET++ is an Eclipse-based simulation tool that generally utilizes some simulation frameworks. Thus, INETMANET framework that is based on INET, is used for the implementation of the proposed hybrid mesh node.

In this study, the proposed hybrid mesh node is realized as a mesh gateway in OMNET++ platform. It consists of an Ethernet mesh interface and wireless mesh interface for the point to point connection and wireless mesh connection, respectively. The implemented hybrid mesh node in OMNET++ is illustrated in Figure 2. In this implementation, the hybrid mesh node receives packets from its neighbour wireless mesh node, and then the received frame is encapsulated as Ethernet frame and forwarded over the wire in order to cross

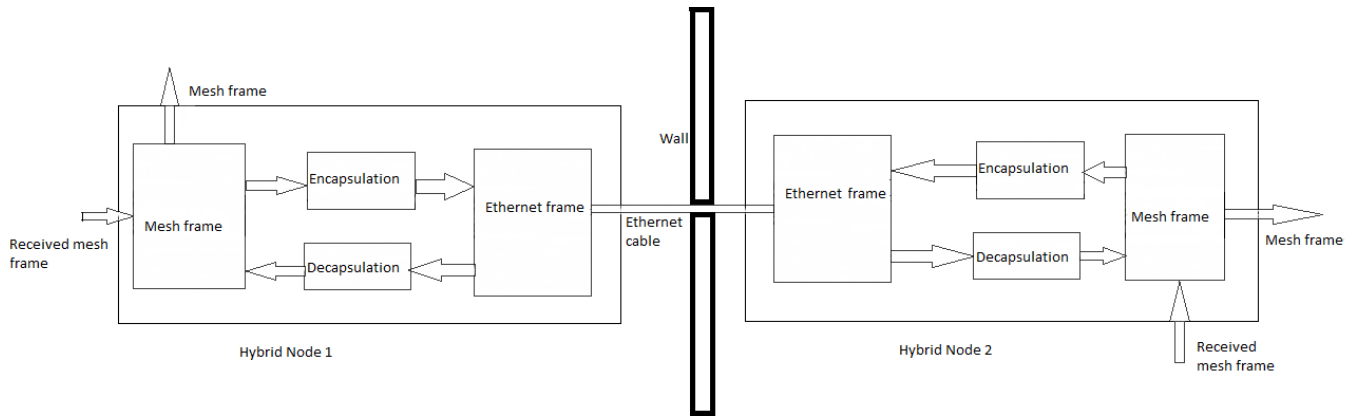


Fig. 1. The proposed hybrid mesh node structure

the metallic walls. On the opposite side of the wall, all the processes are applied reversely, which is the conversion of Ethernet frames into mesh frames through decapsulation module and then transmission of them over wireless interface. It is worth to note that the hybrid mesh node has an adaptive behavior in the sense that it can switch between wireless and wired interfaces if one of them fails. In other words, when the destination is accessible over wireless interface, the hybrid mesh node broadcasts mesh frames to other nodes acting as a mesh node without using wired interface. On the other hand, when the destination is not accessible over wireless interface, then the hybrid mesh node tries to expand the network coverage with the wired interface.

obstacles as shown in Fig 3. The industrial harsh environment is considered as $250m \times 100m$ with all metallic walls. This area has total 6 rooms and the metallic walls are configured in order to pass the data traffic over the hybrid mesh nodes. In the simulation setup, we consider 6 mesh hosts (e.g., mesh IIoT users), 6 Wi-Fi access points (APs) that is called fixhosts in OMNET++ and 10 mesh nodes (called MPP in OMNET++). It is worth to emphasize that the connection of the proposed hybrid mesh nodes is illustrated in OMNET++ as two hybrid mesh nodes connected back-to-back through a wall. Note that the Wi-Fi APs are equipped with dual radios, i.e. 5GHz and 2.4GHz in our study. Based on this dual-radio feature, all the Wi-Fi APs utilize 5GHz radio for the AP-to-AP communications links and 2.4GHz radio for the AP-to-Mesh hosts communications links. Without loss of generality, the mesh hosts, APs and hybrid mesh nodes are positioned sequentially in this setup for tracking multi-hop network easily. Thus, the proposed HMN is designed as a grid topology in OMNET++ environment. In conclusions, the simulation parameters in Table I is used to obtain the results unless otherwise stated.

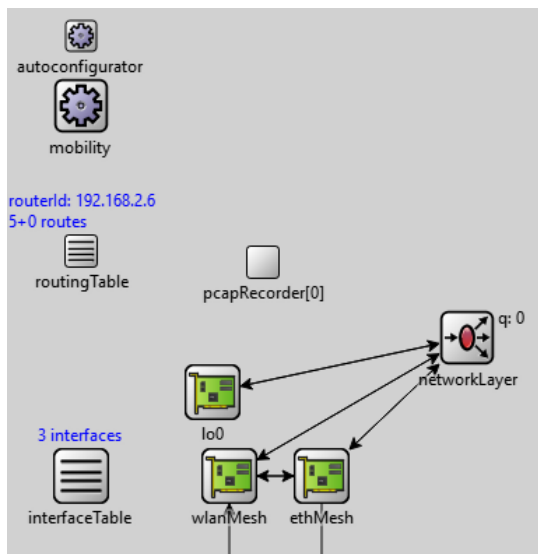


Fig. 2. The implementation of hybrid mesh node structure in OMNET++ environment

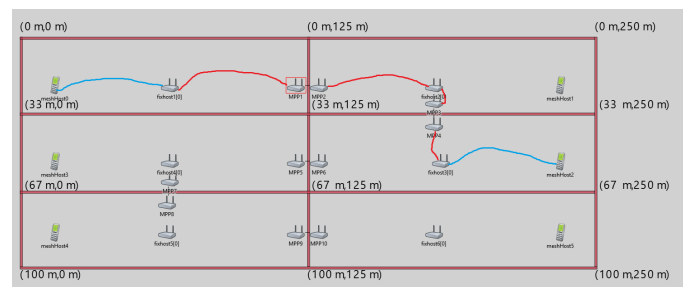


Fig. 3. Simulation environment for the proposed HMN

IV. SIMULATION SETUP AND RESULTS

A. Simulation Setup

In order to demonstrate the performance of the proposed HMN for multi-hop networks in harsh environments, we consider an industrial area consisting of metallic walls and

The simulation is performed for high density obstacle separations in order to show the performance of the proposed HMN in terms of throughput, delay, and jitter. We assume that meshHost0 sends video-stream packets to other Mesh hosts, which are meshHost1, meshHost2, meshHost3, meshHost4, and meshHost5. Note that each mesh host is associated with its mesh node, for example meshHost0 and fixhost1[0], and this is due to the design methodology of OMNET++ plat-

TABLE I
SIMULATION PARAMETERS

Parameters	Parameter values
Simulation Time	10s
Network Toplogy	Mesh
Number of Mesh IIoT Users	6
Area	250m x150m
Material Type	Aluminium
Radio Standard	IEEE 802.11g
Radio Bands	2.4GHz and 5GHz
Channel Propagation Model I	Rician ($K = 8$ dB)
Channel Propagation Model II	Rayleigh Fadings
Path-loss exponent	$\alpha = 2.5$
Radio Maximum Tx Power	11 mW
Data Rate	54Mbps
Application Data	UDP and TCP
Packet Size	1400B, 512B and 256B
Routing Protocol	HWMP
Routing Metric	Hop Count

form. Furthermore, the video-stream packets are sent through User Datagram Protocol (UDP). For the path selection in the proposed HMN shown in Figure 3, the hybrid mesh node broadcasts path request message (PREQ) which includes Medium Access Control (MAC) address of the destination to all other hybrid mesh nodes until it reaches to the destination node. Each PREQ message contains sequence number in order to specify the minimum cost of path selection. Once the PREQ message arrives at the destination, it sends unicast path reply message (PREP) to the source for updating their paths.

In order to compare the performance of the proposed HMN with the performance of the conventional WMN, simulation environment shown in Figure 4 is considered for the conventional WMN. This setup takes into account of the case where the walls are half open allowing the signal to propagate between rooms in order to form direct mesh connectivity. This represents the situations where the topology has walls with some soft separation areas like doors, windows and even some open areas. It is obvious that the conventional wireless mesh nodes are used in this scenario. For instance, all the wireless mesh nodes broadcast PREQ to neighboring wireless mesh nodes and find the path from fixhost1[0] to fixhost3[0]. In this case, it can be observed that the total number of hops required to reach fixhost3[0] is 4. The lack of wire link property of WMN, wireless mesh node needs to broadcast mesh frames all around its neighbourhood to find the path to the destination. In this case, extra wireless mesh nodes might be used with additional hops to support the signal path as shown in Figure 4. As a result, the increase in the traffic load can hamper the throughput and exacerbate the latency of the network and hence deteriorates the performance of the mesh network.

Another scenario is simulated by sending video-stream packet between two pairs simultaneously using both UDP and Transmission Control Protocol (TCP) in order to show the effects of multi-user TCP and UDP traffics on the performance of the proposed HMN. The pairs of mesh hosts for this scenario is given in Table II.

We also investigate the effects of propagation channel on the performance of the proposed HMN by considering the

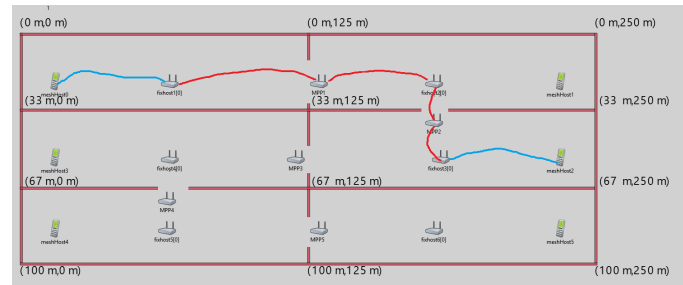


Fig. 4. Simulation environment for the conventional WMN

TABLE II
THE PAIRS OF MESH HOSTS

Server	Client
meshHost0	meshHost2
meshHost4	meshHost1

path loss and small-scale fading effects. For the small-scale fadings, we consider Rayleigh fading channels for Non-line-of-sight (NLOS) and Rician fading channels for Line-of-sight (LOS) environments [31].

B. Simulation Results

The simulation results are obtained using the simulations parameters and setup in the previous section. Initially, we study the performance of the proposed HMN in Rayleigh channel considering the effects of different packet size (1400B and 256B) and the performance results in terms of throughput, delay, and jitter are shown in Figures 5, 6 and 7, respectively. According to the results in Figure 5, the throughput decreases as the number of hops increases for both packet sizes. Furthermore, the throughput increases with the packet size proportionally since the low packet size require more packets need to be transmitted than the higher packet size for the same amount of data. For instance, the proposed HMN achieves 1.62 Mbps and 6.58 Mbps throughput when it utilizes packet sizes of 256B and 1400B, respectively, for 6 hops. Unless otherwise stated, we consider to use 1400B for the subsequent simulations.

The results in Figures 6 and 7 reveal that increasing the number of packet size for a given fixed time increases the delay and jitter. It is obvious that utilizing large packet size in order to send data will increase the delay and jitter. Additionally, it can be seen from the results that delay and jitter is proportional to the number of hops.

Following the performance study of the proposed HMN in harsh environments, now we compare its performance with that of conventional WMN in terms of throughput, delay and jitter considering the simulation environments in Figures 8, 9 and 10.

The results in Figure 8 show that the performance of the proposed HMN is better in Rician channel than Rayleigh channel. This is due to the fact that Rician is a channel modeling the LOS environments which is less harsh than the Rayleigh channel. More importantly, the proposed HMN outperforms the conventional WMN for both Rician and Rayleigh channels.

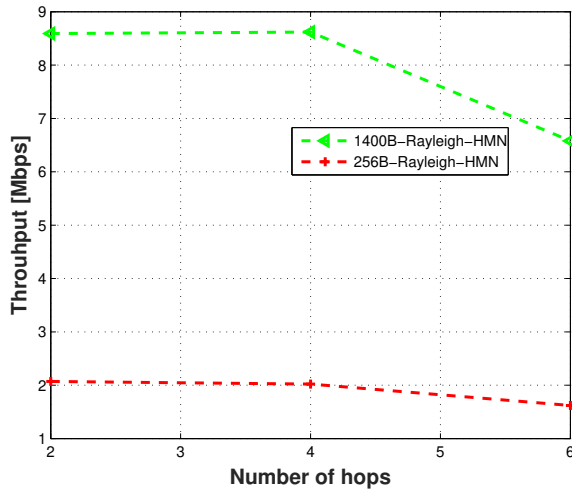


Fig. 5. Throughput vs. Number of Hops performance of HMN for different packet size in Rayleigh channels.

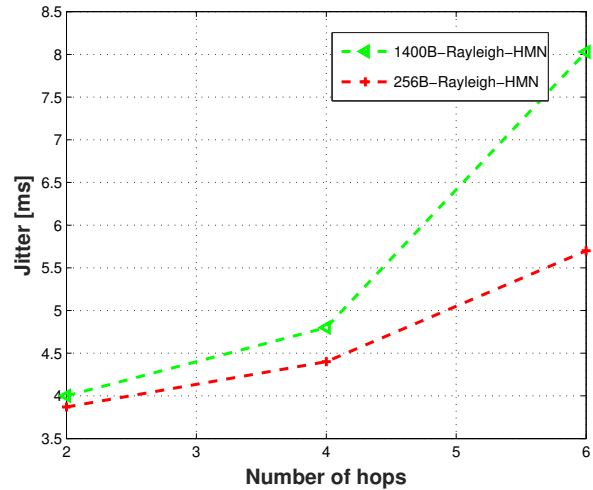


Fig. 7. Jitter vs. Number of Hops performance of HMN for different packet size

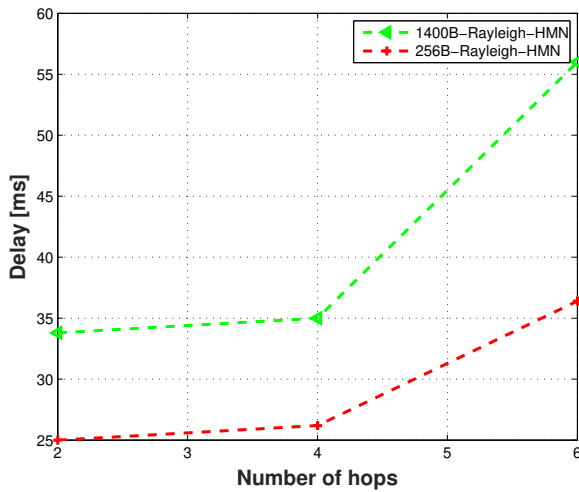


Fig. 6. Delay vs. Number of Hops performance of HMN for different packet size

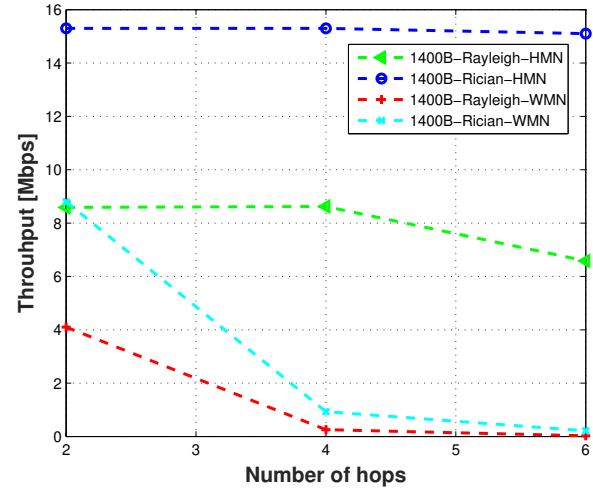


Fig. 8. Throughput performance of HMN and WMN under different channels.

The reason is the wired link reduces the traffic load just near its mesh nodes and also hybrid node capable of obtaining the Ethernet frames and sending mesh frames at the same time in the HMN case. On the other hand, in the WMN case, all the nodes are connected and communicates just only via wireless link that creates heavy traffic loads on each of them. As a result, the throughput of each node drastically decreases when the hop counts increase. Furthermore, it has to be noted that WMN's performance deteriorates for every hop count it undergoes whereas the changes are not that drastic in the HWM case. It is attributed to the usage of wide of nodes, such as MPP, fixhost, and host, rather than multiple simple mesh nodes as the only node utilized. The graph would depict a steeper plunge for WMN in performance and it is owing to the scale of representation in the plot as lower values comes as fractions and would be hard to represent when compared with other higher values.

According to the results in Figure 9, the proposed HMN experiences more delay in Rayleigh channel compared to Rician channel since the former one is more harsh environment than the latter one. Additionally, the proposed HMN performs better than the conventional WMN for both channels in terms of delay, especially for large number of hops. Similarly, the proposed HMN performs better than the conventional WMN for both channels in terms of jitter according to the results in Figure 10.

The performance of the proposed HMN topology is also evaluated under UDP and TCP traffic by sending video-stream packets and the role of nodes are given in Table II. The simulation results are obtained and tabulated in Table III in terms of throughput and number of hops. The results indicate that throughput is higher under UDP traffic than that of TCP for both two cases and it is attributed to the fact that TCP has more overhead inherently than that of UDP.

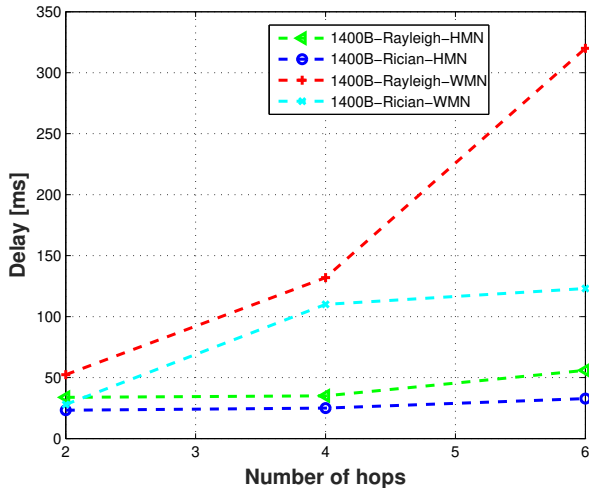


Fig. 9. Delay performance of HMN and WMN under different channels.

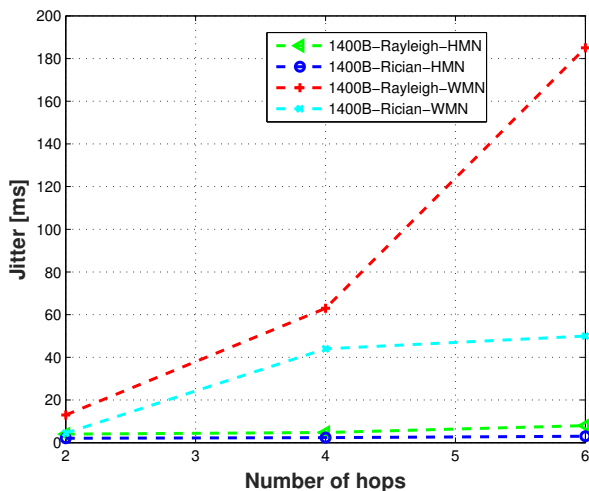


Fig. 10. Jitter performance of HMN and WMN under different channels.

TABLE III
PERFORMANCE OF HMN UNDER UDP AND TCP TRAFFIC

Client nodes	Throughput (Mbps)		Number of hops
	TCP	UDP	
meshHost1	1.3	6	6
meshHost2	0.97	7.19	4

In summary, the results show that the proposed HMN topology enables high performance wireless mesh networking even in harsh environments. This is achieved by the proposed HMN node, which converts the mesh frames into the Ethernet frames. This simply increases the coverage of network and it decreases the network latency and traffic load on the mesh nodes. As a result, the proposed HMN topology improves throughput, delay and jitter of conventional wireless mesh network. On the other hand, the conventional WMN suffers from the number of hops since the higher number of hops degrades the network performance and the complexity of

WMN increases with number of mesh nodes and number of connection.

V. CONCLUSIONS

In this study, we proposed a multi-hop dual-radio hybrid wireless-wired mesh network topology (HMN) in order to address the signal propagation issue in harsh environments by emphasizing on IIoT applications. The proposed HMN topology allows the signal to cross over obstacles, such as metallic walls in harsh environments. In addition, the proposed HMN topology is implemented in OMNET++ environments and its performance is evaluated for different packet sizes under different fading channels. Furthermore, the performance of the proposed HMN is compared with that of conventional WMN. The results reveal that the proposed HMN outperforms the conventional WMN under different fading channels. In addition, the performance of the proposed HMN is investigated under TCP and UDP traffic. The results suggest that the proposed HMN topology is a promising solution for the realization of multi-hop IIoT networks in harsh environments. Inclusion of MIMO and multi-channel to the proposed HMN is expected to improve the performance of the multi-hop IIoT networks in harsh environments further, which can be considered as future directions.

REFERENCES

- [1] M. Raza, N. Aslam, H. Le-Minh, S. Hussain, Y. Cao, and N. M. Khan, "A critical analysis of research potential, challenges, and future directives in industrial wireless sensor networks," *IEEE Communications Surveys Tutorials*, vol. 20, no. 1, pp. 39–95, 2018.
- [2] J.-S. Lee, Y.-W. Su, and C.-C. Shen, "A comparative study of wireless protocols: Bluetooth, uwb, zigbee, and wi-fi," in *IECON 2007 - 33rd Annual Conference of the IEEE Industrial Electronics Society*, 2007, pp. 46–51.
- [3] P. Rawat, K. D. Singh, H. Chaouchi, and J. M. Bonnin, "Wireless sensor networks: a survey on recent developments and potential synergies," *The Journal of supercomputing*, vol. 68, no. 1, pp. 1–48, 2014.
- [4] G. Zhao *et al.*, "Wireless sensor networks for industrial process monitoring and control: A survey," *Netw. Protoc. Algorithms*, vol. 3, no. 1, pp. 46–63, 2011.
- [5] C. Hennebert and J. D. Santos, "Security protocols and privacy issues into Glowpan stack: A synthesis," *IEEE Internet of Things Journal*, vol. 1, no. 5, pp. 384–398, 2014.
- [6] G. R. Hiertz, D. Denteneer, S. Max, R. Taori, J. Cardona, L. Berlemann, and B. Walke, "Ieee 802.11s: The wlan mesh standard," *IEEE Wireless Communications*, vol. 17, no. 1, pp. 104–111, 2010.
- [7] K. K.C., "Wireless mesh network: A survey," in *2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, 2016, pp. 1966–1970.
- [8] J. Karedal, S. Wyne, P. Almers, F. Tufvesson, and A. F. Molisch, "A measurement-based statistical model for industrial ultra-wideband channels," *IEEE Transactions on Wireless Communications*, vol. 6, no. 8, pp. 3028–3037, 2007.
- [9] W. S. Kish, A. Miu, and R. Mok, "Establishing a mesh network with wired and wireless links," May 22 2018, uS Patent 9,979,626.
- [10] A. Ciflone, L. Davoli, L. Belli, and G. Ferrari, "Wireless mesh networking: An iot-oriented perspective survey on relevant technologies," *Future Internet*, vol. 11, no. 4, p. 99, 2019.
- [11] T. Qiu, N. Chen, K. Li, M. Atiquzzaman, and W. Zhao, "How can heterogeneous internet of things build our future: A survey," *IEEE Communications Surveys Tutorials*, vol. 20, no. 3, pp. 2011–2027, 2018.
- [12] M. A. Ertürk, M. A. Aydin, L. Voller, and R. Setola, "Ieee 802.11 s mesh network analysis for post disaster communication," in *International Telecommunications Conference*. Springer, 2019, pp. 53–59.
- [13] M. R. Ericsson, A. Anand, J. Harkulich, and K. A. Helfrich, "Data collection for assistance in an industrial automation environment," Jul. 24 2018, uS Patent 10,031,495.

- [14] M. Thibaud, H. Chi, W. Zhou, and S. Piramuthu, "Internet of things (iot) in high-risk environment, health and safety (ehs) industries: A comprehensive review," *Decision Support Systems*, vol. 108, pp. 79–95, 2018.
- [15] G. d. Campo, S. Calatrava, G. Cañada, J. Olloqui, R. Martinez, and A. Santamaria, "Iot solution for energy optimization in industry 4.0: Issues of a real-life implementation," in *2018 Global Internet of Things Summit (GloTS)*, 2018, pp. 1–6.
- [16] F. Touati, A. S. P. Gonzales, Y. Qiblawey, and K. Benhmed, "A customized pv performance monitoring system in qatar's harsh environment," in *2018 6th International Renewable and Sustainable Energy Conference (IRSEC)*, 2018, pp. 1–6.
- [17] I. F. Akyildiz, X. Wang, and W. Wang, "Wireless mesh networks: a survey," *Computer networks*, vol. 47, no. 4, pp. 445–487, 2005.
- [18] Y. Liu, K.-F. Tong, X. Qiu, Y. Liu, and X. Ding, "Wireless mesh networks in iot networks," in *2017 International Workshop on Electromagnetics: Applications and Student Innovation Competition*, 2017, pp. 183–185.
- [19] V. M. S. and D. K. N., "Routing protocols for wireless mesh networks," *International Journal of Scientific & Engineering Research*, vol. 2, pp. 1–5, 2011. [Online]. Available: <http://www.ijser.org>
- [20] S. Kumar, "Reactive and proactive routing protocols for wireless mesh network using multimedia streaming," in *Proceedings of the International Conference on Recent Advances and Future Trends in Information Technology (iRAFIT 2012)*, 2012.
- [21] M. Manjunath and D. Manjaiah, "Spatial dsdv (s-dsdv) routing algorithm for mobile ad hoc network," in *2014 International Conference on Contemporary Computing and Informatics (IC3I)*, 2014, pp. 625–629.
- [22] —, "Spatial dsdv (s-dsdv) routing algorithm for mobile ad hoc network," in *2014 International Conference on Contemporary Computing and Informatics (IC3I)*, 2014, pp. 625–629.
- [23] M. Ramadhan and M. Davis, "Performance study of a cross-layer optimization to the dsr routing protocol in wireless mesh networks," in *2009 International Conference on Network and Service Security*, 2009, pp. 1–5.
- [24] S. Kolipaka, B. N. Bhandari, and A. Rajani, "Performance analysis of aodv with multi-radio in hybrid wireless mesh network," in *2014 Eleventh International Conference on Wireless and Optical Communications Networks (WOCN)*, 2014, pp. 1–5.
- [25] L. Bello, P. Bakalis, S. J. Manam, T. I. Eneh, and K. A. Anang, "Power control and performance comparison of aodv and dsr ad hoc routing protocols," in *2011 Uksim 13th International Conference on Computer Modelling and Simulation*, 2011, pp. 457–460.
- [26] I. Ibrahim, N. M. A. Latiff, S. K. S. Yusof, N. N. N. A. Malik, and S. H. S. Ariffin, "Performance comparison of aodv and hwmp routing protocols in wireless mesh networks," in *2013 IEEE International RF and Microwave Conference (RFM)*, 2013, pp. 116–120.
- [27] S. Aksu and G. K. Kurt, "Effects of propagation path loss on the quality of voip over wireless mesh networks," in *2011 IEEE 19th Signal Processing and Communications Applications Conference (SIU)*, 2011, pp. 912–915.
- [28] A. A. Pirzada, M. Portmann, and J. Indulska, "Performance comparison of multi-path aodv and dsr protocols in hybrid mesh networks," in *2006 14th IEEE International Conference on Networks*, vol. 2, 2006, pp. 1–6.
- [29] P. Garnepudi, T. Damarla, J. Gaddipati, and D. Veeraiah, "Proactive, reactive and hybrid multicast routing protocols for wireless mesh networks," in *2013 IEEE International Conference on Computational Intelligence and Computing Research*, 2013, pp. 1–7.
- [30] "Omnetpp Home Page," <https://omnetpp.org/>.
- [31] T. Sarkar, Z. Ji, K. Kim, A. Medouri, and M. Salazar-Palma, "A survey of various propagation models for mobile communication," *IEEE Antennas and Propagation Magazine*, vol. 45, no. 3, pp. 51–82, 2003.



Hasari Celebi received his BSc. degree in Electronics and Communications Engineering dept. from Yildiz Technical University in Istanbul, Turkey and his MSc. in EE dept. from San Jose State University in California, USA. Dr. Celebi received his PhD degree in EE dept. from University of South Florida in Florida, USA in 2008. He is currently a Full Professor in Computer Engineering Department at Gebze Technical University in Turkey. He is also the Director of Institute of Information Communications Technologies. Prior to this, he was with Texas A&M University at Qatar Campus as a Research Scientist. He was awarded for "The Research Fellow Excellence Award" at Texas A&M University at Qatar in 2010. He was also the recipient of "The Best Paper Award" in Crowncom 2009 Conference. In addition, he was awarded for "The Outstanding Graduate Achievement Award" at University of South Florida in 2008. His research areas include Statistical Signal Processing, Detection and Estimation theory, Localization, UWB, Cognitive Radio and Software Defined Radio (SDR). He has co-authored over 50 journal and international conference papers including four book chapters and he has 4 issued patents including US patents. He is an IEEE Senior Member.

Fractional Integration Based Feature Extractor for EMG Signals

Ibrahim Ethem Sacu

Abstract—Electromyography (EMG) signals have been extensively used for identification of finger movements, hand gestures and physical activities. In the classification of EMG signals, the performance of the classifier is widely determined by the feature extraction methods. Thus, plenty of feature extraction methods based on time, histogram and frequency domain have been reported in literature. However, these methods have several drawbacks such as high time complexity, high computation demand and user supplied parameters. To overcome these deficiencies, in this work, a new feature extraction method has been proposed to classify EMG signals taken from two different datasets. While one of the datasets includes 14 different finger movements, the other consists of 20 different physical activities. The proposed method is based on numerical fractional integration of time series EMG signals with different fractional orders. K Nearest Neighborhood (KNN) classifier with 8-fold cross validation has been employed for prediction of EMG signals. The derived fractional features can give better results than the two commonly used time domain features, notably, mean absolute value (MAV) and waveform length (WL) in terms of accuracy. The experimental results are also supported with statistical analysis.

Index Terms—Fractional integration, feature extraction, EMG, signal processing.

I. INTRODUCTION

Electromyography (EMG) signals are the measured bioelectrical signals that are generated by contraction of skeletal muscles during neuromuscular activities [1]. EMG signals have been widely used in broad scope of areas and applications such as diagnosis of some diseases [2], in the design of hand prosthesis [3]–[5] and entertainment industry [6]. With the advances of science, technology and speed of computers, EMG signals have become more important due to fact that they make people's life easier. Although the classification of hand gestures [1], [7]–[9] is most common in literature, the classification of finger movements [10], [11], movements of different body parts and body position [12] has been also researched.

The signal classification routine mainly consists of three parts including pre-processing, feature extraction and classification [6], [13]. In pre-processing step, so as to eliminate noises and interferences, the signals have been filtered firstly then, the windowing of datasets has been performed for the next step. On the other hand, in feature extraction step, the distinctive information based on time, frequency and/or time-frequency domains have been extracted. This step is very crucial owing to fact that it effects the performance of classifier, primarily. In the last step, the signals are classified

by an appropriate classifier according to some feature sets. The prevailing classifiers are support vector machines (SVMs), decision trees (DTs), Naive Bayes (NB), K Nearest Neighborhood (KNN), neural networks (NNs), etc. [14], [15].

In order to classify the EMG signals, plenty of feature extraction methods have been reported in literature [6], [13], [15]–[17]. These methods can be split into three categories, namely, time domain, frequency domain and time-frequency domain (TFD). Since time domain features are calculated from the amplitudes of raw EMG data without any additional transformation, features in this category are fast and simple to compute. On the other hand, frequency domain features are usually based on statistical properties of power spectral density or spectrum of the EMG signals. They are usually employed for the muscle fatigue studies. Although TFD features may cause a good classification accuracy, they are very complex and time-consuming to compute. Moreover, they need to reduce their high dimensions before the classification step. Feature extraction methods have shown variable performance for different problems. So, it is very complicated to say that which feature set is the best for a specific problem. Additionally, some of the available feature extraction methods have good classification accuracy but they are usually complex and high computational cost [16]. Therefore, the discovery of new accurate feature extraction methods is still ongoing research. So, the time domain fractional order integration based feature extractor has been proposed in this work.

Fractional calculus, which considers the differential and integral expressions of arbitrary order, has been gained increasing interest, recently. It can find applications in engineering, signal processing, control, biology and modeling [18]. In the field of signal processing, the fractional order derivative and integral operators are usually employed for filtering purpose [19]. By means of fractional derivative, the canceling of electrocardiogram (ECG) artifacts in surface EMG signals [20], the detection of P and T waves [21], the detection of R wave in ECG [22], zero phase filtering for ECG denoising [23] and heart rate detection [24] have been studied. On the other hand, by using fractional integral, zero phase filtering on ECG signal [25] and ECG signal denoising [26] have been investigated. Except usage of fractional derivative and integral operators, electroencephalography and EMG signals have been filtered by fractional order filters [27]–[29]. To the best of the authors' knowledge, there is no study about fractional integration based feature extraction method for EMG signals in the literature yet.

Fractional order integration is the integration of arbitrary order [30]. Even though there are various fractional order integral definitions, Caputo and Riemann-Liouville (R-L) types

I. ETHEM SACU is Clinical Engineering Research and Application Center, Erciyes University, Kayseri, Turkey. E-mail: iesacu@erciyes.edu.tr
<https://orcid.org/0000-0002-8627-8278>

Manuscript received Mar. 18, 2021; accepted Mar 15, 2022.

DOI: 10.17694/bajece.899088

are prominent ones [31]. It should be mentioned this point that these definitions are viable for continues time analog signals. However, their numerical version are also available in literature [32], [33]. The numerical version of R-L fractional integral in [33] is exploited in this work. The well-known time domain features of EMG signals mean absolute value (MAV), zero crossing, slope sign changes, waveform length (WL), Willison amplitude, variance, skewness, kurtosis etc. have been broadly used [1]. However, a number of valuable studies have shown that some of these features are redundant while the features of MAV and WL are usually significant and distinctive in terms of classification performance [1], [6]. Since the introduced feature extraction method is based on time-domain, this method has been compared with MAV and WL features.

Observing the performance of the proposed extraction method clearly, two different EMG signal datasets are taken into consideration. One of the datasets has eight channel fourteen different finger movements [11], while the other dataset has twenty different physical activities. On the other hand, to classify the EMG signals, KNN classifier with 8-fold cross validation has been operated for twenty times.

This work is structured as four sections. In the first section, the main reason of this work is introduced. In the second section, feature extraction methods with the proposed one are considered. In the section three, experimental results are presented and discussed. In the last section, in the lights of experimental results, the conclusion and possible future works are given.



Fig. 1. Finger motions

II. METHOD

In this section, three steps involving pre-processing, feature extraction and classification are performed, respectively.

A. Datasets

The dataset 1 is provided by Khushaba et al. for safe driving. During car driving, the distracting motions such as tuning

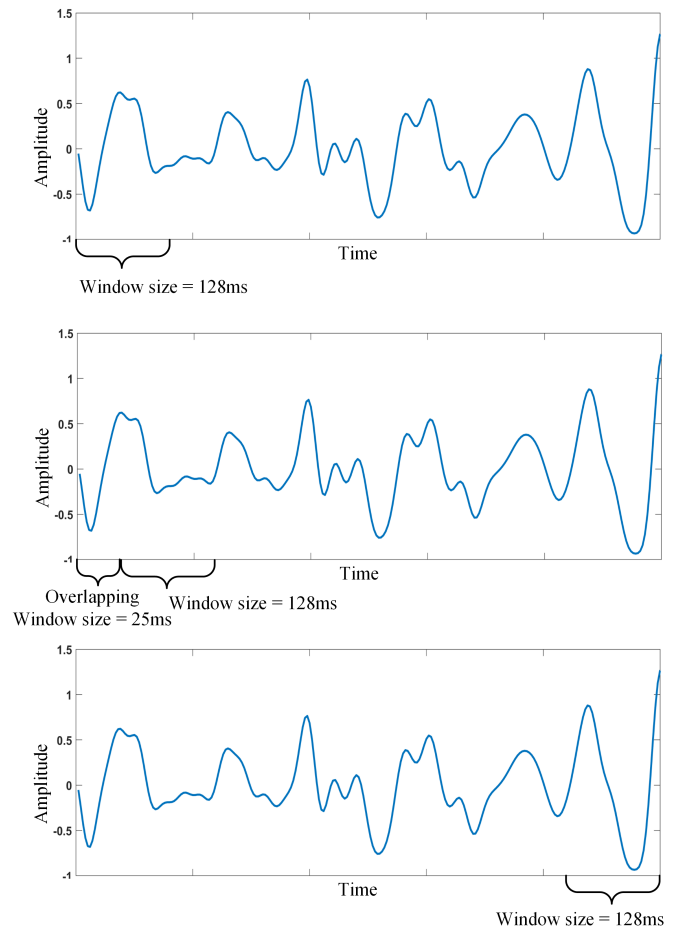


Fig. 2. Sliding windowing of data

the radio, controlling the windows and answering the phone endanger drivers' lives. To tackle this problem, simple finger motions (FM) can be used to perform the aforementioned tasks. However, the classification of FM is not that easy. Fourteen different FM demonstrated in Fig. 1 are recorded from eight channel EMG apparatus that is connected to driver's right arm. The eight subjects are participated in the experience. On the other hand, the dataset 2 involves three male and one female subjects. Each subject performed ten normal and ten aggressive physical activities (PA) during sessions. Eight channel EMG signals are recorded from right and left arms as well as legs.

B. Pre-processing

Pre-processing is split into two steps. In the first step, raw EMG signals are denoised by means of the sixth order band-pass Butterworth filter with $20 - 450Hz$ bandwidth. Therefore, meaningful information is obtained. The second step is windowing of the filtered signals. To do this, a window of $128ms$ wide is sliding on the signals with a time step $25ms$ as shown in Fig. 2. Each window of $128ms$ wide is used for the feature extraction.

C. Feature Extraction

High accuracy classification rate is achieved by effective feature extraction. For dataset 1, a typical signal part of 128ms window wide corresponds to 512 samples per channel due to sampling frequency 4kHz. Totally, 4096 samples are available for eight channel. When whole signal length is considered, such a high sample can not be directly load the classifier. So feature extraction is a vital part of whole signal processing process. On the other hand, for a dataset 2, 331 samples per channel and totally 2648 samples are available. The features are numerical data that represents the original signal as good as possible. In this work, the following feature extraction methods are exploited owing to fast and easy computation.

1) *Mean Absolute Value (MAV)*: MAV is the mean of the absolute values of the respective EMG signal amplitude. For EMG samples x_k ($k = 1, 2, \dots, N$), the mathematical expression of MAV is given as

$$MAV = \frac{1}{N} \sum_{k=1}^N |x_k| \quad (1)$$

where N represents the total data points in a window.

2) *Waveform Length (WL)*: WL is the sum of the amplitude differences between adjacent data points. It is the indicator of degree of amplitude change. WL is expressed as

$$WL = \sum_{k=1}^{N-1} |x_{k+1} - x_k| \quad (2)$$

where x_k and x_{k+1} are two adjacent data points of respective signal and N is the sample size.

3) *The Proposed Feature Extraction Methods*: The proposed new feature extraction methods are based on fractional integration of arbitrary order. It is known that MAV and WL depend on integration of the integer order that is a special case of the fractional order. However, the fractional integration removes this limitation and provides extra freedom for researchers. The numerical version of R-L fractional integration is reported in [33] and computed in the interval $(0, T)$ as

$$I^\alpha x_N(h) = \frac{h^\alpha}{\Gamma(2+\alpha)} \sum_{n=0}^N c_{n,N} x_n \quad (3)$$

$$I^\alpha x(T) = I^\alpha x_N(h) + O(h^2) \quad (4)$$

where $\alpha > 0$ is fractional order, $h = T/N$ is time step, N is total data points $\Gamma(\cdot)$ is Gamma function, $x_{n,N}$ is respective signal, $O(h^2)$ is error term and $c_{n,N}$ is quadrature weights that is defined as below.

$$c_{n,N} = \begin{cases} (1+\alpha)N^\alpha - N^{\alpha+1} + (N-1)^{\alpha+1} & \text{if } n=0 \\ (N-n+1)^{\alpha+1} - 2(N-n)^{\alpha+1} + (N-n-1)^{\alpha+1} & \text{if } 0 < n < N \\ 1 & \text{if } n=N \end{cases} \quad (5)$$

4) *Fractional Absolute Value (FAV)*: FAV is the fractional integration of the absolute values of EMG signal points. Fractional integration order range from 0.1 to 1.5. The general expression of FAV is given below.

$$FAV = I^\alpha |x_k| \quad (6)$$

If (6) is inserted into (3), the following equation is derived as

$$FAV = \frac{h^\alpha}{\Gamma(2+\alpha)} \sum_{k=0}^N c_{k,N} |x_k| \quad (7)$$

where x_k ($k = 0, 1, \dots, N$) is EMG signal amplitude at the k^{th} point.

5) *Fractional Waveform Length (FWL)*: FWL is the fractional integration of the amplitude differences between adjacent data points of EMG signal. Fractional integration order range from 0.1 to 1.5. The general expression of FWL is given below.

$$FWL = I^\alpha |x_{k+1} - x_k| \quad (8)$$

By inserting (8) into (3), the following equation is derived as

$$FWL = \frac{h^\alpha}{\Gamma(2+\alpha)} \sum_{k=0}^N c_{k,N} |x_{k+1} - x_k| \quad (9)$$

III. RESULTS AND DISCUSSIONS

Extensive experiments on two public EMG data sets are conducted on a PC with an Intel I7 4GHz processor with 16GB of RAM to verify the effectiveness and efficiency of the proposed methods, including FAV and FWL. The performance of the proposed methods are investigated for the different fractional orders. The accuracy value, which is obtained by using KNN classification method, is used as a performance metric for all simulations conducted in this study. The simulations are conducted by using two different data sets and are repeated 20 times with 8-cross validation technique for all fractional orders related to proposed methods and traditional methods including MAV and WL.

A. Comparison of the classification performance of the proposed method with respect to fractional order

In this section, we show how to select the fractional order (FO) for EMG data sets in FAV and FWL schemes. Since, the maximum possible accuracy value of the proposed method may vary between different fractional order values. For this reason, the results on each data set for each proposed method are presented in Fig. 3 and 4 for better visual clarity.

The Fig. 3 and 4 show that the changing accuracy values of the proposed methods FAV and FWL with respect to fractional order for the FM and PA data sets, respectively. The accuracy rates are demonstrated in Fig. 3 for FM dataset where the average accuracy rates raised from 36.45% to 97.59% with changing $FO = 0.1$ to 1.07. $FO = 0.1$ score the lowest accuracy rate of 36.45%, while $FO = 1.04$ and 1.07 produce the highest accuracy rate of 97.59%. On the other hand in Fig. 4 we observe the lowest accuracy rate 46.4% with $FO = 0.1$ and the highest one 97.99% with $FO = 1.07$ for FAV method in PA data set. Also, it seen that from the same figure proposed method FWL has the worst and best performance for $FO = 0.1$ and $FO = 1.06$, respectively.

The results in Fig. 3 and 4 show that the classification accuracy of the proposed method for all four cases always yields the smallest classification performance between $FO = 0.1$ and $FO = 1.01$. The accuracy of proposed method is observed to

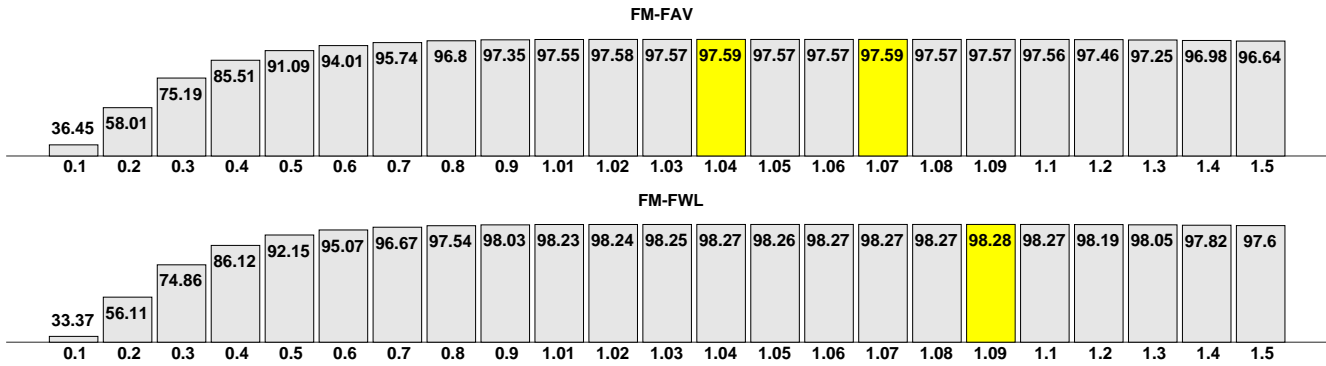


Fig. 3. The changing accuracy values of the proposed methods FAV and FWL with respect to the FO for the FM data set.

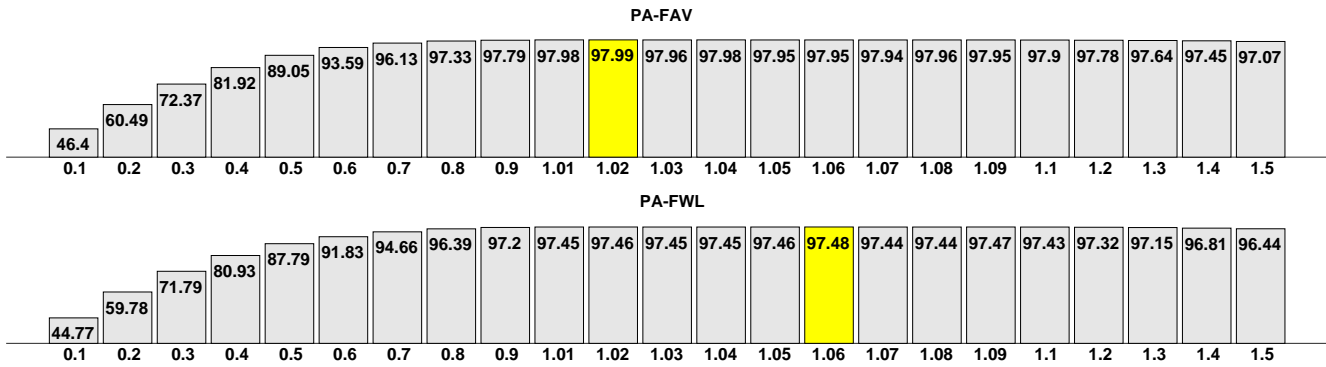


Fig. 4. The changing accuracy values of the proposed methods FAV and FWL with respect to the FO for the PA data set.

increase often monotonically with increasing the FO values up to a point where its location between $FO = 1.01$ and $FO = 1.1$ and its best performance is attained in this range. In simulation, we observe that the typical ranges of appropriate FO parameter do not usually vary dramatically for two EMG sets and adjusting this parameter to values within the interval $FO \in [1.01, 1.1]$

To better assess how accuracies change according to each simulation, we represent results in Fig. 5, 6, 7 and 8 for the nine best FO values. The results for the FM data set are presented in Fig. 5 and 6 with red and black color, along with the result of the conventional methods MAV and WL with blue color, for comparison. The proposed methods performance are better than conventional methods. Particularly proposed methods FAV and FWL exhibit the best classification performance for $FO = 1.04$ and $FO = 1.09$ with red color, respectively for FM data set. In the other data set PA, when comparing the proposed methods FAV and FWL to conventional methods MAV and WL, the proposed methods FAV with $FO = 1.02$ and FWL with $FO = 1.06$ has the best accuracy for each of these measures.

Some detailed descriptive statistical values, including, maximum (max), minimum (min), mean and standard deviation (std) are illustrated in Table I, II, III and IV for FM data set with FAV, FWL methods and PA data set with FAV, FWL methods, respectively. The results for the proposed FAV ($FO = 1.04$) at FM data set given in Table I where its accu-

racy rate are higher than conventional method MAV. Regarding another proposed method FWL ($FO = 1.09$) at the same data set, it can be seen from Table II its classification performance is better than the conventional method WL. In the case of FM data set, the results are listed in Table III and Table IV where the best accuracies values are FAV ($FO = 1.02$) and FWL ($FO = 1.06$), respectively.

The proposed feature extraction methods based on fractional integration are sufficient especially regarding the complexity of the classification of EMG signals in a practical application.

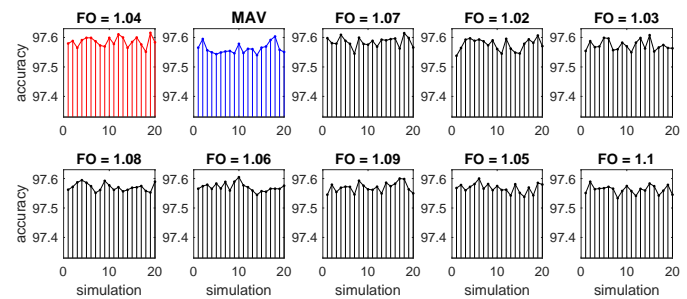


Fig. 5. Proposed method FAV performance throughout simulations for FM data set.

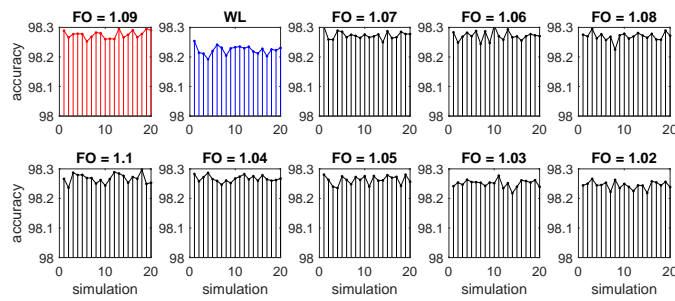


Fig. 6. Proposed method FWL performance throughout simulations for FM data set.

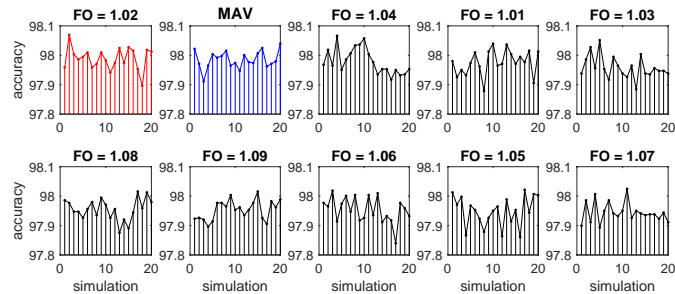


Fig. 7. Proposed method FAV performance throughout simulations for PA data set.

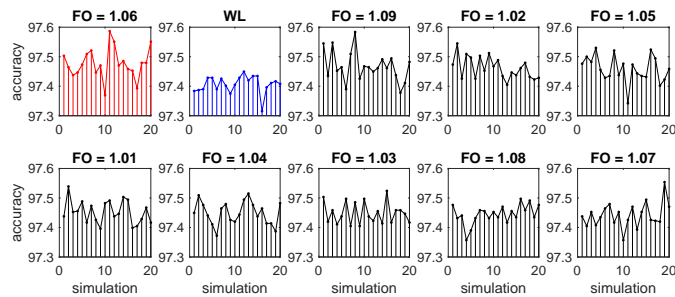


Fig. 8. Proposed method FWL performance throughout simulations for PA data set.

B. Statistical analysis for the proposed method

In line of obtained results from the experimental study and basic descriptive statistical analysis, it should be noted that the proposed method has superior classification performance than the conventional approaches. However, in this section we validate the obtained results by performing some additional statistical analysis. Since we understand whether there is a sufficient difference between the proposed method and the conventional method or not, we have performed the Wilcoxon signed-rank test, which is a non-parametric statistical test utilized to compare two repeated methods to assess whether the two methods are different from each other. This test is an alternative to the paired t-test when the distribution of the differences between the two samples cannot be assumed to be normally distributed. A Wilcoxon signed-rank test is a nonparametric test, which is utilized to determine whether two dependent samples are selected from populations having the same distribution.

The statistical results are listed in Table V and VI for FM

TABLE I
SOME DETAILED DESCRIPTIVE STATISTICAL VALUES, INCLUDING, MAX, MIN, MEAN AND STD FOR FAV AT FM DATA SET

FM-FAV (1.04)				
Method	Max	Mean	Min	std
FO: 0.1	36.5354	36.4508	36.3472	0.0537
FO: 0.2	58.09	58.0064	57.9199	0.0506
FO: 0.3	75.3074	75.1853	75.091	0.0478
FO: 0.4	85.573	85.5058	85.4662	0.0281
FO: 0.5	91.1336	91.089	91.0375	0.0252
FO: 0.6	94.0518	94.0054	93.9647	0.0245
FO: 0.7	95.7804	95.7436	95.7092	0.018
FO: 0.8	96.8417	96.7994	96.7738	0.0168
FO: 0.9	97.3904	97.35	97.318	0.0195
FO: 1.01	97.5978	97.545	97.509	0.022
FO: 1.02	97.6068	97.5758	97.5379	0.0203
FO: 1.03	97.6079	97.5733	97.5497	0.0172
FO: 1.04	97.6158	97.5858	97.5514	0.0165
FO: 1.05	97.6	97.5677	97.5379	0.0159
FO: 1.06	97.6045	97.5705	97.5446	0.0138
FO: 1.07	97.6147	97.5853	97.5452	0.0167
FO: 1.08	97.5944	97.571	97.5514	0.0132
FO: 1.09	97.6	97.5703	97.5452	0.0166
FO: 1.1	97.5893	97.5623	97.5333	0.0146
FO: 1.2	97.4876	97.4585	97.439	0.0122
FO: 1.3	97.3011	97.2539	97.2152	0.0254
FO: 1.4	97.027	96.9759	96.9439	0.0206
FO: 1.5	96.6631	96.6393	96.6009	0.0195
MAV	97.6034	97.562	97.5401	0.0177

and PA data sets. It can be seen that from Table V there is no statistical significance between the proposed method FAV with $FO = 1.04$ and its variants, including, FAV with $FO = 1.02, 1.03, 1.06, 1.07, 1.08$ and 1.09 in FM data set. On the other hand in favor of FAV with 1.04 we observe statistical significance between FAV with 1.04 and conventional technique MAV. In case of FWL with $FO = 1.09$, we observe nearly the same results.

TABLE II
SOME DETAILED DESCRIPTIVE STATISTICAL VALUES, INCLUDING, MAX, MIN, MEAN AND STD FOR FWL AT FM DATA SET

FM-FWL (1.09)				
Method	Max	Mean	Min	std
FO: 0.1	33.4426	33.3686	33.2521	0.0466
FO: 0.2	56.2048	56.115	56.042	0.0524
FO: 0.3	74.9435	74.8614	74.7836	0.0478
FO: 0.4	86.1737	86.1159	86.0432	0.0345
FO: 0.5	92.2061	92.1516	92.088	0.0325
FO: 0.6	95.1091	95.0726	95.0429	0.018
FO: 0.7	96.6936	96.6652	96.6354	0.0167
FO: 0.8	97.5627	97.5358	97.5113	0.0177
FO: 0.9	98.0431	98.0253	98.0024	0.0111
FO: 1.01	98.251	98.2305	98.1951	0.0134
FO: 1.02	98.2663	98.2447	98.2188	0.0126
FO: 1.03	98.277	98.2506	98.2177	0.0126
FO: 1.04	98.2866	98.267	98.2465	0.0105
FO: 1.05	98.2804	98.2627	98.2352	0.0146
FO: 1.06	98.3064	98.2715	98.2442	0.0159
FO: 1.07	98.303	98.2735	98.2499	0.0124
FO: 1.08	98.2951	98.2695	98.225	0.0144
FO: 1.09	98.2957	98.2756	98.2527	0.0125
FO: 1.1	98.2979	98.2676	98.2363	0.0163
FO: 1.2	98.2177	98.1942	98.1736	0.014
FO: 1.3	98.0736	98.0508	98.0182	0.0147
FO: 1.4	97.8447	97.8217	97.7922	0.0168
FO: 1.5	97.6368	97.5984	97.5712	0.0171
WL	98.2538	98.2235	98.1922	0.0144

TABLE III
SOME DETAILED DESCRIPTIVE STATISTICAL VALUES, INCLUDING, MAX, MIN, MEAN AND STD FOR FAV AT PA DATA SET

PA-FAV (1.02)				
Method	Max	Mean	Min	std
FO: 0.1	46.7181	46.4002	46.1878	0.1173
FO: 0.2	60.6382	60.4873	60.3522	0.0852
FO: 0.3	72.568	72.3661	72.1866	0.1107
FO: 0.4	82.0397	81.9231	81.7627	0.0739
FO: 0.5	89.1637	89.0464	88.9015	0.0709
FO: 0.6	93.6627	93.5897	93.4601	0.0504
FO: 0.7	96.1982	96.13	96.0343	0.0447
FO: 0.8	97.387	97.3277	97.2231	0.0448
FO: 0.9	97.8488	97.794	97.7386	0.0327
FO: 1.01	98.0395	97.9759	97.8786	0.0424
FO: 1.02	98.0693	97.9892	97.8965	0.0378
FO: 1.03	98.0514	97.9585	97.8846	0.0386
FO: 1.04	98.0663	97.9799	97.9173	0.044
FO: 1.05	98.0216	97.9485	97.8607	0.0513
FO: 1.06	98.0187	97.9531	97.8399	0.0443
FO: 1.07	98.0246	97.9445	97.8935	0.0338
FO: 1.08	98.0157	97.955	97.8756	0.0366
FO: 1.09	98.0157	97.9533	97.8965	0.0335
FO: 1.1	97.9829	97.899	97.8429	0.0354
FO: 1.2	97.8607	97.7825	97.7058	0.0457
FO: 1.3	97.7088	97.6352	97.5807	0.0309
FO: 1.4	97.5092	97.4506	97.393	0.0273
FO: 1.5	97.1665	97.0683	96.9848	0.0523
MAV	98.0395	97.985	97.9114	0.0299

In case of PA data set, we observed statistically significant results among the different FO values. However, there is no significance between proposed method FAV with $FO = 1.02$ and MAV methods. The possible reason of the this result is that conventional MAV has $FO = 1$, which is near to $FO = 1.02$. Besides, in favor of the proposed method FWL with $FO = 1.06$ it seen that there is statistically significance between FWL and conventional WL.

TABLE IV
SOME DETAILED DESCRIPTIVE STATISTICAL VALUES, INCLUDING, MAX, MIN, MEAN AND STD FOR WL AT PA DATA SET

PA-FWL (1.06)				
Method	Max	Mean	Min	std
FO: 0.1	44.9692	44.7651	44.5491	0.11
FO: 0.2	60.0185	59.7835	59.6043	0.1097
FO: 0.3	71.9632	71.7854	71.5639	0.0987
FO: 0.4	81.0863	80.9338	80.7675	0.1012
FO: 0.5	87.8795	87.7855	87.6263	0.0713
FO: 0.6	91.9048	91.8278	91.6515	0.0597
FO: 0.7	94.7	94.6557	94.5297	0.0602
FO: 0.8	96.4634	96.3914	96.2935	0.0455
FO: 0.9	97.2768	97.1978	97.1189	0.0472
FO: 1.01	97.539	97.4528	97.3959	0.0389
FO: 1.02	97.5449	97.4624	97.4049	0.0376
FO: 1.03	97.5241	97.4485	97.4049	0.0365
FO: 1.04	97.5151	97.4488	97.3721	0.0393
FO: 1.05	97.53	97.4585	97.3423	0.0454
FO: 1.06	97.5866	97.4774	97.3691	0.0513
FO: 1.07	97.5539	97.4418	97.3572	0.0419
FO: 1.08	97.4972	97.4448	97.3572	0.0329
FO: 1.09	97.5836	97.4676	97.3781	0.0517
FO: 1.1	97.5568	97.4317	97.3572	0.056
FO: 1.2	97.381	97.3164	97.2321	0.0441
FO: 1.3	97.2261	97.1478	97.0861	0.0372
FO: 1.4	96.8894	96.8062	96.7256	0.039
FO: 1.5	96.517	96.4379	96.2995	0.0588
WL	97.4496	97.4065	97.3155	0.0294

TABLE V
STATISTICAL RESULTS FOR THE FM DATA SET

Finger Movement Data Set					
MAV-FO: 1.04			WL-FO: 1.09		
Method	p-val	Sig.	Method	p-val	Sig.
FO: 0.1	0	FO: 1.04	FO: 0.1	0	FO: 1.09
FO: 0.2	0	FO: 1.04	FO: 0.2	0	FO: 1.09
FO: 0.3	0	FO: 1.04	FO: 0.3	0	FO: 1.09
FO: 0.4	0	FO: 1.04	FO: 0.4	0	FO: 1.09
FO: 0.5	0	FO: 1.04	FO: 0.5	0	FO: 1.09
FO: 0.6	0	FO: 1.04	FO: 0.6	0	FO: 1.09
FO: 0.7	0	FO: 1.04	FO: 0.7	0	FO: 1.09
FO: 0.8	0	FO: 1.04	FO: 0.8	0	FO: 1.09
FO: 0.9	0	FO: 1.04	FO: 0.9	0	FO: 1.09
FO: 1.01	0	FO: 1.04	FO: 1.01	0	FO: 1.09
FO: 1.02	0.1367	Nan	FO: 1.02	0	FO: 1.09
FO: 1.03	0.0166	Nan	FO: 1.03	0	FO: 1.09
FO: 1.05	0.0025	FO: 1.04	FO: 1.04	0.0274	FO: 1.09
FO: 1.06	0.0055	Nan	FO: 1.05	0.0123	FO: 1.09
FO: 1.07	0.8924	Nan	FO: 1.06	0.4247	Nan
FO: 1.08	0.0055	Nan	FO: 1.07	0.4648	Nan
FO: 1.09	0.0055	Nan	FO: 1.08	0.2183	Nan
FO: 1.1	0.0001	FO: 1.04	FO: 1.1	0.1554	Nan
FO: 1.2	0	FO: 1.04	FO: 1.2	0	FO: 1.09
FO: 1.3	0	FO: 1.04	FO: 1.3	0	FO: 1.09
FO: 1.4	0	FO: 1.04	FO: 1.4	0	FO: 1.09
FO: 1.5	0	FO: 1.04	FO: 1.5	0	FO: 1.09
MAV	0.0003	FO: 1.04	WL	0	FO: 1.09

TABLE VI
STATISTICAL RESULTS FOR THE PA DATA SET

Physical Action Data Set					
MAV-FO: 1.02			WL-FO: 1.06		
Method	p-val	Sig.	Method	p-val	Sig.
FO: 0.1	0	FO: 1.02	FO: 0.1	0	FO: 1.06
FO: 0.2	0	FO: 1.02	FO: 0.2	0	FO: 1.06
FO: 0.3	0	FO: 1.02	FO: 0.3	0	FO: 1.06
FO: 0.4	0	FO: 1.02	FO: 0.4	0	FO: 1.06
FO: 0.5	0	FO: 1.02	FO: 0.5	0	FO: 1.06
FO: 0.6	0	FO: 1.02	FO: 0.6	0	FO: 1.06
FO: 0.7	0	FO: 1.02	FO: 0.7	0	FO: 1.06
FO: 0.8	0	FO: 1.02	FO: 0.8	0	FO: 1.06
FO: 0.9	0	FO: 1.02	FO: 0.9	0	FO: 1.06
FO: 1.01	0.3716	Nan	FO: 1.01	0.1134	
FO: 1.03	0.0065	FO: 1.02	FO: 1.02	0.2232	
FO: 1.04	0.2729	Nan	FO: 1.03	0.0396	FO: 1.06
FO: 1.05	0.0094	FO: 1.02	FO: 1.04	0.0719	FO: 1.06
FO: 1.06	0.0105	FO: 1.02	FO: 1.05	0.2182	
FO: 1.07	0.0004	FO: 1.02	FO: 1.07	0.0122	FO: 1.06
FO: 1.08	0.0065	FO: 1.02	FO: 1.08	0.0127	FO: 1.06
FO: 1.09	0.0053	FO: 1.02	FO: 1.09	0.3864	
FO: 1.1	0	FO: 1.02	FO: 1.1	0.0109	FO: 1.06
FO: 1.2	0	FO: 1.02	FO: 1.2	0	FO: 1.06
FO: 1.3	0	FO: 1.02	FO: 1.3	0	FO: 1.06
FO: 1.4	0	FO: 1.02	FO: 1.4	0	FO: 1.06
FO: 1.5	0	FO: 1.02	FO: 1.5	0	FO: 1.06
MAV	0.6649	Nan	WL	0	FO: 1.06

In general, we can say that proposed method is statistically sufficient than the conventional methods. However proposed methods produce close results according to same FO values ranging 1.02 and 1.09.

IV. CONCLUSION

We proposed a new technique based on fractional integration to enhance the performance of a EMG feature extractors, including MAV and WL. The obtained results show that new MAV and WL, renamed as FAV and FWL, respectively can be used as sufficient feature extraction method. This method

can be applied any feature extractor based on integration for further researches. Only limitation of the proposed method is that it is depend on an user supplied parameter fractional order. In order to solve this problem we performed number of simulations, which lead how to choose this value. On the other hand, the fractional derivative based feaure extraction method may be a candidate for the future studies.

REFERENCES

- [1] C. Sapsanis, G. Georgoulas, and A. Tzes, "Emg based classification of basic hand movements based on time-frequency features," in *21st Mediterranean Conference on Control and Automation*, June 2013, pp. 716–722.
- [2] H. Kataoka and K. Sugie, "Recent advancements in lateral trunk flexion in parkinson disease," *Neurology. Clinical practice*, vol. 9, no. 1, p. 74–82, February 2019.
- [3] F. H. Y. Chan, Yong-Sheng Yang, F. K. Lam, Yuan-Ting Zhang, and P. A. Parker, "Fuzzy emg classification for prosthesis control," *IEEE Transactions on Rehabilitation Engineering*, vol. 8, no. 3, pp. 305–311, Sep. 2000.
- [4] K. Andrianesis and A. Tzes, "Design of an anthropomorphic prosthetic hand driven by shape memory alloy actuators," in *2008 2nd IEEE RAS EMBS International Conference on Biomedical Robotics and Biomechanics*, Oct 2008, pp. 517–522.
- [5] N. Parajuli, N. Sreenivasan, P. Bifulco, M. Cesarelli, S. Savino, V. Niola, D. Esposito, T. J. Hamilton, G. R. Naik, U. Gunawardana, and G. D. Gargiulo, "Real-time emg based pattern recognition control for hand prostheses: A review on existing methods, challenges and future implementation," in *Sensors*, 2019.
- [6] A. Phinyomark, P. Phukpattaranont, and C. Limsakul, "Feature reduction and selection for emg signal classification," *Expert Syst. Appl.*, vol. 39, no. 8, p. 7420–7431, Jun. 2012.
- [7] X. Zhang, X. Chen, Y. Li, V. Lantz, K. Wang, and J. Yang, "A framework for hand gesture recognition based on accelerometer and emg sensors," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 41, no. 6, pp. 1064–1076, Nov 2011.
- [8] J. Qi, G. Jiang, G. Li, Y. Sun, and B. Tao, "Surface emg hand gesture recognition system based on pca and grnn," *Neural Computing and Applications*, Mar 2019.
- [9] S. Raurale, J. McAllister, and J. M. del Rincón, "Emg wrist-hand motion recognition system for real-time embedded platform," *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1523–1527, 2019.
- [10] F. V. G. Tenore*, A. Ramos, A. Fahmy, S. Acharya, R. Etienne-Cummings, and N. V. Thakor, "Decoding of individuated finger movements using surface electromyography," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 5, pp. 1427–1434, May 2009.
- [11] R. N. Khushaba, S. Kodagoda, D. Liu, and G. Dissanayake, "Muscle computer interfaces for driver distraction reduction," *Computer Methods and Programs in Biomedicine*, vol. 110, no. 2, pp. 137 – 149, 2013.
- [12] A. Phinyomark, R. N. Khushaba, and E. Scheme, "Feature extraction and selection for myoelectric control based on wearable emg sensors," *Sensors*, vol. 18, no. 5, p. 1615, May 2018. [Online]. Available: <http://dx.doi.org/10.3390/s18051615>
- [13] M. B. I. Reaz, M. S. Hussain, and F. Mohd-Yasin, "Techniques of emg signal analysis: detection, processing, classification and applications," *Biological Procedures Online*, vol. 8, no. 1, pp. 11–35, Dec 2006.
- [14] K. Anam and A. Al-Jumaily, "Evaluation of extreme learning machine for classification of individual and combined finger movements using electromyography on amputees and non-amputees," *Neural Networks*, vol. 85, pp. 51 – 68, 2017.
- [15] L. Zhang, Y. Shi, W. Wang, Y. Chu, and X. Yuan, "Real-time and user-independent feature classification of forearm using emg signals," *Journal of the Society for Information Display*, vol. 27, no. 2, pp. 101–107, 2019.
- [16] A. Ameri, M. A. Akhaee, E. Scheme, and K. Englehart, "Regression convolutional neural network for improved simultaneous EMG control," *Journal of Neural Engineering*, vol. 16, no. 3, p. 036015, apr 2019.
- [17] J. Too, A. R. Abdullah, N. Mohd Saad, and W. Tee, "Emg feature selection and classification using a pbest-guide binary particle swarm optimization," *Computation*, vol. 7, no. 1, 2019.
- [18] R. Magin, M. D. Ortigueira, I. Podlubny, and J. Trujillo, "On the fractional signals and systems," *Signal Processing*, vol. 91, no. 3, pp. 350 – 371, 2011, advances in Fractional Signals and Systems.
- [19] Y. FERDI, "Some applications of fractional order calculus to design digital filters for biomedical signal processing," *Journal of Mechanics in Medicine and Biology*, vol. 12, no. 02, p. 1240008, 2012.
- [20] N. Miljković, N. Popović, O. Djordjević, L. Konstantinović, and T. B. Šekara, "Ecg artifact cancellation in surface emg signals by fractional order calculus application," *Computer Methods and Programs in Biomedicine*, vol. 140, pp. 259 – 264, 2017.
- [21] A. Goutas, Y. Ferdi, J.-P. Herbeuval, M. Boudraa, and B. Boucheham, "Digital fractional order differentiation-based algorithm for p and t-waves detection and delineation," *ITBM-RBM*, vol. 26, no. 2, pp. 127 – 132, 2005.
- [22] Y. Ferdi, J. Herbeuval, A. Charef, and B. Boucheham, "R wave detection using fractional digital differentiation," *ITBM-RBM*, vol. 24, no. 5, pp. 273 – 280, 2003.
- [23] J. Wang, Y. Ye, X. Pan, and X. Gao, "Parallel-type fractional zero-phase filtering for ecg signal denoising," *Biomedical Signal Processing and Control*, vol. 18, pp. 36 – 41, 2015.
- [24] L. Cimeša, N. Popović, N. Miljković, and T. B. Šekara, "Heart rate detection: Fractional approach and empirical mode decomposition," in *2017 25th Telecommunication Forum (TELFOR)*, Nov 2017, pp. 1–4.
- [25] J. Wang, Y. Ye, X. Pan, X. Gao, and C. Zhuang, "Fractional zero-phase filtering based on the riemann-liouville integral," *Signal Processing*, vol. 98, pp. 150 – 157, 2014.
- [26] J. Wang, Y. Ye, Y. Gao, S. Qian, and X. Gao, "Fractional compound integral with application to ecg signal denoising," *Circuits, Systems, and Signal Processing*, vol. 34, no. 6, pp. 1915–1930, Jun 2015.
- [27] J. Baranowski, W. Bauer, M. Zagorowska, and P. Piatek, "On digital realizations of non-integer order filters," *Circuits, Systems, and Signal Processing*, vol. 35, no. 6, pp. 2083–2107, 2016.
- [28] A. Kawala-Janik, W. Bauer, M. Żobubak, and J. Baranowski, "Early-stage pilot study on using fractional-order calculus-based filtering for the purpose of analysis of electroencephalography signals," *Studies in Logic, Grammar and Rhetoric*, vol. 47, no. 1, 2016.
- [29] T. Moszkowski and E. Pociask, *Comparison of Fractional- and Integer-Order Filters in Filtration of Myoelectric Activity Acquired from Biceps Brachii*. Heidelberg: Springer International Publishing, 2013, pp. 305–312.
- [30] M. D. Ortigueira, "An introduction to the fractional continuous-time linear systems: the 21st century systems," *IEEE Circuits and Systems Magazine*, vol. 8, no. 3, pp. 19–26, Third 2008.
- [31] A. Kilbas, H. Srivastava, and J. Trujillo, *Theory And Applications of Fractional Differential Equations*, ser. North-Holland Mathematics Studies. Elsevier Science & Tech, 2006.
- [32] K. Oldham and J. Spanier, *The Fractional Calculus Theory and Applications of Differentiation and Integration to Arbitrary Order*, ser. ISSN. Elsevier Science, 1974.
- [33] K. Diethelm, N. Ford, A. Freed, and Y. Luchko, "Algorithms for the fractional calculus: A selection of numerical methods," *Computer Methods in Applied Mechanics and Engineering*, vol. 194, no. 6, pp. 743 – 773, 2005.



Ibrahim Ethem SACU was born in Kayseri, Turkey. He received the B.Sc., M.Sc. and Ph.D. degrees in Electrical-Electronics Engineering from the Erciyes University in 2010, 2012 and 2019, respectively. His current research interests include fractional order circuits, fractional order filters and fractional signal processing.

Stabilization of Chaos in a Cancer Model: The Effect of Oncotripsy

Serpil Yilmaz


Abstract—There has been much interest in the development of therapies for the prevention and treatment of tumours. Recently, the method of oncotripsy has been proposed to destroy cancer cells by applying the ultrasound harmonic excitations at the resonant frequency of cancer cells. In this study, periodic disturbances whose frequency tuned to the fundamental frequency and the higher harmonics of the change in the population of tumor cells are applied to a tumour growth model, respectively, and the appearance of periodic behaviors in a three-dimensional chaotic cancer model is investigated as a result of those harmonic excitations. The numerical results show that by choosing the appropriate values of the parameters of periodic disturbances, the chaotic cancer model induces periodic behaviors such as period-one and two limit cycles which may have important implications on cancer treatment. The results also provide a view to understanding the oncotripsy effect within the framework of stabilization of chaos.

Index Terms—Bifurcation, cancer model, chaos, stabilization.

I. INTRODUCTION

CANCER is a disease that is caused by abnormal or uncontrolled cell growth. Understanding the mechanisms of tumour growth is an actively developing field among many researchers from different disciplines. There are several techniques that are used to perform cancer diagnosis such as X-ray mammography, magnetic resonance imaging, ultrasound technique, digital tomosynthesis, and microwave imaging [1–4]. There are also many mathematical models about tumour-immune dynamics which have been extensively studied to provide valuable insights into the evolution of tumours. A deterministic model in the form of second-order ordinary differential equations governing the populations of immune cells and the tumour cells has been proposed in [5]. There have been many other studies on deterministic tumour-immune models [6–11]. A bifurcation analysis has been developed in [12] to explore the effect of immune response and the treatment strategy of periodically pulsed therapies on tumour-immune dynamics. Pulsed immunotherapy has been extended and a novel hybrid model combining chemotherapy and immunotherapy has been proposed in [13] and the effects of duration, dosage and frequency of combined treatment strategies have been investigated on the tumour population. The above studies have been obtained based on the deterministic tumour-immune model. However, in recent years, stochastic models of tumour-immune systems have been also developed.

Serpil YILMAZ is with Department of Computer Engineering and Artificial Intelligence & Data Science Research and Application Center, İzmir Katip Çelebi University, İzmir, 35620 Türkiye, (e-mail: serpil.yilmaz@ikcu.edu.tr)

 <https://orcid.org/0000-0002-6276-6058>

A tumour-immune model driven by symmetric Lévy noise has been studied in [14]. The effects of noise intensity and stability index have been analyzed on the tumour growth dynamics and the numerical results showed that the effects of Lévy stable noise lead to a decrease of the tumour cells compared with the Gaussian noise. A stochastic mathematical model of the coevolution of immune- and tumour cells has been proposed in [15] by considering both interactions and phenotypic plasticity to help guide the treatment protocols. By considering the variability in cellular reproduction, death and the fluctuation of chemotherapy effect, the deterministic differential equation model has been extended to the stochastic one to analyze the dynamics of tumour cells and immune cells under chemotherapy in [16]. The culling rate of effector cells and the intrinsic growth rate of tumour cells have been modeled as stochastic processes and the effect of environmental noise on the dynamic behaviors of the tumour-immune model has been studied in [17]. A stochastic tumour-immune system with a combination of immunotherapy and chemotherapy has been modeled in [18] and the evolution of tumours has been analyzed in the presence of environmental noise and chemotherapeutic dose. The responses of tumour growth to different drug dosing frequencies have been studied in [19] to improve treatment success.

Despite the various treatment modalities in practice such as radiotherapy, chemotherapy, immunotherapy, or their combination, there is a constant search for alternative modes of treatment for cancer. In [20] a new cancer therapy which is referred to as oncotripsy has been proposed by applying the low-intensity ultrasound waves at specific resonance frequencies. The method is based on the differences in morphologies and material properties between the healthy and tumour cells to selectively target cancer cells. Since the resonant frequency is generally obtained by geometric configurations and the mechanical properties of individual cells such as shape, size and stiffness which are altered in disease then the natural resonant frequencies of healthy cells should be significantly different from those of cancer cells. Therefore, the natural frequency of cells can be used as a tool for specifically targeting the cancer cells while sparing the healthy cells. The normal and cancer cells have been modeled as a sphere linear elastic material in [21] and modal analysis has been carried out to determine the natural frequencies of the cells. Frequency responses of healthy and cancer cells to mechanical stimuli (typically low-intensity therapeutic ultrasound) have been studied in [22] and it has been shown that the discrimination of the normal and tumour cells can be amplified at some ultrasound frequencies. The influence of viscoelasticity on the oncotripsy effect has been studied in [23]. A new perspective for cancer

treatment has been studied in [24] based on utilizing resonance interaction mechanisms between an applied electromagnetic field and the resonant frequency of cancer cells. The effect of thermodynamic resonance in the presence of electromagnetic waves with resonant frequency has been experimentally tested in [25] and it has been shown that cancer cell invasion and proliferation can be decreased with the specific electromagnetic field. The method of oncotripsy has been studied in a panel of breast, colon, and leukemia cancer cell models in [26, 27] by conducting mechanistic experiments. A three-dimensional dynamics of the cell has been developed in [28, 29] and the numerical experiments of the dynamic response of a cell in the presence of ultrasound waves have been presented.

Furthermore, the interactions between tumour and immune cells are nonlinear and extremely complicated which exhibit many properties of chaotic systems. A simple tumour growth model which has chaotic behavior in the parameter range of interest has been developed in [11]. The relationship between tumour size and the chaotic behavior of system dynamics has been examined in [30]. The chaotic behavior has been analyzed in a stochastic cancer model in [31] in which Brownian motion has been used to obtain the corresponding stochastic model with the spread of cancer. As it is well known, even a small disturbance in chaotic systems may lead to a significant change in the nature of the system behavior.

With the above discussions, the aim of this study is to analyze the stabilization of chaotic behavior in the presence of some external periodic disturbances. Our research question is therefore how the dynamics of tumour cells evolve if periodic disturbances are applied at the natural resonant frequency of the change in the population of tumour cells. The effects of fundamental frequency and higher harmonics on the periodic behavior of the cancer model have been also compared. To the best of our knowledge, this is the first study to investigate the effect of oncotripsy in a chaotic cancer model.

The paper is organized as follows. In Section II, the dynamics of a chaotic cancer cell have been modeled in the presence of periodic disturbances. In Section III, the fundamental frequency and higher harmonics of the change in the population of tumour cells have been calculated. The bifurcation analysis has been carried out and the dynamic responses of the model have been obtained numerically. The discussions of the analysis have been presented in Section IV.

II. THE CANCER MODEL IN THE PRESENCE OF PERIODIC DISTURBANCES AT HARMONICS

Periodic disturbances have been applied to the cancer model proposed in [11] which is described by the system of differential equations given as:

$$\begin{aligned} \frac{dT}{dt} &= r_1 T \left(1 - \frac{T}{k_1}\right) - \alpha_{12} T H - \alpha_{13} T E + g(t) + \epsilon \xi(t) \\ \frac{dH}{dt} &= r_2 H \left(1 - \frac{H}{k_2}\right) - \alpha_{21} T H \\ \frac{dE}{dt} &= \frac{r_3 T E}{T + k_3} - \alpha_{31} T E - d_3 E \end{aligned} \quad (1)$$

where $T(t)$ is the number of tumour cells at time t with the growth rate of r_1 and maximum carrying capacity k_1 . $H(t)$ is the number of healthy cells at time t with the growth rate of r_2 and maximum carrying capacity k_2 and $E(t)$ is the number of effector immune cells at time t . The parameters α_{12} and α_{13} denote the tumour cells killing rate by the healthy cells and effector cells, respectively. The parameter r_3 denotes the effector cell production rate in response to the presence of tumour cells. The parameter d_3 is the decay rate of effector cells and k_3 is the half-saturation constant rate for immune production. The rates of inactivation of healthy cells and effector cells by tumour cells are given by the parameters α_{21} and α_{31} , respectively.

The periodic disturbances $g(t)$ have been applied to the tumour growth in (1). The periodic disturbances are modeled in the form of

$$g(t) = \begin{cases} A \sin(2\pi f t), & \sin(2\pi f t) \geq 0 \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

with amplitude A and frequency f . The positive half cycle of periodic signals have been considered since the use of negative half cycles can lead to the negative values of populations which has no biological meaning. Furthermore, the negative cycles of periodic signals have been deliberately modeled by zero intervals which imply the rest period after each intervention.

Since the studies have shown that the impact of environmental factors such as temperature, radiations, oxygen and nutrition are unavoidable on the tumour growth rate, these environmental fluctuations are represented by $\xi(t)$ which is Gaussian white noise satisfying the statistical properties $\langle \xi(t) \rangle = 0$ and $\langle \xi(t) \xi(s) \rangle = \delta(t - s)$ and ϵ is the noise intensity.

III. ANALYSIS AND NUMERICAL RESULTS

The non-dimensionalized cancer model of (1) has been obtained by following the same rescaling in [11] and the corresponding Itô type stochastic differential equation can be written as

$$\begin{aligned} dx_1 &= (x_1(t)(1 - x_1(t)) - \alpha_{12}x_1(t)x_2(t) - \alpha_{13}x_1(t)x_3(t) + \dots \\ &\quad \dots + g(t)) dt + \epsilon dW(t) \\ dx_2 &= (r_2x_2(t)(1 - x_2(t)) - \alpha_{21}x_1(t)x_2(t)) dt \\ dx_3 &= \left(\frac{r_3x_1(t)x_3(t)}{x_1(t) + k_3} - \alpha_{31}x_1(t)x_3(t) - d_3x_3(t) \right) dt \end{aligned} \quad (3)$$

where $W(t)$ is the Wiener process and the effect of changing environmental conditions is achieved by replacing $\xi(t)dt$ with the increments of Wiener process $dW(t)$ which are Gaussian random variables.

The numerical solution of (3) is obtained by using the Euler-Maruyama method. Table I presents the system parameters involved in (3). The parameter values are chosen as in [11] such that with this parameter set the system exhibits chaotic behavior in the absence of periodic disturbances $g(t) = 0$ and the environmental fluctuations $\epsilon = 0$.

TABLE I
THE PARAMETER VALUES OF MODEL (3).

α_{12}	α_{13}	r_2	α_{21}	r_3	k_3	α_{31}	d_3
1	2.5	0.6	1.5	4.5	1	0.2	0.5

The phase portrait of the system (3) for the initial conditions $(x_1(0), x_2(0), x_3(0)) = (0.4, 0.6, 0.1)$ is shown in Fig. 1 when $g(t) = 0$ and $\epsilon = 0$. As seen from Fig. 1 the cancer model has chaotic behavior in the absence of periodic disturbances.

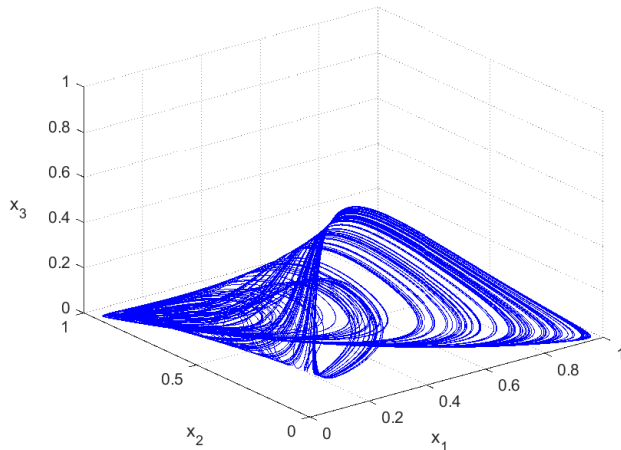


Fig. 1. Chaotic attractor with the initial condition $x_1(0) = 0.4, x_2(0) = 0.6, x_3(0) = 0.1$ in the absence of periodic disturbances and environmental fluctuations.

However, when the external periodic disturbances are allowed to act during a specific period of time, the evolution of cancer cells can be changed. In this study, in particular, periodic signals are applied at the fundamental frequency and harmonics of the change in the population of tumour cells because under small periodic forces at natural frequencies, a condition known as resonance occurs and the system can produce a large-amplitude response. Further, the effect of oncotripsy is discussed by analyzing how the presence of periodic disturbances at the fundamental frequency and harmonics of the change in the population of tumour cells affects the chaotic nature of the system.

The fundamental frequency of the change in the population of tumour cells is estimated by calculating the amplitude spectrum of system state x_1 of (3) for $g(t) = 0$ and $\epsilon = 0$. The amplitude spectrum has a continuous broadband spectrum due to the chaotic nature and the fundamental frequency also known as the first harmonic is indicated as a maximum in the spectrum. The spectrum is also characterized by several peaks at other relatively high frequencies which refer to the harmonic frequencies. From the amplitude spectrum, the fundamental frequency is estimated as 0.022 Hz and higher harmonics are estimated as 0.0356, 0.0572 Hz, 0.069 Hz, and 0.0756 Hz, respectively.

The responses of the system states are obtained by using the Euler-Maruyama scheme with step size 0.01. The numerical

simulations are performed with 500000-time steps for the initial conditions of the system states $x_1(0) = 0.4, x_2(0) = 0.6$ and $x_3(0) = 0.1$ and the noise intensity $\epsilon = 0.0001$. The first 10000-time points are ignored to avoid the transient portion of the data.

The frequency of the periodic signal is set to $f = 0.022$ Hz which is the value of the fundamental frequency of the change in the population of tumour cells. Then to determine the values of amplitude parameter of periodic signal A at which a qualitative change occurs in the dynamics of cancer cells, the bifurcation diagram of the system (3) is obtained by recording the local maxima x_2 for the range of $A = [0, 0.2]$ as shown in Fig. 2. It is seen from Fig. 2 that when the frequency of periodic disturbances is chosen as the fundamental frequency, a transition occurs from chaos to periodicity in the dynamics of the chaotic cancer model.

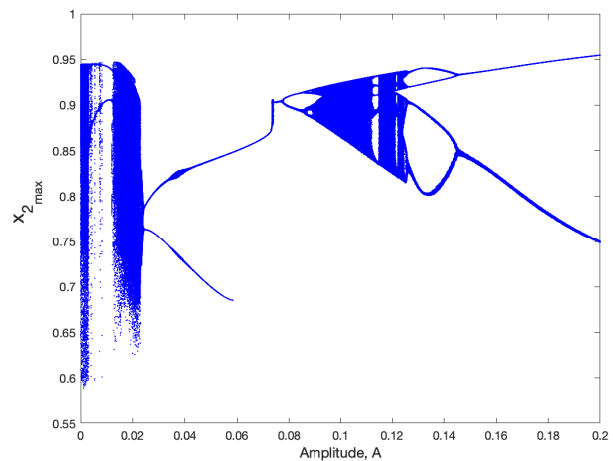


Fig. 2. Bifurcation diagram of model (3) with respect to the bifurcation parameter A . The frequency of periodic signal is chosen as the fundamental frequency of the change in the population of tumour cells, $f = 0.022$ Hz.

The bifurcation diagram in Fig. 2 can be explored in more detail by zooming into the various regions of amplitude A . Figures 3a-3c show the zoomed plots on the particular portions of the bifurcation diagram. It is observed that the system of (3) exhibits rich dynamics such as period-doubling, period-halving and chaotic regions by tuning the amplitude parameter A in the presence of periodic disturbances with the fundamental frequency of the change in the population of tumour cells.

Consider the bifurcation diagram given in Fig. 3a. When the amplitude parameter A is in the range $[0, 0.004)$, the system has chaotic behavior. If the amplitude parameter A exceeds some threshold levels such as $A \approx 0.004$ then the chaos disappears and a period-2 limit cycle occurs. As the amplitude parameter A increases, the periodic solution becomes unstable and the chaos appears again. Then consider the bifurcation diagram given in Fig. 3b. When the amplitude parameter A approaches 0.0230, chaos disappears suddenly and period-8 limit cycles occur. As the amplitude parameter A increases, a very narrow chaotic window appears again with A in the range $[0.0233, 0.0240)$ as shown in Fig. 3b.

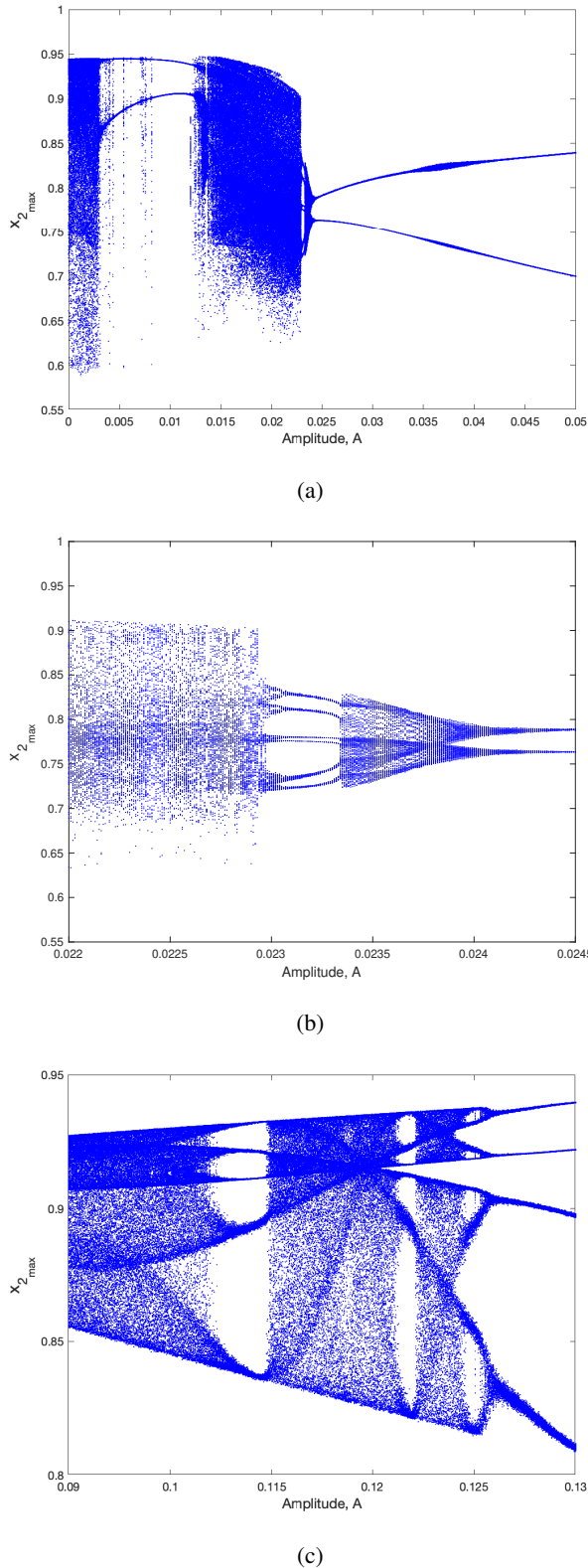


Fig. 3. Bifurcation diagram: Plot of local maxima of x_2 using the amplitude of periodic signal as the bifurcation parameter. The frequency of periodic signal is chosen as the fundamental frequency of the change in the population of tumour cells, $f = 0.022$ Hz. The amplitude of periodic signal is chosen in the range (a) $A = [0, 0.05]$ (b) $A = [0.022, 0.0245]$ (c) $A = [0.09, 0.13]$.

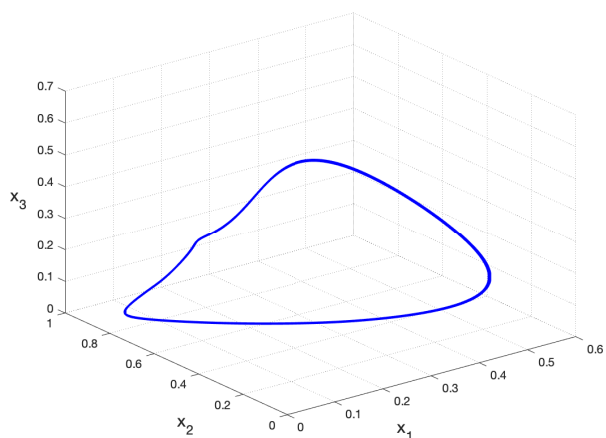
After that, the transition from chaos to the period-2 limit cycle is observed at $A \approx 0.0240$ and the period-2 limit cycle remains in the region which corresponds to the values of A in the range $[0.0240, 0.058)$. A period-halving bifurcation occurs when A approaches 0.058 as shown in Fig. 2 which means that a period-2 limit cycle disappears and a period-1 limit cycle appears at this point.

With the increase of the amplitude parameter A period-doubling bifurcations lead to limit cycles with period 2 and period 4 for $A \in [0.07728, 0.087)$ and $A \in [0.087, 0.09)$, respectively. When A is increased to the range $[0.09, 0.126)$, it can be seen from Fig. 3c that the system (3) has chaotic behavior at a large range of parameter A , except for three narrow ranges $[0.112, 0.1115)$, $[0.1217, 0.122)$ and $[0.124, 0.126)$ which correspond to the limit cycles with periods 5, 6 and 8 respectively. As A approaches 0.0126 the system (3) evolves into a periodic state through the period-halving bifurcations which result in limit cycles with period 4 and period 2 for $A \in [0.126, 0.145)$ and $A \in [0.145, 0.2]$, respectively as shown in Fig. 2.

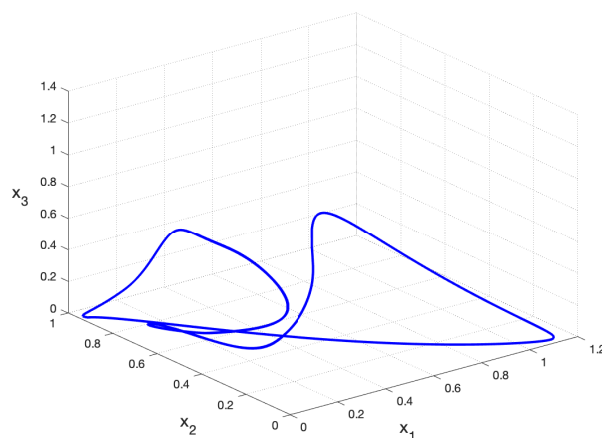
Table II presents the dynamical behaviors of the system (3) with the change of amplitude parameter A which are obtained through the bifurcation analysis. The phase portraits of the system (3) with different amplitude values of A are shown in Fig. 4.

TABLE II
THE DYNAMICS OF SYSTEM (3) FOR THE RANGE OF AMPLITUDE A WHEN THE FREQUENCY IS SET TO $f = 0.022$ Hz.

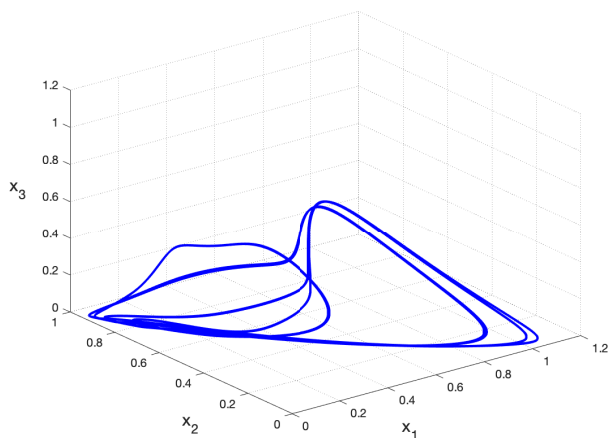
Range of A	Attractor of the system
$[0, 0.004)$	Chaotic
$[0.004, 0.0125)$	Period-2 limit cycle
$[0.0125, 0.0230)$	Chaotic
$[0.0230, 0.0233)$	Period-8 limit cycle
$[0.0233, 0.0240)$	Chaotic
$[0.0240, 0.058)$	Period-2 limit cycle
$[0.058, 0.07728)$	Period-1 limit cycle
$[0.07728, 0.087)$	Period-2 limit cycle
$[0.087, 0.09)$	Period-4 limit cycle
$[0.09, 0.112)$	Chaotic
$[0.112, 0.1115)$	Period-5 limit cycle
$[0.1115, 0.1217)$	Chaotic
$[0.1217, 0.1222)$	Period-6 limit cycle
$[0.1222, 0.124)$	Chaotic
$[0.124, 0.126)$	Period-8 limit cycle
$[0.126, 0.145)$	Period-4 limit cycle
$[0.145, 0.2]$	Period-2 limit cycle



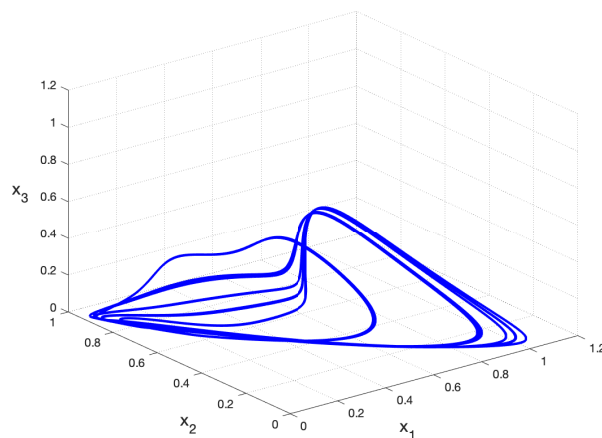
(a)



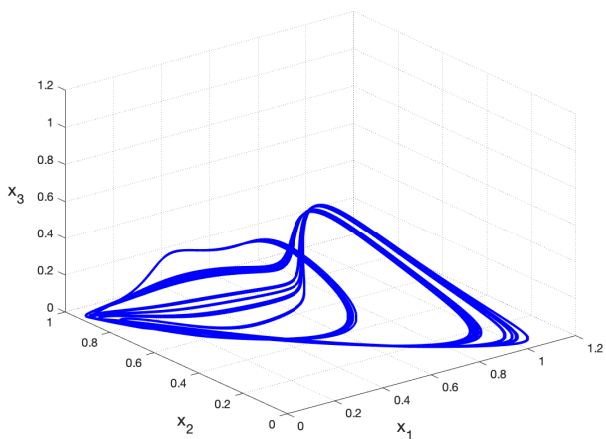
(b)



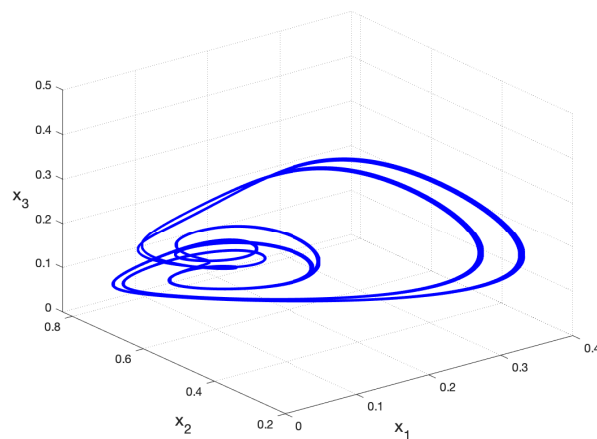
(c)



(d)



(e)



(f)

Fig. 4. 3D phase portrait of system (3) in the presence of periodic disturbances with the frequency $f = 0.022$ Hz and the amplitude (a) $A = 0.064$ (b) $A = 0.2$ (c) $A = 0.13$ (d) $A = 0.114$ (e) $A = 0.122$ (f) $A = 0.125$.

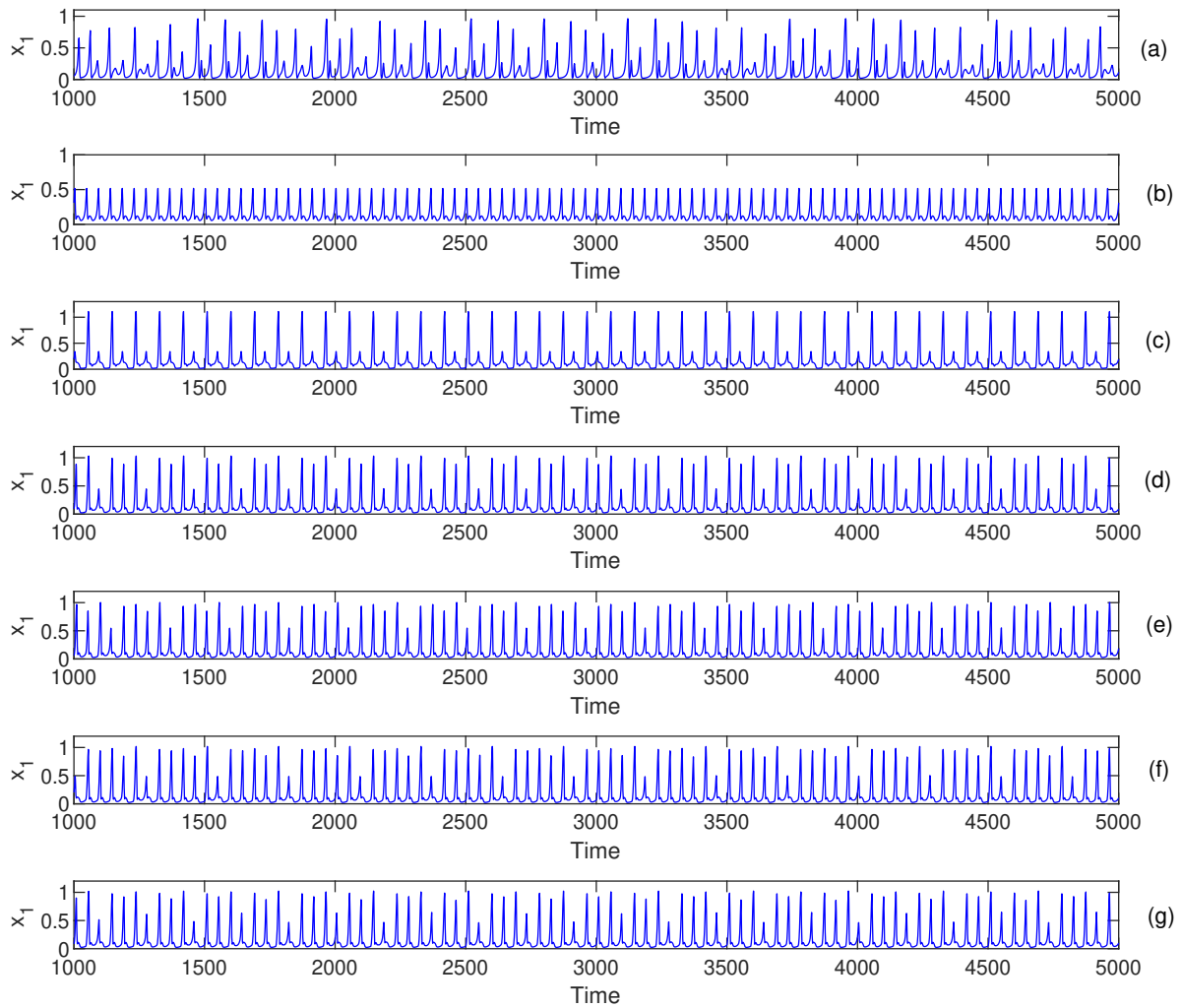


Fig. 5. Time response of the system state x_1 in the absence of periodic disturbances (a) $A = 0$, $\epsilon = 0$ and when the periodic disturbances are present with the frequency $f = 0.022$ Hz, $\epsilon = 0.0001$ and the amplitude (b) $A = 0.064$ (c) $A = 0.2$ (d) $A = 0.13$ (e) $A = 0.114$ (f) $A = 0.122$ (g) $A = 0.125$.

Time responses of the state x_1 are illustrated in Fig. 5 for the initial conditions $x_1(0) = 0.4$, $x_2(0) = 0.6$, $x_3(0) = 0.1$. It is seen from Fig. 5 that the solutions of the system have periodic behaviors when the external periodic disturbances are applied at the fundamental frequency with an appropriate amplitude.

In addition, to analyze the sensitivity to the initial conditions, the parameters of the amplitude and frequency of periodic disturbances are set to $A = 0.064$ and $f = 0.022$ Hz which lead to a period-1 limit cycle behavior of the system. Then the bifurcation diagrams of the system (3) are obtained by changing the initial condition of each state separately as shown in Fig. 6. Each of the initial conditions ($x_1(0)$, $x_2(0)$, $x_3(0)$) are varied at the range $[0.1, 1]$ while the others are set to 0.4. It is seen from Fig. 6 that the period-1 limit cycle behavior of the system remains unchanged with the change of initial conditions $x_1(0)$, $x_2(0)$ and $x_3(0)$.

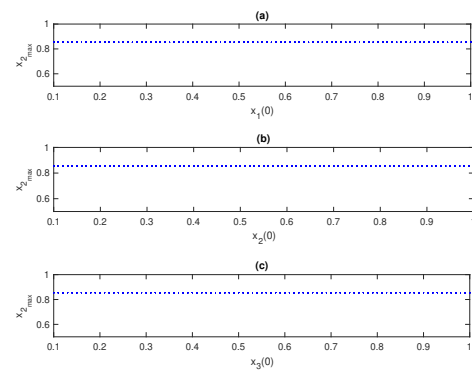


Fig. 6. Bifurcation diagrams with the change of initial conditions: (a) $x_1(0)$ (b) $x_2(0)$ (c) $x_3(0)$.

The frequency of the periodic signal is set to the second harmonics of the change in the population of tumour cells $f = 0.0356$ Hz and the corresponding bifurcation diagram shown in Fig. 7a is obtained by recording the local maxima x_2 for the range of $A = [0, 0.5]$. Fig. 7b shows the zoomed plot of bifurcation diagram on the range of $A = [0, 0.05]$. It can be seen from Fig. 7 that the system of (3) starts from the chaotic state and then suddenly jumps to a periodic state at $A \approx 0.0076$. With the amplitude parameter A in the range $[0.0076, 0.5]$ the system presents a period-doubling bifurcation followed by a period-halving bifurcation. The dynamic behaviors of the system (3) are given in Table III when the amplitude A is changed at the specified ranges.

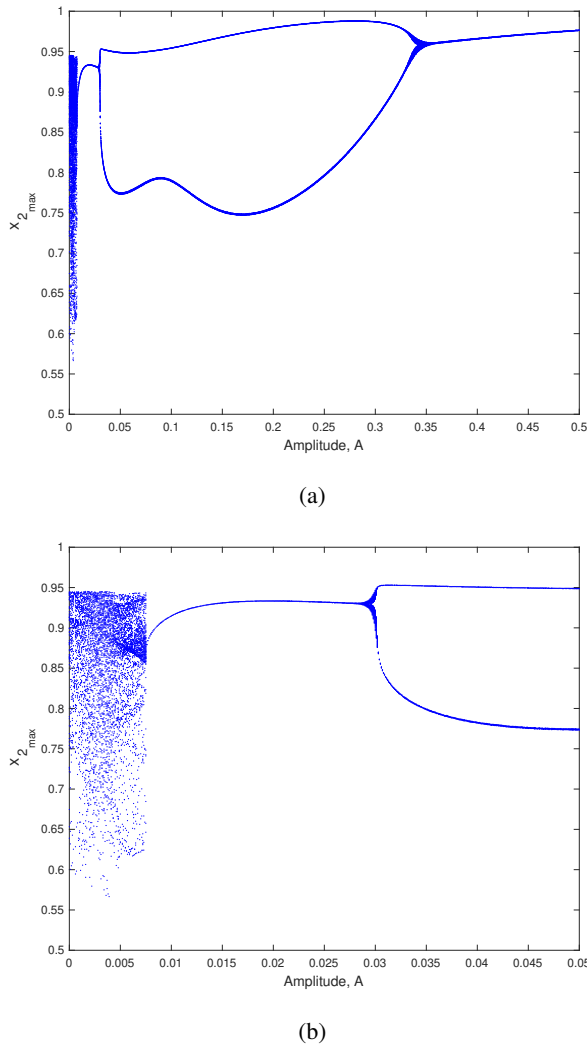


Fig. 7. Bifurcation diagram: Plot of local maxima of x_2 using the amplitude of periodic signal as the bifurcation parameter. The frequency of periodic signal is chosen as the second harmonic of the change in the population of tumour cells $f = 0.0356$ Hz. The amplitude of periodic signal is chosen in the range (a) $A = [0, 0.5]$ (b) $A = [0, 0.05]$.

Figure 8 presents the phase portraits of system (3) in the presence of periodic disturbances with frequency $f = 0.0356$ Hz and amplitude $A = 0.02$ and $A = 0.04$ for period-1 and period-2 limit cycles, respectively.

TABLE III
THE DYNAMICS OF SYSTEM (3) FOR THE RANGE OF AMPLITUDE A WHEN THE FREQUENCY IS SET TO $f = 0.0356$ Hz.

Range of A	Attractor of the system
$[0, 0.0076)$	Chaotic
$[0.0076, 0.0297)$	Period-1 limit cycle
$[0.0297, 0.3412)$	Period-2 limit cycle
$[0.3412, 0.5]$	Period-1 limit cycle

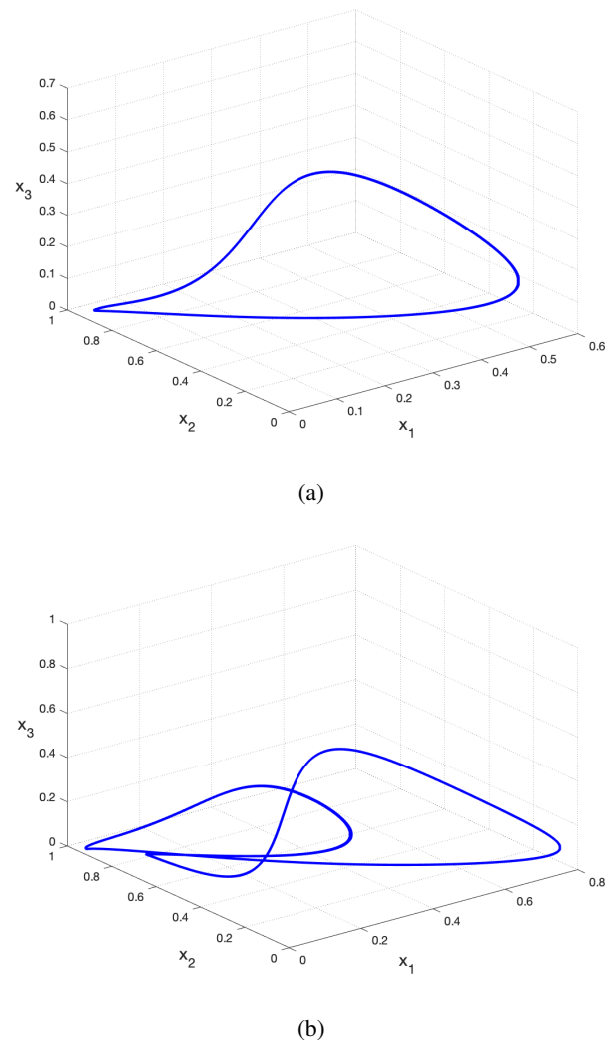
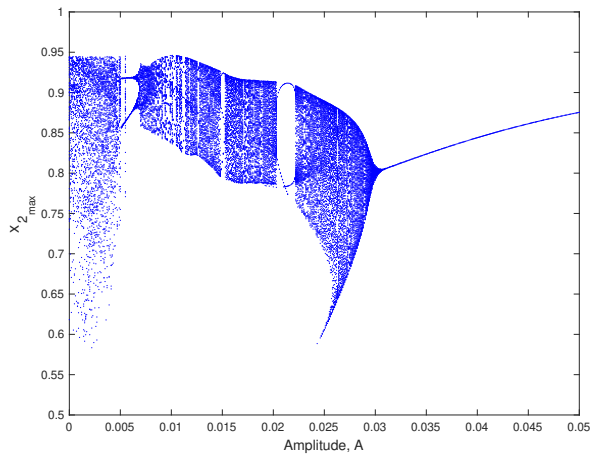
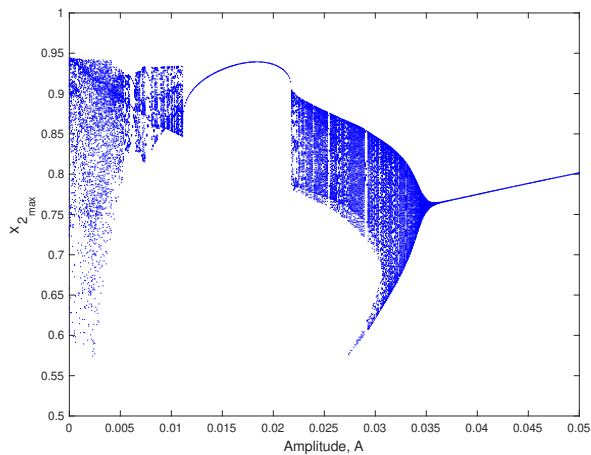


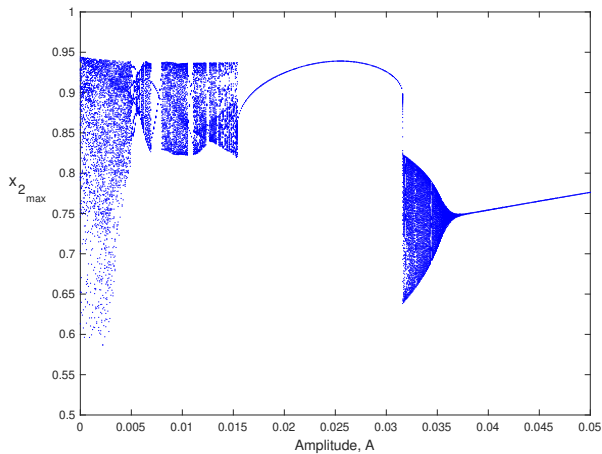
Fig. 8. 3D phase portrait in the presence of periodic disturbances with frequency $f = 0.0356$ Hz and amplitude (a) $A = 0.02$ (b) $A = 0.04$.



(a)



(b)



(c)

Fig. 9. Bifurcation diagram: Plot of local maxima of x_2 using the amplitude of periodic signal as the bifurcation parameter. The frequency of periodic signal is chosen as the (a) 3rd harmonic $f = 0.0572$ Hz (b) 4th harmonic $f = 0.069$ Hz and (c) 5th harmonic $f = 0.0756$ Hz.

TABLE IV
THE DYNAMICS OF SYSTEM FOR THE RANGE OF AMPLITUDE A WHEN THE FREQUENCY IS SET TO $f = 0.0572$ Hz.

Range of A	Attractor of the system
$[0, 0.0052)$	Chaotic
$[0.0052, 0.0066)$	Period-2 limit cycle
$[0.0066, 0.0149)$	Chaotic
$[0.0149, 0.0152)$	Period-2 limit cycle
$[0.0152, 0.0204)$	Chaotic
$[0.0204, 0.022)$	Period-2 limit cycle
$[0.022, 0.0309)$	Chaotic
$[0.0309, 0.05]$	Period-1 limit cycle

TABLE V
THE DYNAMICS OF SYSTEM FOR THE RANGE OF AMPLITUDE A WHEN THE FREQUENCY IS SET TO $f = 0.069$ Hz.

Range of A	Attractor of the system
$[0, 0.006)$	Chaotic
$[0.006, 0.0063)$	Period-3 limit cycle
$[0.0063, 0.0075)$	Chaotic
$[0.0075, 0.0079)$	Period-4 limit cycle
$[0.0079, 0.0112)$	Chaotic
$[0.0112, 0.0218)$	Period-1 limit cycle
$[0.0218, 0.0375)$	Chaotic
$[0.0375, 0.05]$	Period-1 limit cycle

TABLE VI
THE DYNAMICS OF SYSTEM FOR THE RANGE OF AMPLITUDE A WHEN THE FREQUENCY IS SET TO $f = 0.0756$ Hz.

Range of A	Attractor of the system
$[0, 0.007)$	Chaotic
$[0.007, 0.0079)$	Period-2 limit cycle
$[0.0079, 0.0155)$	Chaotic
$[0.0155, 0.0315)$	Period-1 limit cycle
$[0.0315, 0.0387)$	Chaotic
$[0.0387, 0.05]$	Period-1 limit cycle

When the periodic disturbances are applied at the higher harmonics up to the 5th harmonic, the corresponding bifurcation diagrams for the amplitude parameter A in the range $[0, 0.05]$ as shown in Fig. 9. It can be seen from Fig. 9 that the system (3) evolves from chaotic into a periodic state by tuning the amplitude of periodic disturbances. It can be also seen that chaotic regions contain some very narrow periodic windows and when the frequency is set to the higher harmonics there are successive transitions from chaotic to periodic state or periodic to chaotic state as the amplitude

of the periodic disturbances gradually increases. Through bifurcation analysis, the amplitude of periodic disturbances to observe the periodic dynamics such as period-1 limit cycle can be also clearly specified.

Table IV, Table V and Table VI present the ranges of the amplitude A which correspond to the regions of periodic and chaotic behaviors when the frequency of periodic disturbances is set to $f = 0.0572$ Hz, $f = 0.069$ Hz and $f = 0.0756$ Hz, respectively.

Figure 10 shows the phase portraits of the system (3) when the amplitude of periodic disturbances is set to $A = 0.04$ and the frequency of periodic disturbances is tuned to the values of 3rd, 4th and 5th harmonics. Figure 11 shows the structures of period-2 limit cycles when the periodic disturbances are applied at 3rd and 5th harmonics with a sufficient amplitude such as $A = 0.021$ and $A = 0.0075$, respectively.

It is seen from Fig. 10 that the system exhibits period-1 limit cycles for those frequencies and a change in the frequency results in a change in the amplitude of the limit cycles. It is also observed that not only the amplitude of the limit cycle but also the period of the limit cycle can be changed with the change of frequency of periodic disturbances.

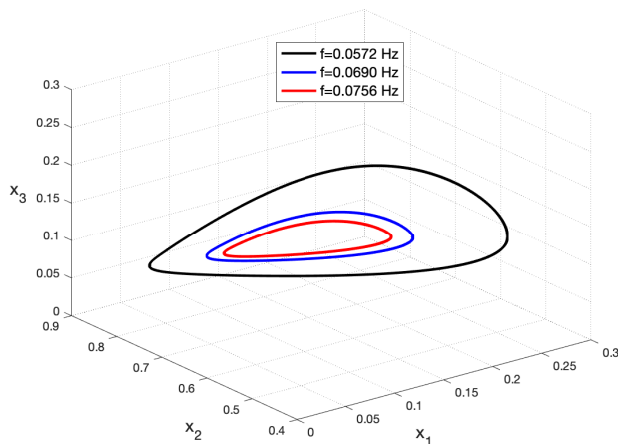
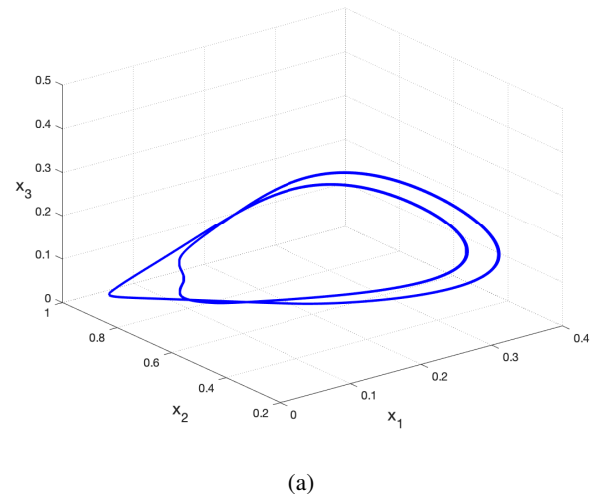


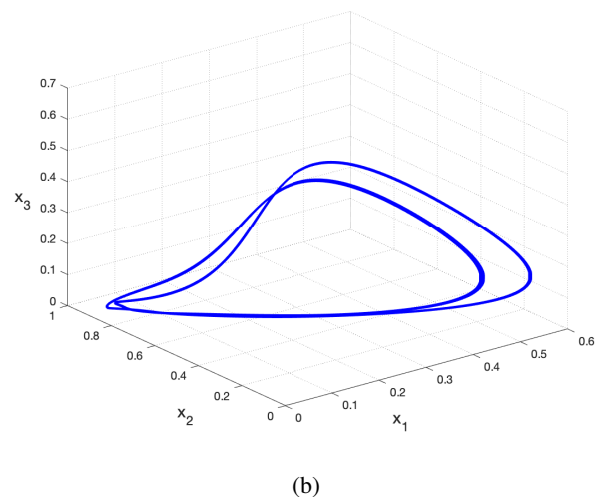
Fig. 10. 3D phase solution in the presence of periodic disturbances with amplitude $A = 0.04$ and the frequency $f = 0.0572$ Hz (black line); $f = 0.069$ Hz (blue line); $f = 0.0756$ Hz (red line).

Specifically, by choosing the frequency of periodic disturbances as 1st or 2nd harmonic the system exhibits period-2 limit cycles at $A = 0.04$. Thus, the system exhibits different sensitivities to the same amplitude of periodic disturbances depending on the frequency of periodic disturbances.

Therefore, it can be concluded that the chaotic cancer model (3) is highly dependent on the frequency of periodic disturbances. This just ensures that if the tumour cells are disturbed with a periodic signal at the fundamental frequency or harmonics, the dynamics of the system can be switched from chaotic to periodic behavior.



(a)



(b)

Fig. 11. 3D phase solution in the presence of periodic disturbances when the frequency f and amplitude A are set to (a) $f = 0.0572$ Hz and $A = 0.021$ (b) $f = 0.0756$ Hz and $A = 0.0075$.

IV. CONCLUSION

In this study, the method of oncotripsy is discussed to stabilize chaotic behavior in a cancer model. To analyze the effect of the resonance phenomenon on the chaotic cancer model, the fundamental and higher harmonic frequencies of the change in the tumour cell population have been calculated, and the external periodic disturbances have been applied at those frequencies to the tumour growth. The amplitude of periodic disturbances has been considered as the bifurcation parameter, and the associated bifurcation diagrams have been obtained for each of the harmonics. The numerical results have been utilized to obtain the dynamic responses of the system.

It has been observed that when the periodic disturbances are applied at the fundamental frequency, the chaotic system can be stabilized to a limit cycle with period-one. Moreover, when the frequency is set to the fundamental frequency of the change in the population of tumour cells, the system may have rich dynamic behaviors such as limit cycles exhibiting different periods just by tuning the amplitude parameter. An

interesting phenomenon has been observed when the frequency of the periodic disturbances is chosen as the second harmonics of the change in the population of tumour cells. In this case, the system has a very narrow chaotic region, and as the amplitude of periodic disturbances exceeds some low level, then the chaos disappears, and the periodic behaviors occur. Furthermore, when the frequency is set to higher harmonics, it has been observed that the dynamics of the system can be switched between chaotic and periodic states with the change of the amplitude parameter.

Finally, it is important to note that our results also provide insights on a possible effect of oncotripsy for the stabilization of chaos in a cancer model. Periodic disturbances acting on a chaotic cancer model are capable of exhibiting periodic behaviors, which may help adjust treatment regimes. Therefore, it will be extremely important to understand the parameters and conditions that lead to limit cycles.

DATA AVAILABILITY

All figures were generated using MATLAB R2021b.

REFERENCES

- [1] S. G. Orel and M. D. Schnall, "Mr imaging of the breast for the detection, diagnosis, and staging of breast cancer," *Radiology*, vol. 220, no. 1, pp. 13–30, 2001.
- [2] C. K. Kuhl, S. Schrading, C. C. Leutner, N. Morakkabati-Spitz, E. Wardelmann, R. Fimmers, W. Kuhn, and H. H. Schild, "Mammography, breast ultrasound, and magnetic resonance imaging for surveillance of women at high familial risk for breast cancer," *Journal of clinical oncology*, vol. 23, no. 33, pp. 8469–8476, 2005.
- [3] S. Vedantham, A. Karellas, G. R. Vijayaraghavan, and D. B. Kopans, "Digital breast tomosynthesis: state of the art," *Radiology*, vol. 277, no. 3, pp. 663–684, 2015.
- [4] A. R. Celik, M. B. Kurt, and S. Helhel, "An experimental performance investigation of an ultra-wideband directional antenna in the microwave imaging of breast cancer tumor," *The Applied Computational Electromagnetics Society Journal (ACES)*, pp. 1549–1560, 2019.
- [5] V. A. Kuznetsov, I. A. Makalkin, M. A. Taylor, and A. S. Perelson, "Nonlinear dynamics of immunogenic tumors: parameter estimation and global bifurcation analysis," *Bulletin of mathematical biology*, vol. 56, no. 2, pp. 295–321, 1994.
- [6] D. Kirschner and J. C. Panetta, "Modeling immunotherapy of the tumor-immune interaction," *Journal of mathematical biology*, vol. 37, no. 3, pp. 235–252, 1998.
- [7] V. A. Kuznetsov and G. D. Knott, "Modeling tumor regrowth and immunotherapy," *Mathematical and Computer Modelling*, vol. 33, no. 12-13, pp. 1275–1287, 2001.
- [8] L. G. De Pillis and A. Radunskaya, "The dynamics of an optimally controlled tumor model: A case study," *Mathematical and computer modelling*, vol. 37, no. 11, pp. 1221–1244, 2003.
- [9] A. d'Onofrio, "A general framework for modeling tumor-immune system competition and immunotherapy: Mathematical analysis and biomedical inferences," *Physica D: Nonlinear Phenomena*, vol. 208, no. 3-4, pp. 220–235, 2005.
- [10] L. G. de Pillis, W. Gu, and A. E. Radunskaya, "Mixed immunotherapy and chemotherapy of tumors: modeling, applications and biological interpretations," *Journal of theoretical biology*, vol. 238, no. 4, pp. 841–862, 2006.
- [11] M. Itik and S. P. Banks, "Chaos in a three-dimensional cancer model," *International Journal of Bifurcation and Chaos*, vol. 20, no. 01, pp. 71–79, 2010.
- [12] H.-C. Wei and J.-T. Lin, "Periodically pulsed immunotherapy in a mathematical model of tumor-immune interaction," *International Journal of Bifurcation and Chaos*, vol. 23, no. 04, p. 1350068, 2013.
- [13] J. Yang, S. Tang, and R. A. Cheke, "Modelling pulsed immunotherapy of tumour-immune interaction," *Mathematics and Computers in Simulation*, vol. 109, pp. 92–112, 2015.
- [14] Y. Xu, J. Feng, J. Li, and H. Zhang, "Stochastic bifurcation for a tumor-immune system with symmetric lévy noise," *Physica A: Statistical Mechanics and its Applications*, vol. 392, no. 20, pp. 4739–4748, 2013.
- [15] M. Baar, L. Coquille, H. Mayer, M. Hölzel, M. Rogava, T. Tüting, and A. Bovier, "A stochastic model for immunotherapy of cancer," *Scientific reports*, vol. 6, no. 1, pp. 1–10, 2016.
- [16] X. Liu, Q. Li, and J. Pan, "A deterministic and stochastic model for the system dynamics of tumor-immune responses to chemotherapy," *Physica A: Statistical Mechanics and its Applications*, vol. 500, pp. 162–176, 2018.
- [17] X. Li, G. Song, Y. Xia, and C. Yuan, "Dynamical behaviors of the tumor-immune system in a stochastic environment," *SIAM Journal on Applied Mathematics*, vol. 79, no. 6, pp. 2193–2217, 2019.
- [18] J. Yang, Y. Tan, and R. A. Cheke, "Modelling effects of a chemotherapeutic dose response on a stochastic tumour-immune model," *Chaos, Solitons & Fractals*, vol. 123, pp. 1–13, 2019.
- [19] P. Schulthess, V. Rottschäfer, J. W. Yates, and P. H. van Der Graaf, "Optimization of cancer treatment in the frequency domain," *The AAPS journal*, vol. 21, no. 6, p. 106, 2019.
- [20] S. Heyden and M. Ortiz, "Oncotripsy: Targeting cancer cells selectively via resonant harmonic excitation," *Journal of the Mechanics and Physics of Solids*, vol. 92, pp. 164–175, 2016.
- [21] S. K. Jaganathan, A. P. Subramanian, M. V. Vellayappan, A. Balaji, A. A. John, A. K. Jaganathan, and E. Supriyanto, "Natural frequency of cancer cells as a starting point in cancer treatment," *Current Science*, pp. 1828–1832, 2016.
- [22] M. Fraldi, A. Cugno, L. Deseri, K. Dayal, and N. Pugno, "A frequency-based hypothesis for mechanically targeting and selectively attacking cancer cells," *Journal of the Royal Society Interface*, vol. 12, no. 111, p. 20150656, 2015.
- [23] S. Heyden and M. Ortiz, "Investigation of the influence of viscoelasticity on oncotripsy," *Computer Methods in*

Applied Mechanics and Engineering, vol. 314, pp. 314–322, 2017.

- [24] E. Calabrò and S. Magazù, “New perspectives in the treatment of tumor cells by electromagnetic radiation at resonance frequencies in cellular membrane channels,” *The Open Biotechnology Journal*, vol. 13, no. 1, 2019.
- [25] U. Lucia, G. Grisolia, A. Ponzetto, L. Bergandi, and F. Silvagno, “Thermomagnetic resonance affects cancer growth and motility,” *Royal Society open science*, vol. 7, no. 7, p. 200299, 2020.
- [26] D. R. Mittelstein, J. Ye, E. F. Schibber, A. Roychoudhury, L. T. Martinez, M. H. Fekrazad, M. Ortiz, P. P. Lee, M. G. Shapiro, and M. Gharib, “Selective ablation of cancer cells with low intensity pulsed ultrasound,” *Applied Physics Letters*, vol. 116, no. 1, p. 013701, 2020.
- [27] D. R. Mittelstein, “Modifying ultrasound waveform parameters to control, influence, or disrupt cells,” Ph.D. dissertation, California Institute of Technology, 2020.
- [28] E. Schibber, D. Mittelstein, M. Gharib, M. Shapiro, P. Lee, and M. Ortiz, “A dynamical model of oncotripsy by mechanical cell fatigue: selective cancer cell ablation by low-intensity pulsed ultrasound,” *Proceedings of the Royal Society A*, vol. 476, no. 2236, p. 20190692, 2020.
- [29] E. Figueroa-Schibber, “High-cycle dynamic cell fatigue with applications on oncotripsy,” Ph.D. dissertation, California Institute of Technology, 2020.
- [30] S. Abernethy and R. J. Gooding, “The importance of chaotic attractors in modelling tumour growth,” *Physica A: Statistical Mechanics and its Applications*, vol. 507, pp. 268–277, 2018.
- [31] M. Fahimi, K. Nouri, and L. Torkzadeh, “Chaos in a stochastic cancer model,” *Physica A: Statistical Mechanics and its Applications*, vol. 545, p. 123810, 2020.

BIOGRAPHY



Serpil Yılmaz is currently an assistant professor of Computer Engineering at İzmir Katip Çelebi University. She received her M.Sc. and Ph.D. degrees in Electronics and Communication Engineering from İzmir Institute of Technology in 2012 and 2019, respectively. Her research interests include chaotic

systems, stochastic models, artificial intelligence and embedded systems.

Assessment of Epileptic Seizures and Non-Epileptic Seizures via Wearable Sensors and Priori Detection of Epileptic Seizures


Ömer Faruk Ertuğrul, Yasin Sönmez, Necmettin Sezgin and Eşref Akil

Abstract— Epilepsy is one of the most prevalent neurological disorders whose causes are not exactly known. Diagnosis and treatment of epilepsy are closely related to the patient's story, and the most important indicator is the frequency and severity of seizures. Since the disease does not only affect the patients but also the lives of their environment seriously, it is very important to make the diagnosis and treatment correctly. However, sometimes misrecognition from patients and their relatives, unnecessary epilepsy treatment to the patient in non-epileptic seizures mixed with epileptic seizures, or increasing the dose of the drugs used for the patient are the situations frequently encountered.


The so-called video-EEG method is used in the detection and segregation of epileptic / non-epileptic seizures. In this method, the patient is kept in an environment where video recording is continuously taken until the seizure occurs, and EEG, EMG, and ECG records of the patient are taken. When the patient has a seizure, the seizure type is separated by examining these records. In this project, seizure detection and seizure type (epileptic / non-epileptic) detection is aimed to be done by using wearable sensors increasingly applied in the field of health. The achievable benefits from the project and data set will provide a different perspective on the epilepsy illness, as well as reduce the number of epilepsy patients who are not in fact epilepsy patients needing treatment, and keep epileptic seizure recordings constantly in the electronic environment so that the treatment processes are monitored more closely.

Index Terms— Epilepsy; Epileptic Seizures; Non-Epileptic Seizures; Priori Detection of Epileptic Seizures

ÖMER FARUK ERTUĞRUL, is with Department of Electrical Engineering University of Batman University, Batman, Turkey, (e-mail: omerfaruk.ertugrul@batman.edu.tr).

 <https://orcid.org/0000-0003-0710-0867>


YASİN SÖNMEZ, is with Department of Computer Technologies University of Batman University, Batman, Turkey, (e-mail: yasin.sonmez@batman.edu.tr).

 <https://orcid.org/0000-0001-9303-1735>

NECMETTİN SEZGİN, is with Department of Electrical Engineering University of Batman University, Batman, Turkey, (e-mail: necmettin.sezgin@batman.edu.tr).

 <https://orcid.org/0000-0002-2306-6008>

EŞREF AKIL, is with Department of Medicine University of Dicle University, Diyarbakır, Turkey, (e-mail: esrefakil@gmail.com).

 <https://orcid.org/0000-0001-9669-6804>

Manuscript received January 17, 2022; accepted March 5, 2022.

DOI: [10.17694/bajece.1054818](https://doi.org/10.17694/bajece.1054818)

I. INTRODUCTION

EPILEPSY is one of the most common neurological diseases in society and is generally analyzed using EEG signals (Sezgin et al., 2013; Kaya et al., 2012; Proix et al., 2021). According to WHO, the prevalence of epilepsy disease is in the range of 0.4-1% (Çakıl et al., 2013, Kanas et al., 2015). Epilepsy can be divided into two sub-groups and the reason behind the more common subgroup, which is the idiopathic epilepsy subgroup, is unknown. The other subgroup is based on some physiological or hereditary reasons. Since epilepsy is a disease that requires long-term treatment, its diagnosis has great importance and the most important symptom for diagnosis is seizures. However, in addition to epileptic seizures, which is associated with simultaneous neuronal activity, non-epileptic seizures (pseudo-epileptic, psychogenic, or physiological seizures), which are not associated with dysfunction of the central nervous system, may present symptoms similar to epileptic seizures and make it difficult to diagnose (Arıkanoğlu, 2011; Kanas et al., 2015; Lai et al., 2022). An epileptic and non-epileptic EEG signals, which were recorded from epileptic and non-epileptic patients in Dicle University, Faculty of Medicine, Department of Neurology, were given in Figure 1(a), and 1(b), respectively.

As seen in Figure 1(a) and 1(b), the EEG signal, which can be taken during a seizure is enough to classify it as an epileptic or non-epileptic seizure, and the main difference between epileptic and non-epileptic seizures is that epileptic seizures are accompanied by neuronal discharges, on the other hand, no special changes are observed in EEG in non-epileptic seizures (Arıkanoğlu, 2011; Tatlı, 2004; Cobb and Beebe, 2022).

However, the most important problems in differentiating epileptic and non-epileptic seizures are that the patient comes to the medical doctor after having a seizure and sometimes non-epileptic seizures occur in patients who have had epileptic seizures (Kotsopoulos et al., 2003; D'Alessio et al., 2006). For this reason, the first diagnosis is mostly made according to the patient and his environment (medical history) and the age of the patient (Çakıl et al., 2013; Arıkanoğlu, 2011; Tatlı, 2004), since the EEG recordings of the patient cannot be obtained at the time of the seizure.

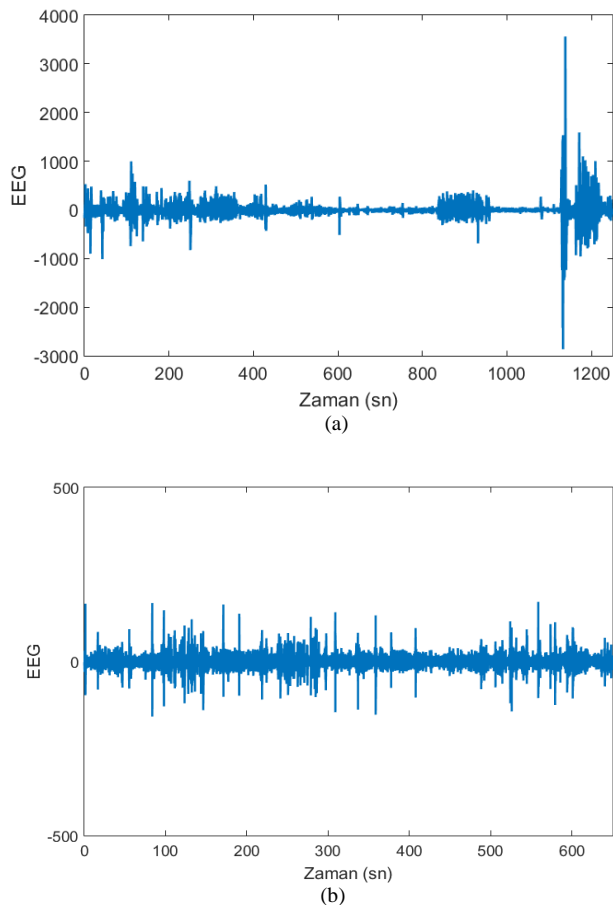


Figure 1.(a) EEG signal change during epileptic seizure process, (b) EEG signal change during non-epileptic seizure process

Bodde et al. (2009) stated in their review study that it is not easy to distinguish between epileptic and non-epileptic seizures, and the cost per patient of misdiagnosis is around \$231,432. It has been stated that the estimated annual loss from misdiagnosis of epilepsy in the UK alone is around £185 million (Kanas et al., 2015). It has been reported that 10-20% of patients who applied to a physician with the suspicion of having epileptic seizures were not epileptic (Cuthill and Espie, 2005). It was stated that in case of misdiagnosis (diagnosis of epilepsy in a patient with non-epileptic seizures), the treatment was stopped based on the patient's response to epileptic drugs, and this process has established an average of 7.2 years after the patient started the treatment (Çakıl et al., 2013; Reuber and Elger, 2003).

Considering that many unnecessary antiepileptic drugs are given to the patient during this wrong treatment process, many side effects occur in patients due to these drugs, it causes significant labor, time, and cost, and the patient's quality of life decreases, it is seen how important it is to differentiate non-epileptic seizures from epileptic seizures in the early period. (Çakıl et al., 2013; Arıkanoğlu, 2011; Kanas et al., 2015).

For these reasons, it is vital to detect epileptic and non-epileptic seizures with a good differential method (Çakıl et al., 2013; Arıkanoğlu, 2011; Araz et al., 2009; Buchanna et al.,

2022). Many EEG-based differential tests have been recommended for differential diagnosis due to their low cost, and success rates of 60-90% have been reported with these tests (Çakıl et al., 2013). On the other hand, no adequate method has been proposed to differentiate epileptic and non-epileptic seizures except video-EEG (Kotsopoulos et al., 2003; Bodde et al., 2009; Cuthill and Espie, 2005; Reuber, 2008; Çakıl et al., 2013). Video-EEG, on the other hand, is an expensive and long-term test routine that bothers the patient. In order to obtain distinctive results in video-EEG, the patient must have had a seizure during the recording process. In addition to video-EEG, age-related autonomic nervous system tests, serum prolactin levels, neurocardiogenic and psychological tests are also used for this purpose (Çakıl et al., 2013; Arıkanoğlu, 2011; Tatlı, 2004; Araz et al., 2009; Bodde et al., 2009; Cuthill and Espie, 2005; Cragar et al., 2002; Reuber and Elger, 2003; Reuber, 2008). These suggested tests are based on the statistical analysis of the physiological characteristics of the patients or the test results obtained (Kotsopoulos et al., 2003; D'Alessio et al., 2006; Turner et al., 2011). For example, patients' response to epileptic drugs, seizure frequency, genetic epileptic history, seizure characteristics, physiological and neurophysiological tests, and similar parameters were used for this purpose (Cragar et al., 2002; Reuber and Elger, 2003; Reuber, 2008).

With the decrease in size and cost of wearable sensors, they have started to be used in many areas, including health applications (Akbulut and Akan, 2015; Worrell et al., 2021). It has been reported that successful results were obtained by using signals from wearable sensors (especially accelerometers) during the seizure process as a differential diagnosis (Gubbi et al., 2015; Kusmakar et al., 2015-a, 2015-b). The methods recommended in the literature and the success rates obtained for the differentiation of epileptic and non-epileptic seizures are summarized in Table 1.

Table 1. Success rates in differentiating epileptic and non-epileptic seizures

Reference	Seizures	Data Type	Method	Accuracy
Kanas et al. (2015)	126	EEG	Support Vector Machines	%96
Xu et al. (2014)	25	EEG	Support Vector Machines	%92
Pippa et al. (2014)	205	EEG	Bayes	%86
Gubbi et al. (2015)	27	Accelerometer	k Means and Support Vector Machines	%93.3
Kusmakar et al. (2015-a)	34	Accelerometer	Support Vector Machines	% 89
Kusmakar et al. (2015-b)	34	Accelerometer	Support Vector Machines	%80

When Table 1 is examined, it is observed that differential diagnosis is achieved with high success when EEG signals are used. However, it is not easy to obtain a differential diagnosis with this method, since it is difficult to predict the moment of seizure and the EEG signals received for differential diagnosis must be at the time of seizure. On the other hand, wearable sensors are used for many different purposes in integration with many systems such as smartwatches or mobile phones without reducing the comfort of life.

The main reason that motivates us for this study is to provide a test routine for the differential diagnosis of epilepsy that will cause less discomfort to the patient and their relatives and can be applied at a low cost. The development of such a test method will not only differentiate between epileptic and non-epileptic seizures but will also enable the digital recording of seizures and seizure characteristics, opening the door to many possibilities that could improve the quality of life of epileptic patients.

In addition, seizure detection (Gonzalez-Velldn et al., 2003; Yadav et al., 2009; Hussain et al., 2021) is an important process for the detection and treatment of the disease, as well as an important output to improve the quality of life of the patient and their relatives. Because the relatives of the patients put limits on the patient's and their own economic and social lives due to the patient's desire to be taken into consideration, as the patient has the possibility of having a seizure, and naturally their quality of life decreases. Detecting the moment of epilepsy of the patient will open the way of informing/warning the relatives of the patient with simple software, and the number and characteristics of seizures also contain important information in terms of the course of the disease.

II. MATERIAL AND METHODS

A. Dataset

Patients who routinely apply to the Dicle University Hospital, Department of Neurology for diagnosis and treatment and whose seizures cannot be differentiated as epileptic or non-epileptic seizures are subjected to video-EEG testing. In this test, both EEG, ECG, and EMG recordings of a total of 88 patients are taken and they are recorded with the video. In this process, after the patient has a seizure, the images of this seizure process and the recorded EEG, ECG, and EMG signals are examined, and a diagnosis is made about whether the patient has epilepsy or not.

In addition to this routine test, additional information from the patient was recorded with a sensor node to be attached to each leg and arm of the patient, as seen in Figure 2. Obtained sensor records were recorded on the computer. In addition, each record was recorded by associating the diagnosis of the expert neurologist with the record in question, and the data set used in the study was formed.



Figure 2. Obtained sensor records were recorded on the computer.

In summary, while creating the data set, 18-channel EEG, ECG, and EMG signals received during the routine video-EEG tests taken from the subjects during epileptic and non-epileptic seizures, signals received from wearable sensors attached to the arms and legs, and the diagnosis of the specialist neurologist (epileptic or non-epileptic seizure) information. has been recorded in accordance with the permission of the regional ethical committee (Per. no.: 2016/352, Provider: Dicle University Ethical Committee).

B. Method

Wearable sensors have started to be used in many areas with the decrease in the size, cost, and energy needs of the sensors and the developments in communication and processor technologies. In the literature review, it was seen that wearable sensors were used successfully in detecting seizures and seizure types (Nei, 2009; Zijlmans et al., 2002; Behbahani et al., 2012). The connection of the sensors is shown in Figure 3.



Figure 3. The connection of the sensors.

In addition, the data from the sensor nodes were transferred to the portable computer with an Arduino Due and RF transceiver. The energy needs of the sensor nodes are met by a rechargeable lipo battery. In Figure 4, there is a sample epileptic seizure and EEG signal of non-epileptic seizure process taken from patients treated in Dicle University, Faculty of Medicine, Department of Neurology.

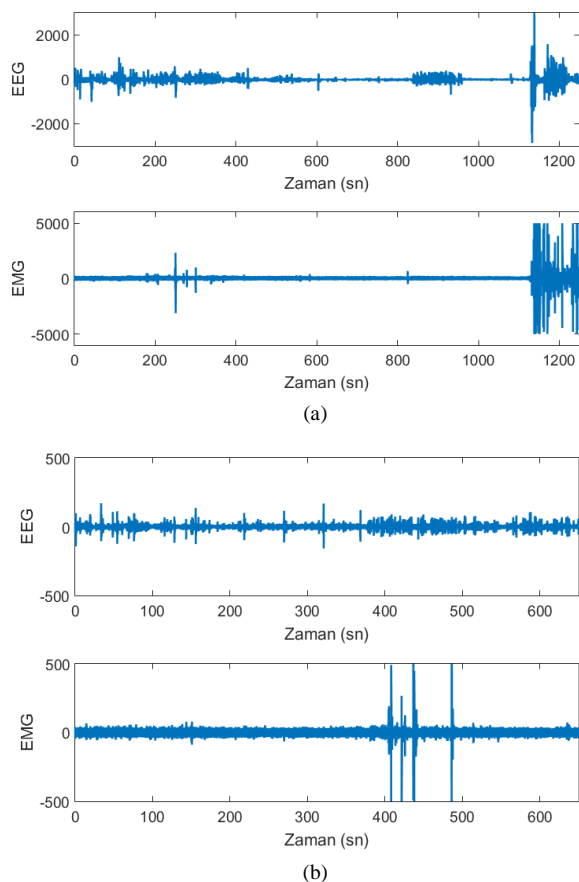


Figure 4. Example for Seizure Segregation (a) Epileptic Seizure, (b) Non-Epileptic Seizure

For the purpose of seizure detection, a priori seizures detection, and seizures type detection:

1. Extracting features from the signal,
2. Determining the features that are active for each target of the extracted features,
3. Classification with artificial intelligence, stages were carried out.

III. RESULTS AND DISCUSSION

Statistical features extracted from the processed signals are classified by artificial intelligence. All classification results were obtained and reported by 5-part cross-validation. First of all, the artificial intelligence method used in all experiments was optimized and the obtained accuracy results were shared. The following classification processes were performed in the simulations

- Seizures Detection: It is the determination of whether the patient has seizures (epileptic or non-epileptic).
- Seizures Type Detection: Separation of epileptic / non-epileptic seizures,
- Classification of all data: Normal/Epileptic/Non-Epileptic signals,

As an example of optimization, the results obtained in the process of seizures detection of EEG, ECG + EMG signals, and wearable sensors in detecting seizure types, and separation of all data are summarized in Table 2.

Table 3. Accuracies (%) in the detection of a seizure before it occurs

Dataset	Detection of Seizures	Detection of Seizure Type	Classification of Signal
EEG	100	100	100
ECG + EMG	100	98,2	95,4
Wearable S.	100	100	100

As can be seen from the literature, it is possible to achieve higher success with fewer features. This is because higher achievements can be achieved by not using complex, noisy or irrelevant features. In addition, using fewer features reduces the operational load and memory requirement, which allows more successful and real-time usable hardware to be developed with fewer data and processor capacity.

In addition, seizures were detected a priori using wearable sensors. The signal was processed with the method used previously and the obtained features were classified and the accuracy rates obtained with the wearable sensor attached to the upper arm are summarized in Table 3.

Table 3. Accuracies (%) in the detection of a seizure before it occurs.

Task	Accuracy (%)
Detection of an epileptic seizure 0-10 seconds before it occurs	100
Detection of an epileptic seizure 10-20 seconds before it occurs	100
Detection of an epileptic seizure 20-30 seconds before it occurs	100
Detection of an epileptic seizure 60-70 seconds before it occurs	98,5
Detection of an epileptic seizure 90-100 seconds before it occurs	97,6
Detection of an epileptic seizure 120-130 seconds before it occurs	95,6
Detection of an epileptic seizure 150-160 seconds before it occurs	94,9
Detection of an epileptic seizure 180-190 seconds before it occurs	92,2

As it can be seen from Table 3, the proposed approach can be successfully employed in the detection of a seizure before it occurs. It was seen from the results that although the human does not know that a seizure will have occurred, his body knows and tries to protect itself. Furthermore, a device was designed and its details can be found in www.insense.com.tr.

The designed device was employed in order to send the sensor data to a mobile phone that was paired. In mobile phone, the received signals were processed and relevant features were extracted. The extracted features were classified

by a machine learning method. Each of the trained three learning methods have been employed to generate a decision about epileptic status, type of the seizure and the pre-epileptic status. Generated alarms were also saved to a database and may be shared with medical doctors according to the permission of the user.

IV. CONCLUSION

This paper reports the results of detecting, classifying, and pre-detecting an epileptic seizure. The obtained results showed that the employed methodology is good enough to be successfully detecting, classify, and pre-detecting an epileptic seizure. Furthermore, the proposed method was employed in a real-time device

REFERENCES

- [1] Akbulut, F. P., Akan, A. 2015. "Akıllı Giyilebilir Hasta Takip Sistemleri", *Vogue*, 15(18), 440-443.
- [2] Araz, N. Ç., Yılmaz, K., Ölmez, A., Kılınç, M. 2009. "Epilepsi Ayırıcı Tanısında üç Olgu ile Kardiyovasküler Nedenler", *ADÜ Tıp Fakültesi Dergisi*, 10(2), 37 – 40.
- [3] Arkanoglu, A. 2011. "Epileptik nöbet ve psödonöbetlerin ayırıcı tanısına güncel yaklaşım", *Journal of Clinical and Experimental Investigations*, 2(3), 330-334.
- [4] Behbahani, S., Dabanloo, N. J., Nasrabadi, A. M., Attarodi, G., Teixeira, C. A., Dourado, A. 2012. "Epileptic seizure behaviour from the perspective of heart rate variability", *IEEE Computing in Cardiology (CinC)*, 117-120.
- [5] Bodde, N. M. G., Brooks, J. L., Baker, G. A., Boon, P. A., Hendriksen, J. G., Aldenkamp, A. P. 2009. "Psychogenic non-epileptic seizures—diagnostic issues: a critical review", *Clinical neurology and neurosurgery*, 111(1), 1-9.
- [6] Buchanna, G., Premchand, P., & Govardhan, A. (2022). Classification of Epileptic and Non-Epileptic Electroencephalogram (EEG) Signals Using Fractal Analysis and Support Vector Regression. *Emerging Science Journal*, 6(1), 138-150.
- [7] Cragar, D. E., Berry, D. T. R., Fakhoury, T. A., Cibula, J. E., Schmitt, F. A. 2002. "A Review of Diagnostic Techniques in the Differential Diagnosis of Epileptic and Nonepileptic Seizures", *Neuropsychology Review*, 12(1), 31-63.
- [8] Cobb, S. J., & Beebe, L. H. (2022). Quality of Life in Psychogenic Nonepileptic Seizures: An Evolutionary Concept Analysis Nonepileptic Seizures: An Integrative Review. *Issues in Mental Health Nursing*, 1-7.
- [9] Cuthill, F. M., Espie, C. A. 2005. "Sensitivity and specificity of procedures for the differential diagnosis of epileptic and non-epileptic seizures: a systematic review", *Seizure*, 14(5), 293-303.
- [10] Çakıl, D., İnanır, S., Baykan, H., Aygün, H., Kozan, R. 2013. "Epilepsi ayırıcı tanısında psikojenik non-epileptik nöbetler", *Göztepe Tıp Dergisi*, 28(1), 41-47.
- [11] D'Alessio, L., Giagante, B., Oddo, S., Wa, W. S., Solis, P., Consalvo, D., Kochen, S. 2006. "Psychiatric disorders in patients with psychogenic non-epileptic seizures, with and without comorbid epilepsy", *Seizure*, 15(5), 333-339.
- [12] Ertuğrul, Ö.F. (2016-a). Determining the Order of Risk Factors in Diagnosing Heart Disease by Extreme Learning Machine, *International Conference on Natural Science and Engineering (ICNASE'16)*, 10-19.
- [13] Ertuğrul, Ö. F. (2016-b). Forecasting electricity load by a novel recurrent extreme learning machines approach. *International Journal of Electrical Power & Energy Systems*. 78: 429–435, DOI:10.1016/j.ijepes.2015.12.006.
- [14] Ertuğrul, Ö. F. (2016-c). Aşırı Öğrenme Makineleri ile Biyolojik Sinyallerin Gizli Kaynaklarına Ayırıştırılması. *Dicle Üniversitesi Mühendislik Dergisi*, 7:41-40.
- [15] Ertuğrul, Ö. F., Altun, Ş. (Baskıda). Developing Correlations by Extreme Learning Machine for Calculating Higher Heating Values of Waste Frying Oils from their Physical Properties. *Neural Computing and Applications*. DOI: 10.1007/s00521-016-2233-8.
- [16] Ertuğrul, Ö. F., Kaya, Y. (Baskıda). Determining the Optimal Number of Body-Worn Sensors for Human Activity Recognition. *Soft Computing*, DOI: 10.1007/s00500-016-2100-7.
- [17] Ertuğrul, Ö. F., Kaya, Y., Tekin, R. (2016-a). A novel approach for SEMG signal classification with adaptive local binary patterns. *Medical & Biological Engineering & Computing*, 54(7): 1137-1146, DOI: 10.1007/s11517-015-1443-z.
- [18] Ertuğrul, Ö.F., Tağluk, M.E., Kaya, Y. (2012). Enerji İletim Hatlarında Wigner Ville Dağılımı, Gri Düzey Eş Oluşum Matrisi Ve Örüntü Tanıma Yöntemleri İle Arıza Analizi. 20. Sinyal İşleme ve İletişim Uygulamaları Kurultayı (SIU2012), 1-4.
- [19] Ertuğrul, Ö.F., Tağluk, M.E., Kaya, Y., Tekin, R. (2013). EMG Sinyallerinin Aşırı Öğrenme Makinesi ile Sınıflandırılması. 21. Sinyal İşleme ve İletişim Uygulamaları Kurultayı (SIU2013), 1-4.
- [20] Gonzalez-Velldn, B., Sanei, S., Chambers, J. A. 2003. "Support vector machines for seizure detection", 3rd IEEE International Symposium on Signal Processing and Information Technology (ISSPIT 2003), 126-129.
- [21] Gubbi, J., Kusmakar, S., Rao, A. S., Yan, B., O'Brien, T. J., Palaniswami M. 2015. "Automatic Detection and Classification of Convulsive Psychogenic Non-epileptic Seizures Using a Wearable Device", *IEEE Journal of Biomedical and Health Informatics*, doi:10.1109/JBHI.2015.2446539.
- [22] Hinich, M.J. and Clay, C.S., —The application of the discrete Fourier transform in the estimation of power spectra, coherence and bispectra of geophysical data, *Reviews of Geophysics*, 6(3):347-363, 1968.
- [23] Huang G. B., Zhu Q. Y. Siew C. K. (2006). Extreme learning machine: Theory and applications. *Neurocomputing*, 70:489–501.
- [24] Hussain, W., Sadiq, M. T., Siuly, S., & Rehman, A. U. (2021). Epileptic seizure detection using 1 D-convolutional long short-term memory neural networks. *Applied Acoustics*, 177, 107941.
- [25] Kanas, V. G., Zacharaki E. I., Pippa, E., Tsirka, V., Koutroumanidis, M., Megalooikonomou, V. 2015. "Classification of epileptic and non-epileptic events using tensor decomposition", *IEEE 15th International Conference on Bioinformatics and Bioengineering (BIBE)*, 1-5.
- [26] Kaya, Y., Ertuğrul, Ö. F. (2016). A novel approach for spam email detection based on shifted binary patterns. *Security and Communication Networks*, 9(10): 1216–1225.
- [27] Kaya, Y., Ertuğrul, Ö. F., Tekin, R. (2012). Epileptik EEG İşaretlerinin Sınıflandırılmasında Karar Kuralları ve Karar Ağaçlarının Kullanılması. *Yaşam Bilimleri Dergisi*, 1(2):403-413.
- [28] Kotsopoulos, I. A. W., de Krom, M. C. T. F. M., Kessels, F. G. H., Lodder, J., Troost, J., Twellaar, M., van Merode, T., Knottnerus, A. J. 2003. "The diagnosis of epileptic and non-epileptic seizures", *Epilepsy research*, 57(1), 59-67.
- [29] Kusmakar, S., Gubbi, J., Rao, A. S., Yan, B., O'Brien, T. J., Palaniswami, M., 2015-b. "Classification of convulsive psychogenic non-epileptic seizures using histogram of oriented motion of accelerometry signals", 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 586-589.
- [30] Kusmakar, S., Gubbi, J., Yan, B., O'Brien, T. J., Palaniswami, M. 2015-a. "Classification of convulsive psychogenic non-epileptic seizures using muscle transforms obtained from accelerometry signal", 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 582-585.
- [31] Lai, W., Du, D., & Chen, L. (2022). Metabolomics Provides Novel Insights into Epilepsy Diagnosis and Treatment: A Review. *Neurochemical Research*, 1-16.
- [32] Nei, M. 2009. "Cardiac Effects of Seizures", *Epilepsy Currents*, 9:(4), 91–95.
- [33] Nikias, C.L. and Petropulu, A.P., *Higher order spectral analysis: A nonlinear signal processing framework*, Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [34] Ning, T. and Bronzino, J.D., Bispectral analysis of the rat EEG during various vigilance states, *IEEE Trans Biomed Eng*, 36(4):497–499, 1989.
- [35] Pippa, E., Zacharaki, E. I., Mporas, I., Megalooikonomou, V., Tsirka, V., Richardson, M., Koutroumanidis, M., 2014. "Classification of epileptic and non-epileptic EEG events", *IEEE 4th International Conference on Wireless Mobile Communication and Healthcare (Mobihealth)*, 87-90.
- [36] Proix, T., Truccolo, W., Leguia, M. G., Tchong, T. K., King-Stephens, D., Rao, V. R., & Baud, M. O. (2021). Forecasting seizure risk in adults with focal epilepsy: a development and validation study. *The Lancet Neurology*, 20(2), 127-135.

- [37] Raghuveer, M.R. and Nikias, C.L., Bispectrum estimation: A parametric approach, IEEE Trans. Acoust., Speech, Signal Processing, 33(4):1113-1230, 1985.
- [38] Reuber, M. 2008. "Psychogenic nonepileptic seizures: answers and questions", *Epilepsy & Behavior*, 12(4), 622-635.
- [39] Reuber, M., Elger, C. E. 2003. "Psychogenic nonepileptic seizures: review and update", *Epilepsy & Behavior*, 4(3), 205-216.
- [40] Sezgin, N., Estimation and classification of sleep apnea in adults by developed preprocessing-neural network models, PhD thesis, Inonu University, Malatya, Turkey, 2010.
- [41] Sezgin, N., Tađluk, M. E., Ertuđrul, Ö. F., Kaya, Y. (2013). Epileptik EEG Őaretlerinin İviz-spektrum Analizi. 21. Sinyal Őleme ve İletiŐim Uygulamaları Kurultayı (SIU2013), 1-4.
- [42] Sigl, J.C. and Chamoun, N.G., An introduction of bispectral analysis for the electroencephalogram, *Journal of Clinical Monitoring*, 10(6):392-404, 1994.
- [43] Tagluk M.E., Sezgin N., A new approach for estimation of obstructive sleep apnea syndromel Expert Systems with Applications, 38(5):5346-5351, 2011.
- [44] Tatlı, B., Aydınlı, N., ÇalıŐkan, M., Özmen, M. 2004. "Non epileptik paroksizmal olaylar: olgu sunumları ile derleme", *Türk Pediatri ArŐivi*, 39(2), 58-64.
- [45] Turner, K., Piazzini, A., Chiesa, V., Barbieri, V., Vignoli, A., Gardella, E., Tisi, G., Scarone, S., Canevini, M. P., Gambini, O. 2011. "Patients with epilepsy and patients with psychogenic non-epileptic seizures: video-EEG, clinical and neuropsychological evaluation", *Seizure*, 20(9), 706-710.
- [46] Worrell, G. A., Brinkmann, B. H., Lundstrom, B. N., & Stead, M. (2021). Emerging Treatments for Epilepsy. *Epilepsy*, 441-449.
- [47] Xu, P., Xiong, X., Xue, Q., Li, P., Zhang, R., Wang, Z., Valdes-Sosa, P. A., Wang, Y., Yao, D. 2014. "Differentiating between psychogenic nonepileptic seizures and epilepsy based on common spatial pattern of weighted EEG resting networks", *IEEE Transactions on Biomedical Engineering*, 61(6), 1747-1755.
- [48] Yadav, R., Agarwal, R., Swamy, M. N. S., 2009. "A new improved model-based seizure detection using statistically optimal null filter", *Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2009)*, 1318-1322.
- [49] Zijlmans, M., Flanagan, D., & Gotman, J. (2002). Heart rate changes and ECG abnormalities during epileptic seizures: prevalence and definition of an objective clinical sign. *Epilepsia*, 43(8), 847-854.

BIOGRAPHIES



ÖMER FARUK ERTUĐRUL, was born in Batman, Turkey in 1978. He received the B.S. degree from the Hacettepe University, Department of Electrical and Electronics Engineering in 2001, M.S. and Ph.D. degrees in Electrical and Electronics Engineering in 2010, and 2015, respectively. His research interests include machine learning and signal processing.



YASIN SÖNMEZ, was born in Diyarbakır, Turkey in 1986. He received the B.S. degree from the Firat University, Technical Education Faculty, Department of Electronics and Computer Education in 2010, M.S. degree in computer science from the Firat University in 2012 and Ph.D. degree department of software engineering at Firat University in 2018. His research interests include, artificial intelligence, and information security.



NECMETTİN SEZGİN, was born in Batman, Turkey in 1978. He received the B.S. degree from the Hacettepe University, Department of Electrical and Electronics Engineering in 2001, M.S. and Ph.D. degrees in Electrical and Electronics Engineering in 2010, and 2015, respectively. His research interests include machine learning and signal processing.



EŐREF AKIL, was born in Batman, Turkey in 1977. He received the B.S. degree from the Uludađ University, Department of Medicine in 2003, His research interests include paralysis and epilepsy.

A Prospective Look on Optimization Methods For RFID Systems: Requirements, Challenges, and Implementation Aspects

Lida Kouhalvandi

Abstract—The radio frequency identification (RFID) is a configuration of wireless communication that uses radio frequency (RF) waves for following up and recognizing data. The RFID system includes important parts as antenna and integrated circuit (IC) for radiating and storing data, respectively. Hence, high performance antenna and IC circuits must be designed for assembling the energy from the radio waves and feeding the RFID chip. One of the important circuit/block in the IC part is the amplifier where the antenna is sensing the radiated output power. For this case, it is substantially important to design high performance antenna and amplifier where the specifications of these circuits must be optimized in a professional way. In this paper, we collect the recently published optimization methods that are employed for designing antenna and RF/analog-based amplifiers. Any researcher by referring to these algorithms can access and find the solutions for their problems, straightforwardly.

Index Terms—Amplifier, antenna, radio frequency identification (RFID), radio frequency (RF), optimization methods, power amplifier (PA).

I. INTRODUCTION

THE need of radio frequency identification (RFID) systems is growing day-by-day, and can be used in the internet of things (IoT) networks, healthcare, agriculture, industrial environments, and smart houses [1], [2], [3], [4], [5], [6], [7], [8], [9], [10]. These systems are preferred due to the contactless communication, dense read/write operation, and low-cost profits in modern technologies [4], [11]. Typically, the RFID system includes the tags, readers, and a database server where the reader can prompt the tags and then send/receive data to/from the tags [12], [13] as Fig. 1 shows. The RFID reader is the device that receives data from the identifiers where the electromagnetic provided from the reader antenna can be used as an energy source [14]. As Fig. 2 shows, in the active RFID systems the two significant components are antenna and amplifiers that is connected to the smart card chip. For proper working of RFID system, designing and optimizing high performance antennas [15], [16], [17] and amplifiers [18], [19], [20], [21], [22], [23], [24], [25] are required where advanced optimization methods must be employed for configuring and sizing these designs [26], [27], [28], [29], [30], [31], [32].

Optimization methods are the techniques focusing on minimizing or maximizing aimed objective functions. At the system and circuit level designs, required constrains and

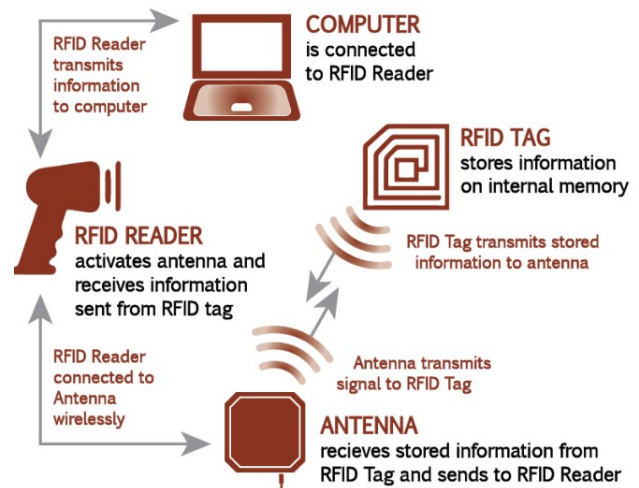


Fig. 1. The RFID system architecture [34].

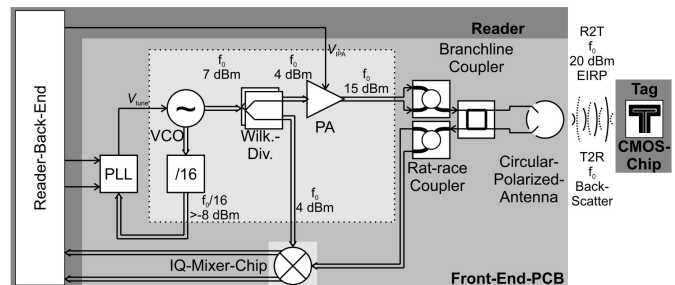


Fig. 2. Components of active RFID system [35]; voltage-controlled oscillator (VCO) and phase-locked loop (PLL).

goals must be determined and afterwards suitable optimization methods can be considered for designing various designs and circuits. Over the last decade, various optimization methods have been employed for designing and optimizing antennas and amplifiers that can be also used in the RFID applications [33]. The reported optimization methods have significant advantages as speeding-up the process, reducing cost and errors with the least efforts.

This survey devotes to provide a comprehensive summary on various optimization methods that can be employed in the design of RFID systems. Typically, any RFID systems include two important designs as antenna and amplifier circuits. Hence, this paper is to summarize various optimization methods and algorithms used in improving the overall performance of RFID systems. Such kind of survey will enable engineers

to figure out successfully the various optimization methods at once and will decide easily on the accurate methods for solving their problems.

The remainder of the article is organized as follows. Section II presents the motivation of employing optimization methods in the RFID systems. Section III devotes to summarize the various optimization methods employed in antenna and amplifier designs. Finally, Sec. IV concludes this manuscript.

II. OPTIMIZATION METHODS AND RFID DESIGNS

Optimization methods have proved their practical uses in the coupled antennas that are used for the RFID systems. As Fig. 3 presents, a large ferrite core is inserted into the tag coil antenna for improving the magnetic coupling that is between the reader and tag antennas.

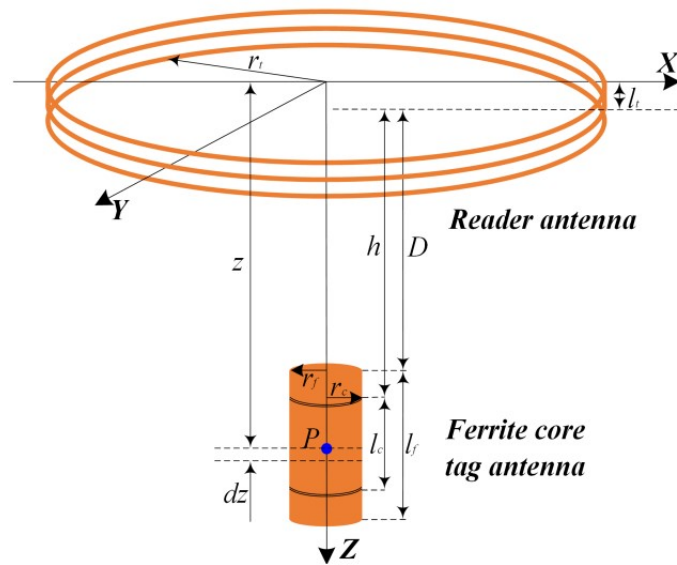


Fig. 3. Schematic of two RFID solenoid antennas presented in [36].

In [37], self-tuning performance is done for solving the optimization problems in the design of ultra-high frequency (UHF) RFID antennas. The presented method leads to maximize the performance of the RFID and to make it independent at the surrounding environment. Figure 4 presents in detail the fingertip tag. For UHF RFID reader, the Quasi-Yagi antenna is presented in [38]. It is a wearable device and can be used in wireless sensing (see Fig. 5) by determining a centered frequency band at 915 MHz.

The RFID sensing antenna is presented in [39] where the detail of antenna structure is shown in Fig. 6, and it has dimensions of $(65 \times 65 \times 4 \text{ mm}^3)$. The type of antenna differs with the injected liquid that is inside the presented antenna.

One of the amplifier architecture that can be used in the RFID systems is depicted in Fig. 7 that includes the active transistor, switch, and an antenna in the overall configuration. This circuit can be used in the field of communications by means of reflected power. Another use of amplifier is proved in [40] where a low-power amplifier is presented using tunnel diodes results in stable reflection gain (see Fig. 8).

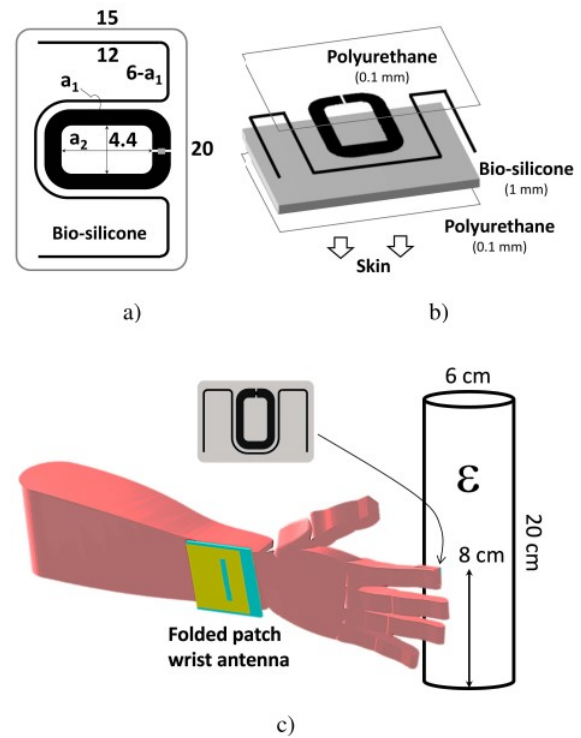


Fig. 4. Presented fingertip tag in [37] to be used as a UHF-RFID Antenna with a) Tag layout in mm size, b) practical use of antenna in the bio-medical applications, c) Tag layout.

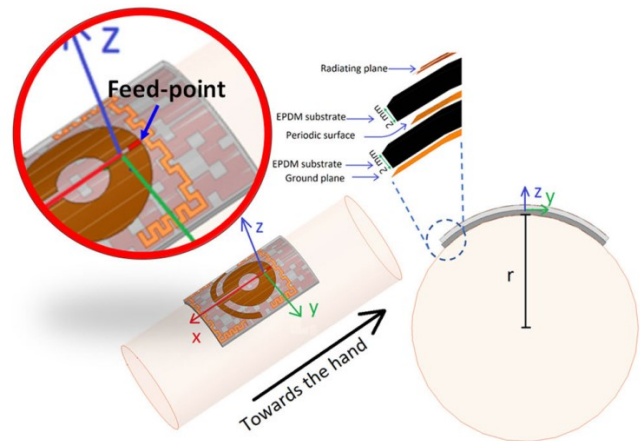


Fig. 5. Wearable quasi-Yagi antenna as UHF RFID reader [38].

III. OPTIMIZATION METHODS

This section provides an overview on various optimization methods used for designing and optimizing antennas and amplifiers. The recently published studies include advanced optimization methods are summarized in a comprehensive aspect.

The chicken swarm optimization (CSO) method is a bio-inspired algorithm aims to follow the hierarchical order and the behaviors of the chicken swarm [41]. This optimization method is employed in [42] for deployment of phased array antennas to be used in the RFID networks. In this method, a new indicator is used to reflect the disturbance of frequency selective fading. The general structure of the CSO method is summarized in Fig. 9.

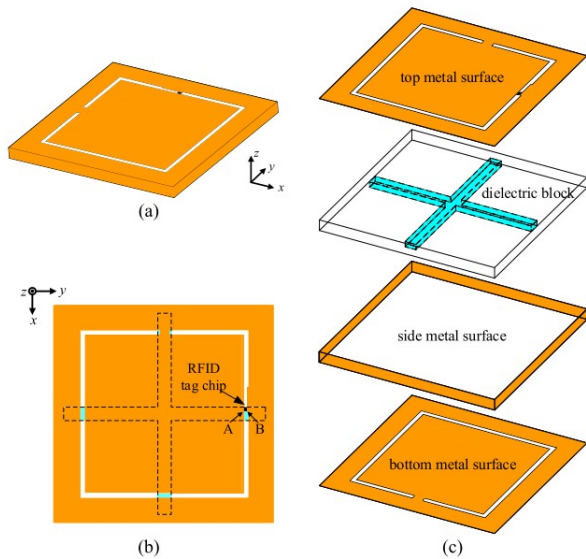


Fig. 6. Presented RFID tag antenna with a) overall structure, b) top, and c) exploded view [39].

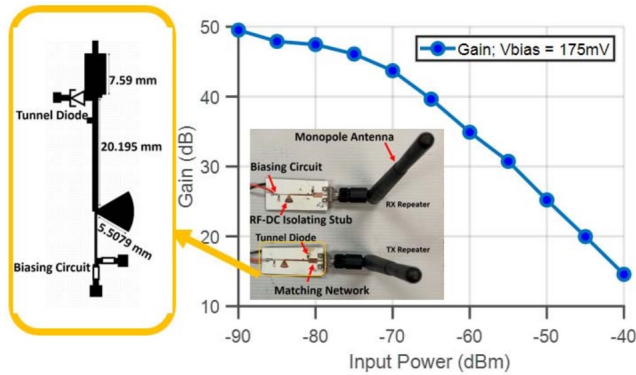


Fig. 7. Presented RFID amplifier presented in [18].

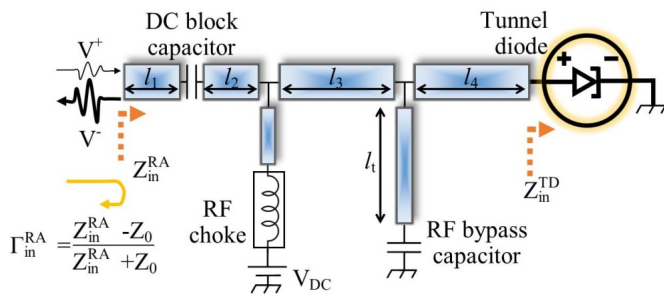


Fig. 8. Tunnel diode-based reflection amplifier presented in [40] to be used in the RFID systems.

The Particle swarm optimization (PSO) method is performed by iteratively trying to enhance a candidate solution and is a stochastic optimization technique based on swarm [43]. This method is employed for the real-time 3D localization of UHF-RFID tags in [28]. In another study presented in [44], the PSO method is used for improving the RFID anticollision model and for optimizing the high-dimensional problem that Fig. 10 illustrates the feature of this optimization.

The hierarchical multilevel bottom-up method becomes important for radio frequency (RF) designs in the recent years where the accuracy simulation of overall system especially for

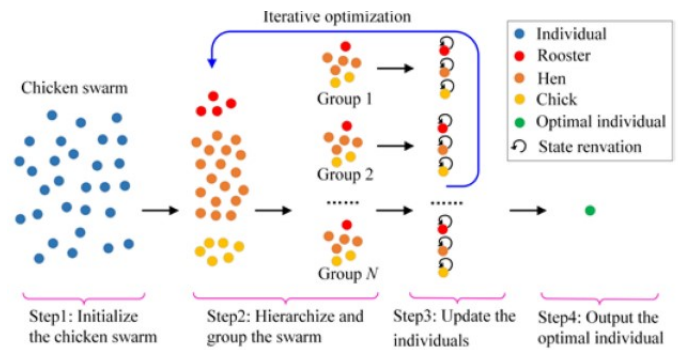


Fig. 9. The CSO method employed in [42] for designing phased array antenna.

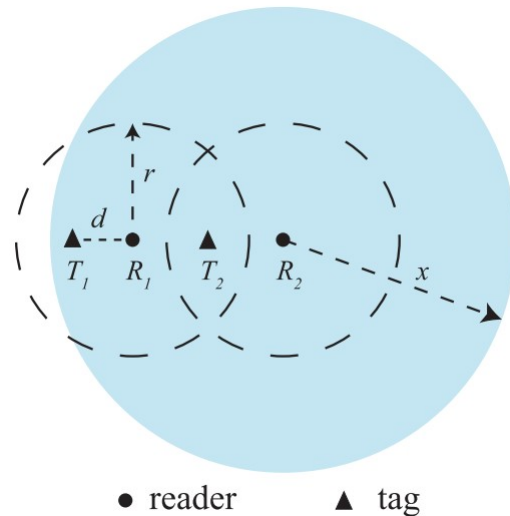


Fig. 10. Employed PSO method in [44].

the electromagnetic simulations have been increased, substantially [45].

The bottom-up optimization (BUO) is the method where it starts with smallest parts and develops to the largest sections. In [46], by using inductor-capacitor networks shown in section A and B of Fig. 11.a, matching networks (MNs) for the power amplifier (PA) with Gallium Nitride (GaN) high electron mobility transistor (HEMT) is designed and optimized. In this method, input and output MNs are started with one MN and sequentially the number of MNs are increasing up to achieving desired output specifications. The detail implementation of this method is shown in Fig. 11.b.

This method is also employed for designing and optimizing the antenna, presented in [47], where the transmission lines (TLs) are increased sequentially and are configuring the structure of antenna. The detail implementation of BUO method used for optimizing antenna is presented in Fig. 12. Additionally, this method is used for designing array antenna [48] where the number of single antennas is enhancing in a latter style as Fig. 13 presents. Hence, the optimal number of single antennas with suitable feeding point can be determined.

The BUO method can be employed in the system-level design as well [49], [45] where Fig. 14 illustrates one of the useful aspect of this method. This method splits the system into sub-blocks as low-noise amplifier (LNA), voltage controlled

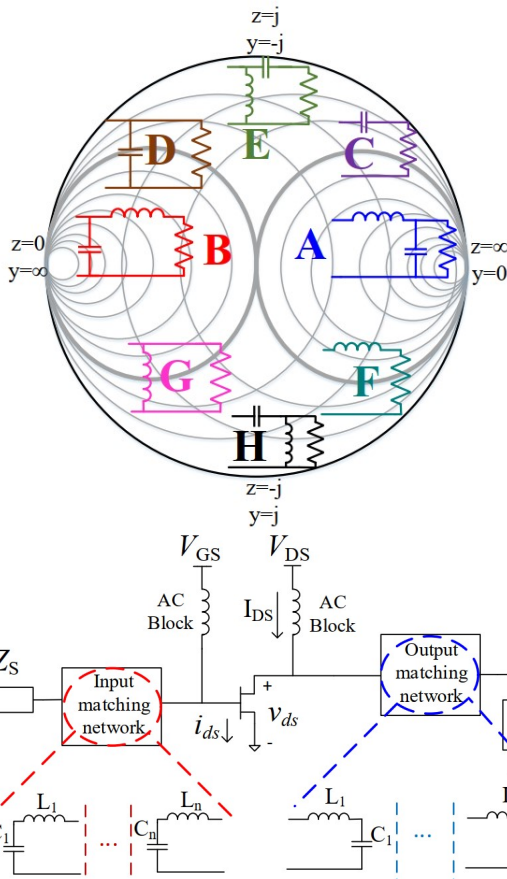


Fig. 11. a) Smith chart with various passive MNs [top]; b) BUO method for designing and optimizing an amplifier [bottom] [46].

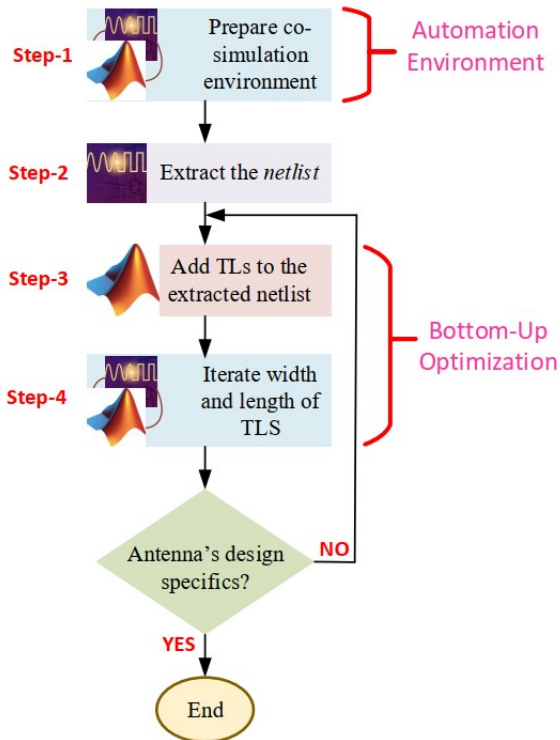


Fig. 12. BUO method for optimizing antenna where the TLs are employed [47].

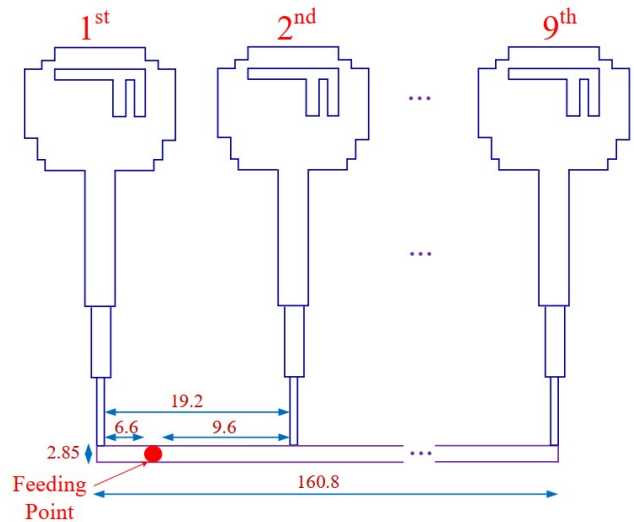


Fig. 13. BUO method for optimizing array antenna [48].

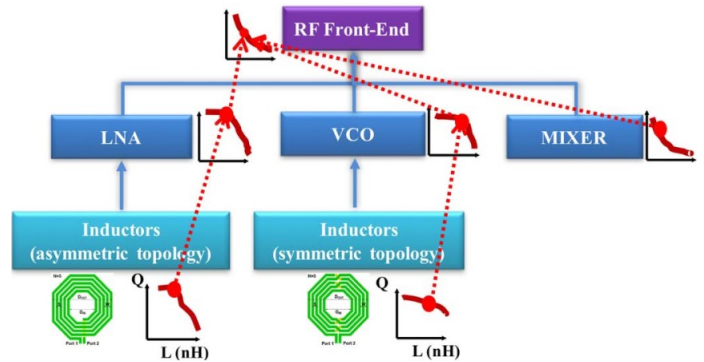


Fig. 14. Designed RF front-end with BUO method [45].

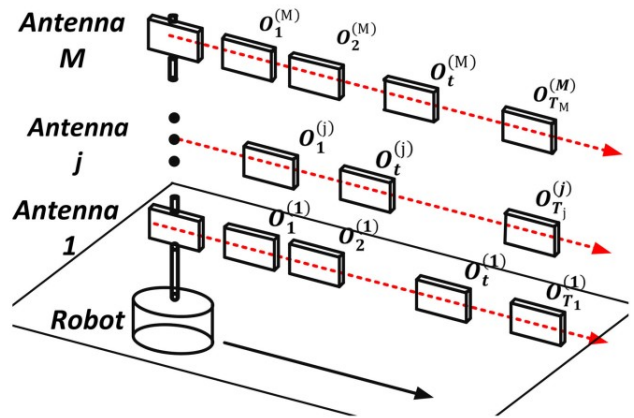


Fig. 15. An illustration of multi-antenna synthetic [2].

oscillators (VCO), and mixer leads to generate and size the layout, effectively.

In [2] for 3D localization of UHF RFID tags, a nonlinear optimization leads to assemble measurements from various antennas is presented (see Fig. 15). This method aims to optimize the cost function where phase measurements are gathered by minimum two antennas, and finally a multi-antenna synthetic aperture is generated.

During designing any antenna or amplifier circuit, time

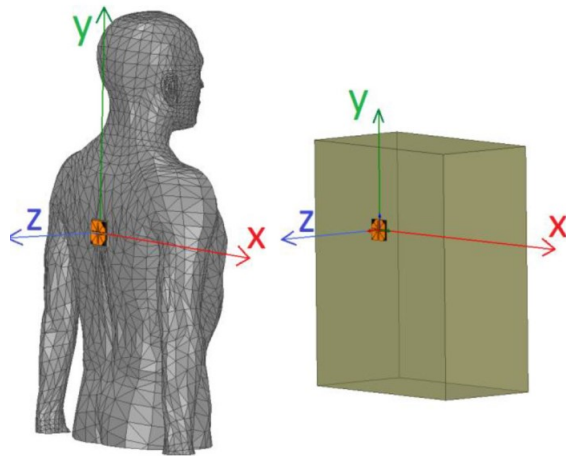


Fig. 16. The comparison between anatomical model (presented in left side) and the cuboid model (presented in right side) [3].

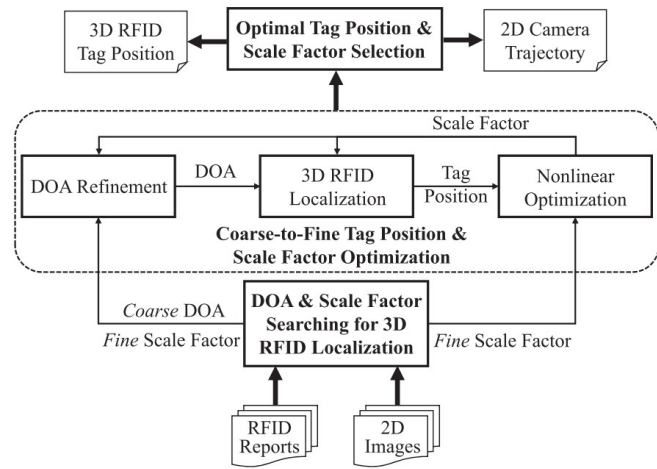


Fig. 17. Presented configuration of RF-MVO in [15].

computation is also another important specification along with the output performances. In [3], the antenna is designed based on the cuboid model since with this model the simulation time with mesh size are reduced, substantially compared with anatomical model (see Fig. 16).

The coarse-to-fine optimization method is employed in [15] for improving the localization accuracy in RFID systems. The presented optimization process is based on the space mapping [50] and uses two models of the same physical system namely as: expensive model (fine model) and cheaper (coarse) model. Based on this method, RF-monocular visual odometry (MVO) is presented that is connecting a light-weight 2D monocular camera to two reader antennas in a parallel version results in reduced localization error. Figure 17 shows the presented RF-MVO architecture with the direction of arrival (DOA) that is azimuth and elevation angles.

QI et al., presents optimization techniques based on the queuing theory and control feedback for minimizing the energy cost where it is illustrated in Fig. 18 that can be used in the industrial internet of things (IIoT) [5].

In [30] efficient global optimization (EGO) method is applied for sizing the designed antenna as Fig. 19 presents. This antenna provides cost-effective option for sensing by using

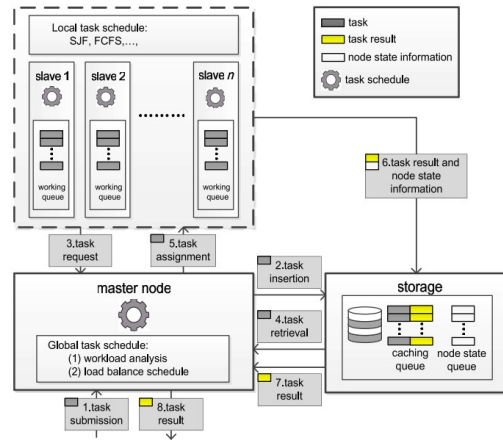


Fig. 18. Configuration of computing service provider (CSP) on Redis in [5].

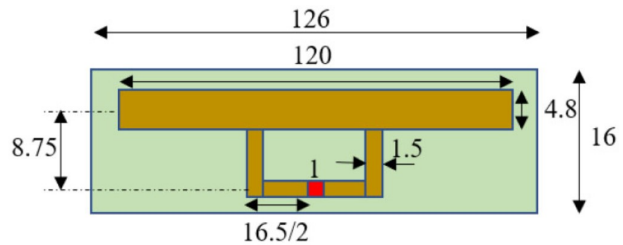


Fig. 19. optimized antenna with EGO method presented in [30].

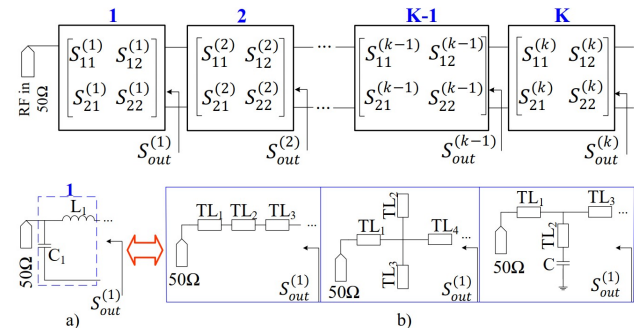


Fig. 20. a) Illustration of TDO method for dividing input/output MNs with S-parameter blocks [top]; b) Replacement choices of each block presented above with various TL-based blocks [bottom] [52].

multidimensional differential measurement.

The top-down optimization (TDO) method is opposite to the BUO method and it starts with largest sections and decreases to the smallest sections [51]. This method is employed in [52] for designing and optimizing amplifier where the input/output MNs are divided into subsections presented with scattering parameters (S-parameter) (see Fig. 20.a) and each of the subsections are replaced and evaluated with various TLs presented in Fig. 20.b.

This method is employed for optimizing the implanted multiple-input and multiple-output (MIMO) antennas leads to generate the configuration of the implanted antenna used for bio-medical applications [53]. As Fig. 21 presents the structure of the implanted MIMO antenna where firstly the various bio-medical tissues as bone, muscle, fat, and skin surroundings are stucked over together. Afterwards, the MIMO antenna mounted on the ground and substrate planes are assembling to

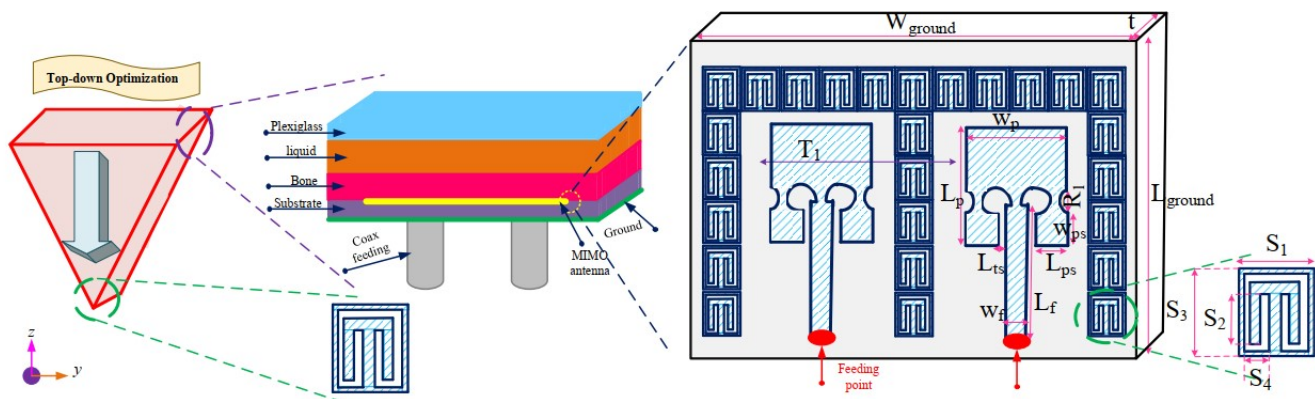


Fig. 21. Presented TDO method for configuring and optimizing implanted MIMO antenna [53].

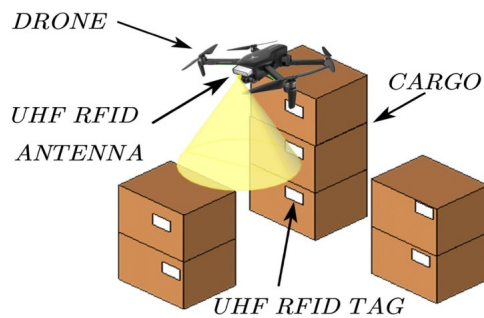


Fig. 22. Stocktaking task-planning with RFID readers [32].

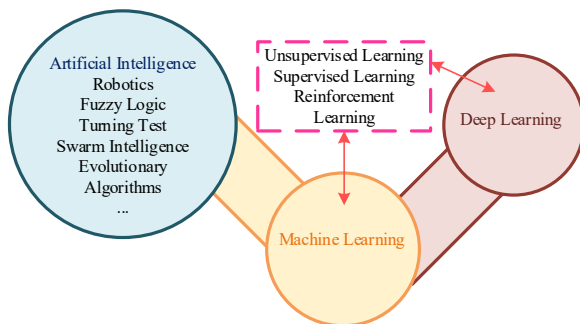


Fig. 23. Subsets of AI and its applications in various domains and fields.

the bio-medical tissues. With this procedure, the difficulty of configuring implanted devices can be solved and the general structure of these bio-medical devices can be easily generated.

Liu et al., present the use of lion swarm optimization method for presenting a trajectory planning model for equipment with RFID readers [32]. This method enhances the efficiency and safety of overall system by minimizing the length of the flight path. Figure 22 shows the general view of drone stocktaking of goods where it is optimized to the minimal flight path.

For the practical use of RFID systems, artificial intelligence (AI) has gotten attention of researchers in the recent years [54], [55] and the subsets of AI are described in Fig. 23. From another point of view in the developed fifth and sixth generation (5G and 6G) networks, high performance amplifiers and antennas play important roles to improve the communication performance [56], [57]. Typically, the figure of merit (FoM) of each amplifier can be determined by the output power (P_L) that is challenged by the power gain (G_p), drain efficiency

(η_D), phase distortion (AM/PM), and amplitude distortion (AM/AM). Concurrently, considering these specifications are not straightforward and requires strong optimization-based approaches. Conventional electronic design automation (EDA) tools such as ADS, AWR, etc. are good candidates for providing optimizations; however, when the design parameters and specifications are lot, these tools can not be powerful enough.

Recently, learning-based methods that are based on the machine learning (ML) have been successfully used in the RF designs. In this track, shallow neural network (SNN) and deep neural network (DNN) are candidates to model the RF circuits such as amplifiers and antennas by determining the relationship between the input and out data that are design parameters and design specifications, respectively.

The SNN is the neural with one hidden layer that it proves its benefits in designing and optimizing complex circuits. In [58], the SNN is employed for designing an amplifier with TLs. In the presented method, the amplifier design is initialized by using the simplified real frequency technique (SRFT) [59]. Firstly, the suitable SNN for the amplifier with lumped elements is designed. Then the suitable SNN for this amplifier is trained, and finally with the help of constructed SNN, the amplifier with TLs is optimized. Figure 24 presents the general flowchart for optimizing an amplifier with TLs through a pre-constructed SNN.

The SNN is also used in the domain of antenna designs and it is useful in sizing the patch antennas. In [60] based on the Bayesian optimization (BO), the SNN is trained and the configured antenna with the BUO method is optimized results in flat-gain performance. Figure 25 demonstrates the presented optimization method that is based on the BUO and BO methods for configuring and sizing antennas, respectively.

By using the neural network, the antenna shown in Fig. 26 is optimized and sized where the optimal design parameters are predicted [61]. This network is employed using the particle swarm optimization (PSO). Another use of neural network is presented in [62] for GPS Beidou dual-mode Yagi microstrip antenna (see 27).

W. Su et al., design and optimize antenna that can be used for cross-body communication using neural network. The comprehensive antenna model with simulation results are depicted in Fig. 28 [63]. In another study, presented in [64], the BO method is employed for converting the lumped

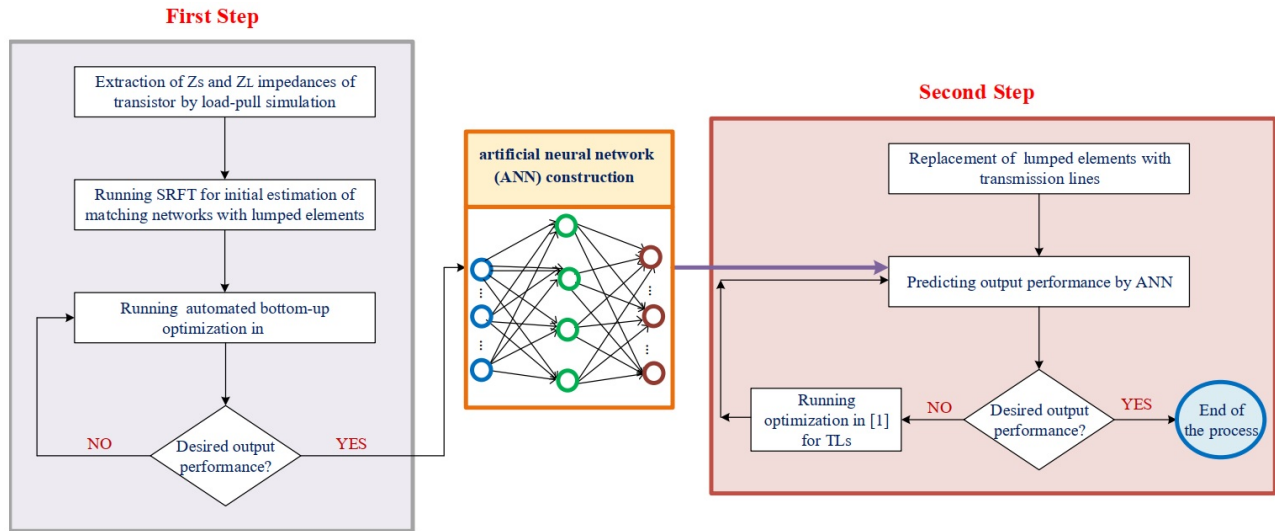


Fig. 24. Optimizing an amplifier with the pre-constructed SNN[58].

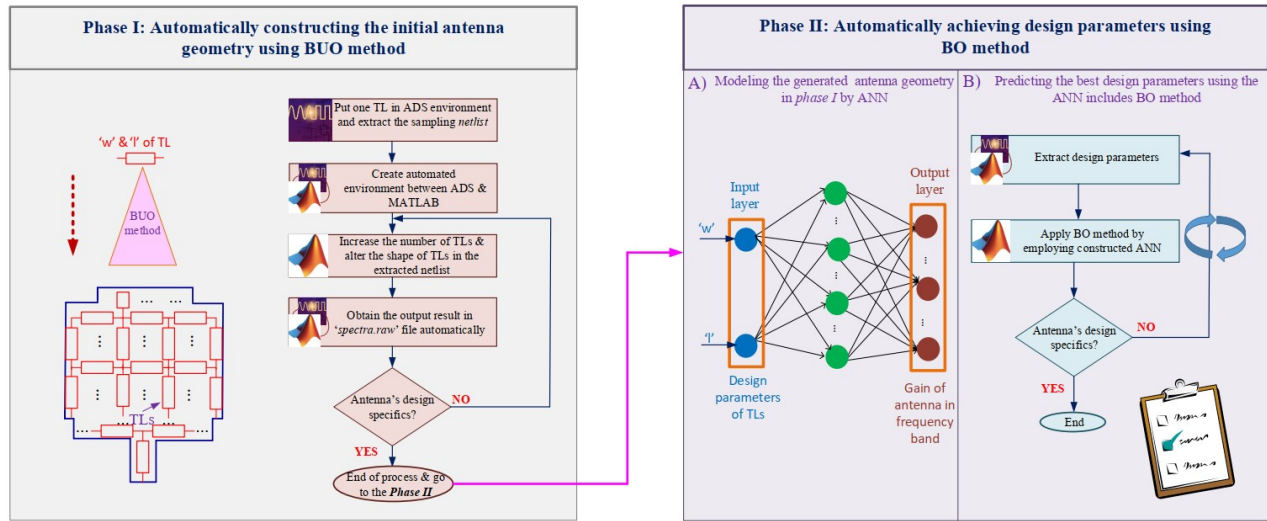


Fig. 25. Presented flowchart for configuring and sizing a patch antenna with the BUO method and SNN network [60].

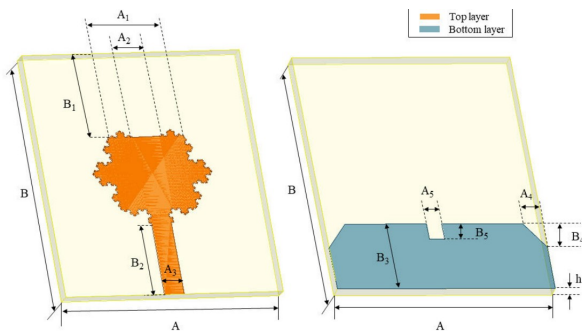


Fig. 26. Optimized antenna in [61].

element (LE) PA to the amplifier with distributed elements (DE). Figure 29 presents the BO-method for generating a PA with DEs where the initial design is with LEs.

The DNN is the multi-layer neural network leads to accurately modeling RF designs with lot design parameters [65], [66]. The practical use of this network is provided in the recently published studies.

The touch screen interaction system can be employed in

the innovation of RFID sensing. Hence, many applicant are requesting high quality touch screens. In [54], the DNN structure is used for predicting the the hand finger gesture coordinates (see Fig. 30).

The application of DNNs with previously reported optimization methods is summarized in Fig. 31. It shows that based on the AI methods, once the transistor model is selected, automatically the high power amplifier (HPA) configuration with the electromagnetic (EM)-verified post-layout is generated. The summarized methods in the following will help the designers to achieve the ready-to-fabricate layouts without human interruption and dependency to designer's experience. The main objective is to provide the novel optimization-oriented methods where the transistor model is selected, the optimal configuration and design parameters are predicted automatically that result in post-layout generations. All the process is performed automatically where this platform is created with the combination of electronic design automation (EDA) tool as ADS and numerical analyzer as MATLAB. The EDA tool, working in the background, generates the

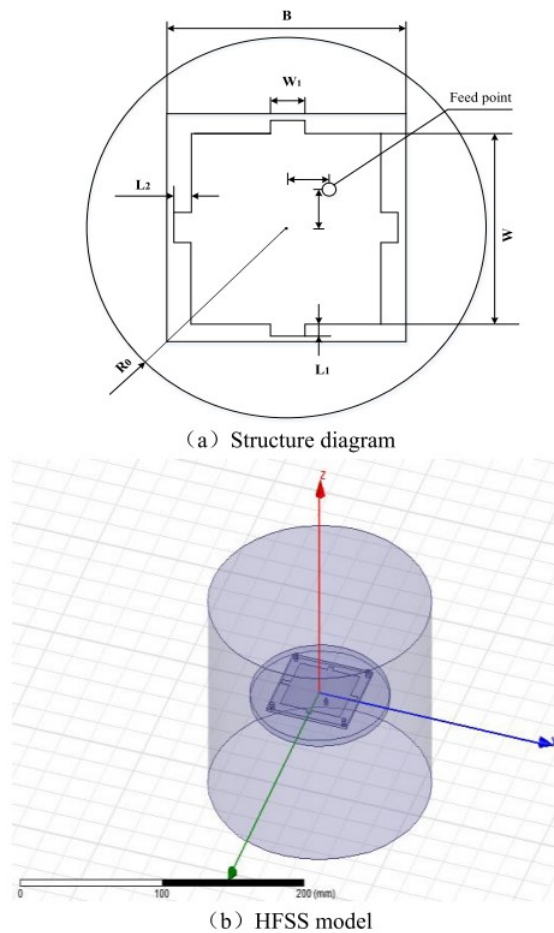


Fig. 27. Sizing of antenna using the neural network presented in [62].

nonlinear simulation results of the PA. Besides, the numerical analyzer gets these data and do mathematical analysis and optimization. These optimized data are used for constructing the DNN to predict the optimal design parameters. Typically, optimizing nonlinear design specifications as P_L , G_p , η_D , and phase distortion are not straightforward and requires intelligent algorithms and methods. Hence, this method proposes the use of artificial intelligent (i.e., DNNs) and also implementation of strong multi-objective algorithms that paves the way of RF designers and generates ready-to-fabricate layouts.

The benefits of using DNNs in the RF designs can be divided into four subsections as:

- Providing an automated optimization-oriented environment which decreases the manual interruptions;
- Employing multi-objective algorithms for concurrently optimizing P_L , G_p , η_D , and phase distortion;
- Constructing the sequential classification and regression DNNs for predicting the suitable PA configuration and also optimal component values;
- Generating the ready-to-fabricate layout of HPAs after providing the transistor model.

Following devotes to present in detail the use of DNNs:

In [67] for enhancing the measurement performance of RFID system, DNN is employed to optimize the 3-D structure of multitag as Fig. 32 presents. This performance will avoid the collision of tags communication importantly and will

reinforce the overall specification of the subsequent multitag batch reading.

In another study, the classification DNN and regression DNN are employed for determining the suitable topology of PA and sizing the design parameters, respectively [68]. The regression DNN, presented in [68], is based on the thompson sampling efficient multi-objective optimization (TSEMO) algorithm and the regression DNN, presented in [69], is based on the multi-objective pareto front using modified quicksort (PFUMQ) algorithm. The TSEMO-based DNN is for optimizing output power (P_L) that is challenged by the power gain (G_p), and drain efficiency (η_D). Additionally, the PFUMQ-based DNN is for optimizing the phase distortion (AM/PM), and amplitude distortion (AM/AM). Figure 33 presents the classification DNN employed in [68] and the two regression DNNs based on the TSEMO [68] (Fig. 34.a) and PFUMQ [69] (Fig. 34.b) algorithms. The input layer features of regression DNN presented in Fig. 34.b are AM-to-AM and AM-to-PM specifications where they depend on the variation of source power (P_S). The output features are the pareto optimal front (POF) of two functions (f_1 and f_2) by using the PFUMQ algorithm.

In summary, the presented summarized methods can be used in designing and optimizing the amplifier and antennas.

- 1) Constructing and predicting the optimal configuration by:
 - Bottom-up optimization [46], [47];
 - Top-down pruning optimization [52];
 - Classification DNN [68];
- 2) Predicting optimal design parameters by:
 - Bayesian optimization with SNN [64], [58], [60];
 - Regression DNN based on the Thompson Sampling Efficient Multi-objective Optimization (TSEMO) algorithm [68];
 - Regression DNN based on the multi-objective particle swarm optimization (PSO) and multi-objective pareto front using modified quicksort (PFUMQ) algorithms [69];

In [70], a new method called MWISBAIL is presented that is the distributed anti-collision algorithm and is based on the idea of a centralized collision avoidance algorithm. This method is combined with the machine learning leads to improved output responses. An example of RFID system with four readers as (R_1 , R_2 , R_3 , and R_4) is presented in Fig. 35 where R_2 and R_3 are active and R_1 with R_4 are inactive. The overall performance of RFID system is improved by employing presented neural network in Fig. 36. For this case, a single neural network model is trained and the constricted model to every RFID reader is applied. The details of proposed flowchart are presented in Fig. 37.

C. Peng et al., present the deep convolutional neural network (CNN)-based approach for locating multiple tags with acceptable accuracy and stability in the indoor environment [71]. Figure 38 presents the designed CNN model for location scene of RFID readers and reference tags (see Fig. 39). In another use of deep learning, in [72] convolutional neural network structure is presented for activity recognition from

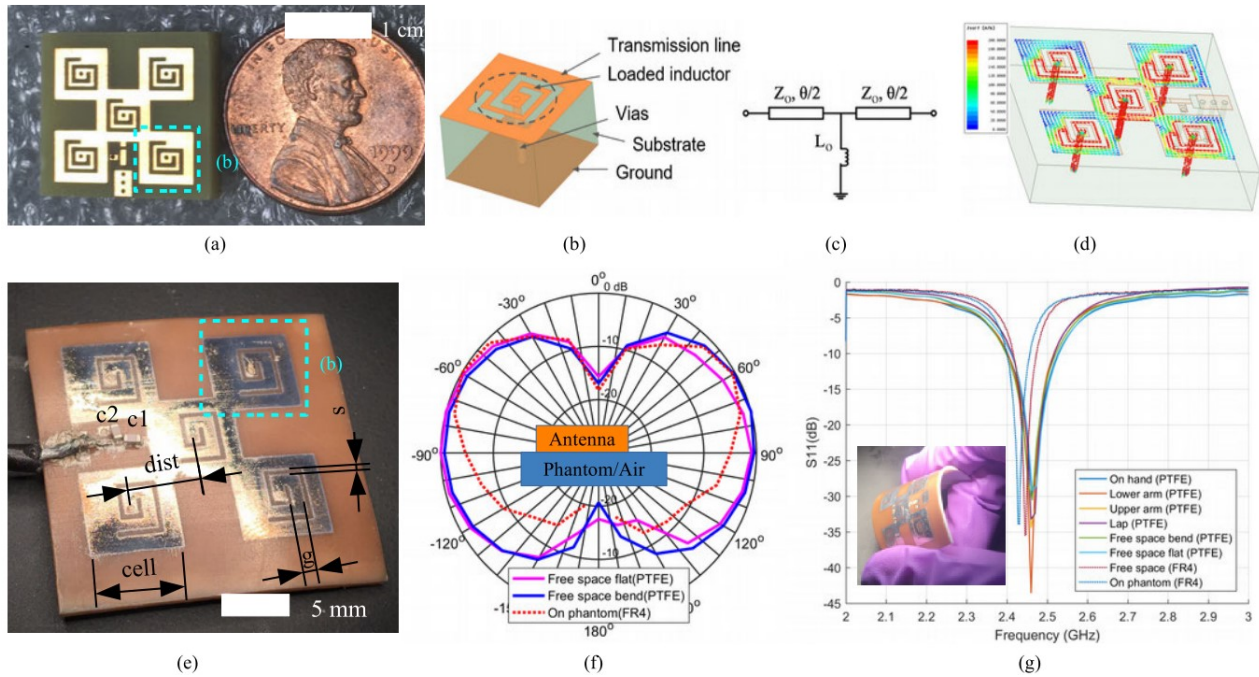


Fig. 28. Presented Mosaic antennas with their performances in [63]. Photo of (a) the presented antenna on FR4 substrate, (e) the flexible antenna on PTFE substrate. (b) the configuration of unit cell of antennas shown in (a) and (e). (c) equivalent circuit of the unit cell in (b). (d) simulated surface current distribution of the antenna. (f) measured radiation patterns of the flexible mosaic antenna. (g) the measured S11 of the flexible mosaic antenna. The inserted photo in (g) presents bending of the mosaic antenna.

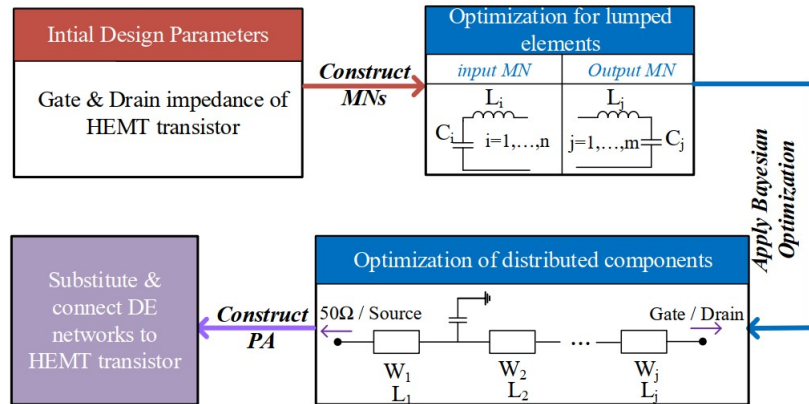


Fig. 29. Presented method for converting the LE power amplifier to the amplifier with the DEs through the BO method [64].



Fig. 30. a) DNN structure employed for the use of touch screens [right]; b) Gestures in touch screens [left] [54].

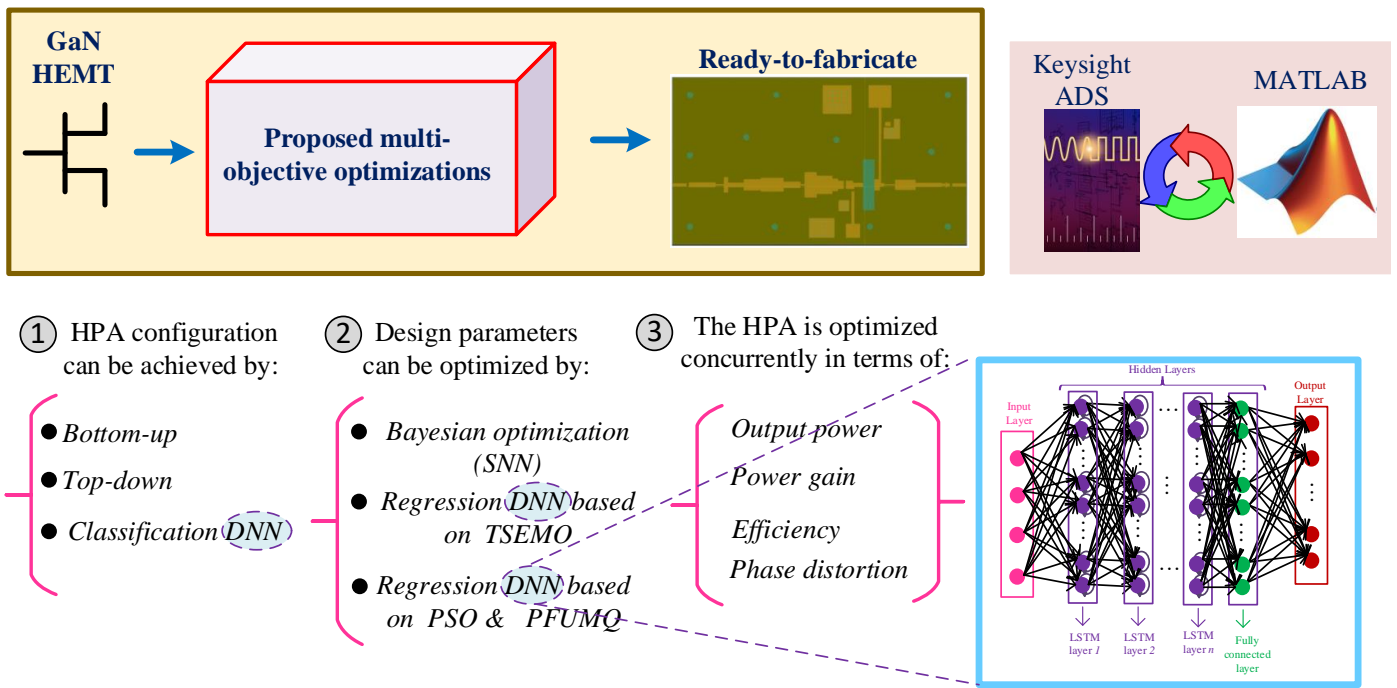


Fig. 31. An overview of the proposed optimization method where the transistor is selected, the EM-verified layout is generating automatically.

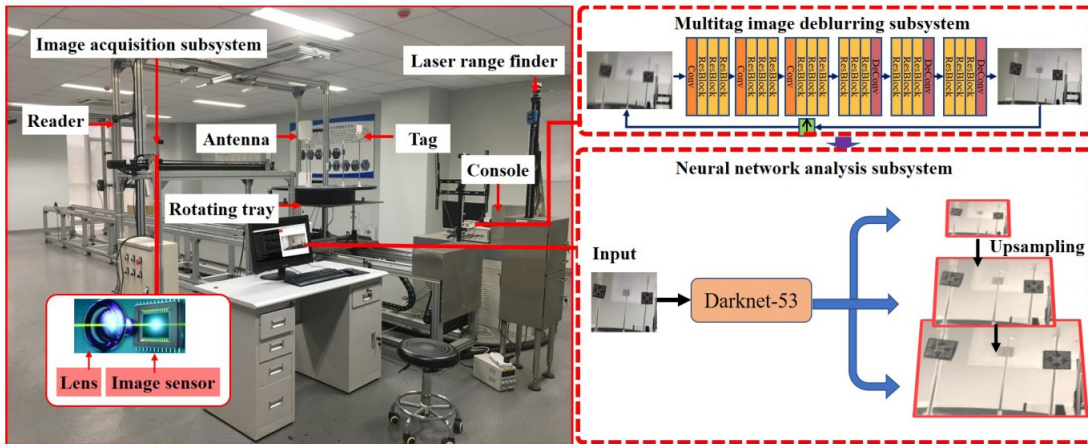


Fig. 32. Presented RFID dynamic performance measurement system through the DNN [67].

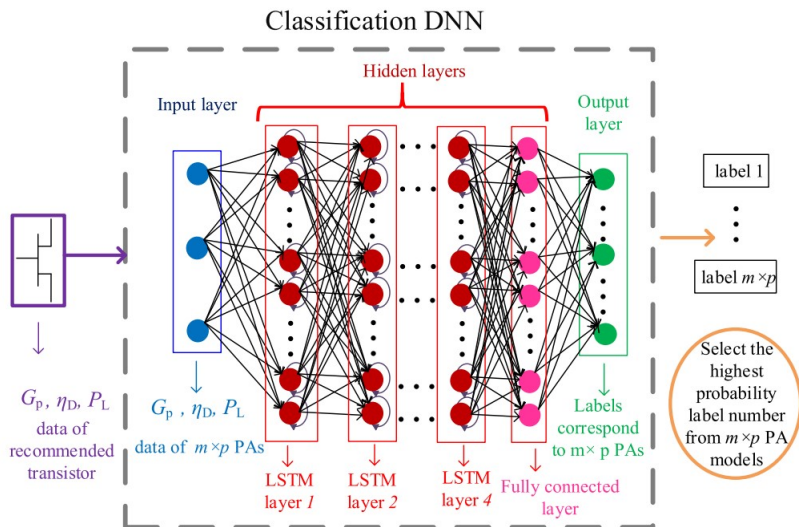


Fig. 33. Presented classification DNN for modeling the power amplifier [68].

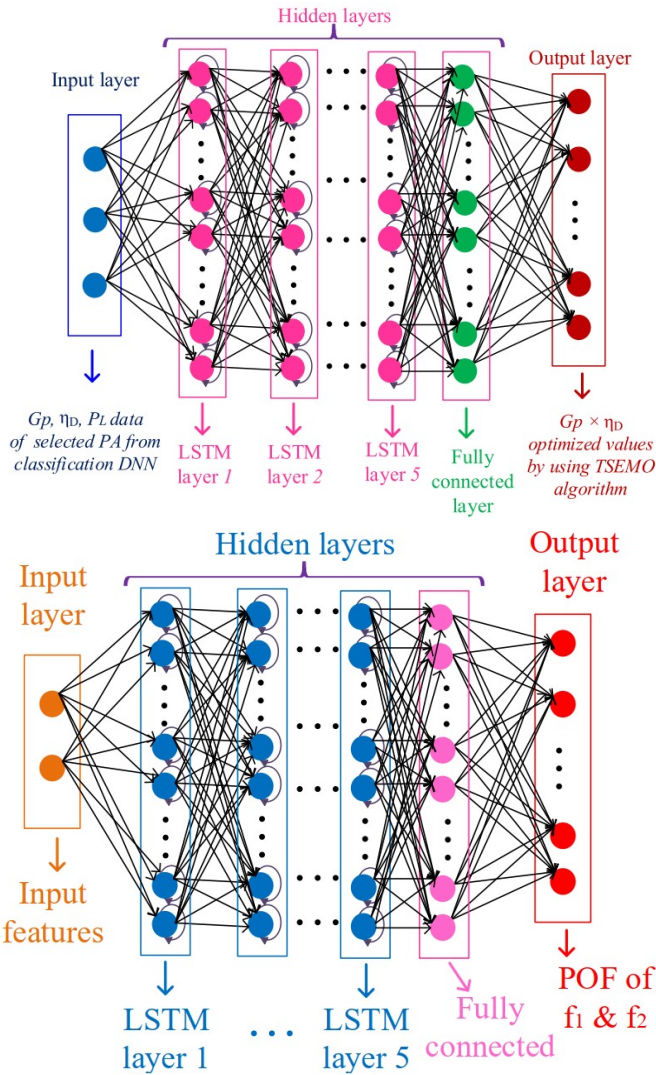


Fig. 34. a) Presented TSEMO-based DNN in [68] [top], and b) presented PFUMQ-based DNN in [69] where $f_1=AM$ -to- $AM(P_S)$ and $f_2=AM$ -to- $PM(P_S)$ [bottom].

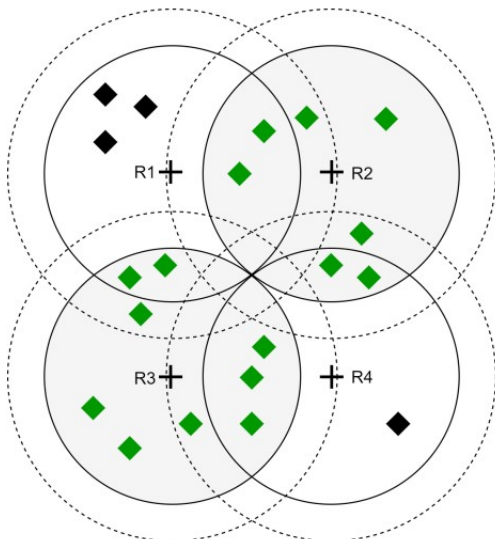


Fig. 35. Illustration of RFID system [70].

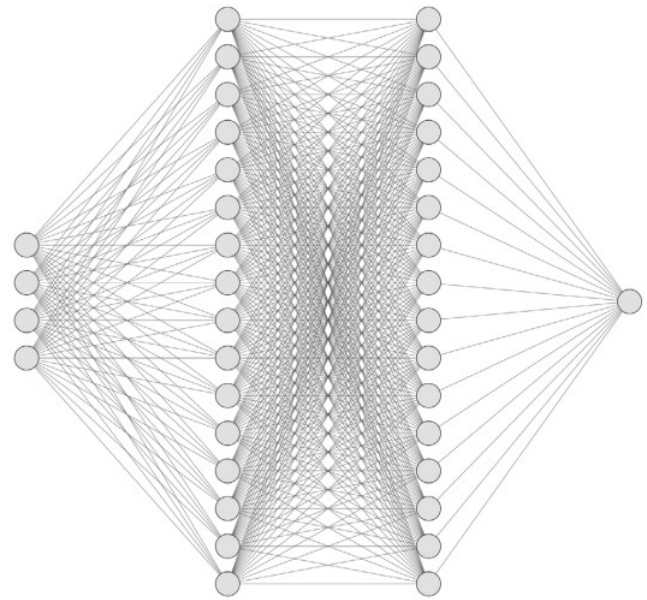


Fig. 36. Proposed flowchart in [70] with 4 neurons, 16 neurons, 16 neurons, and 1 neuron in the input layer, first hidden layer, second hidden layer, and output layer, respectively.

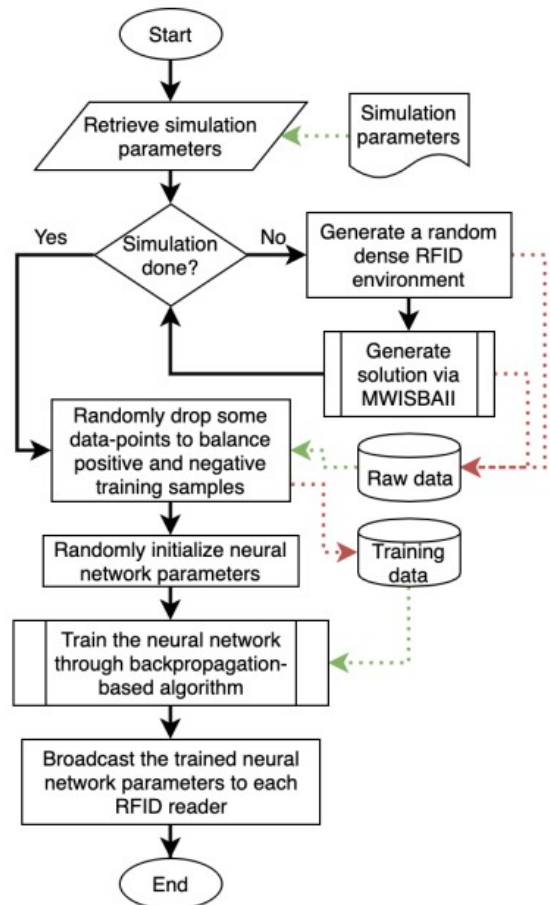


Fig. 37. Proposed flowchart in [70].

passive RFID data and it is appropriate for applications with large number of activity classes (see Fig. 40).

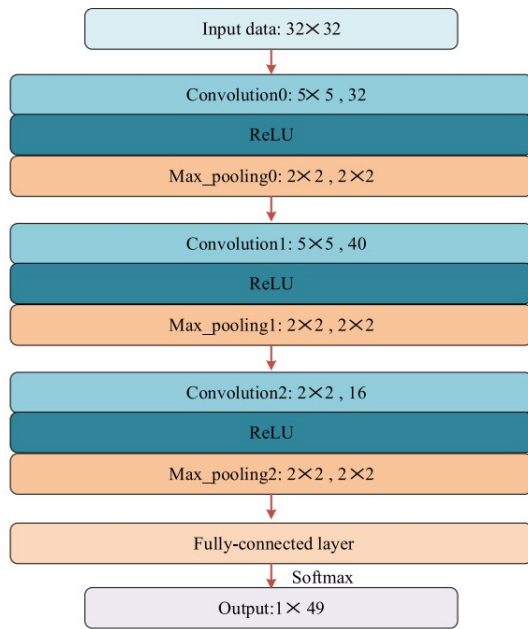


Fig. 38. Training procedure using the designed CNN model [71].

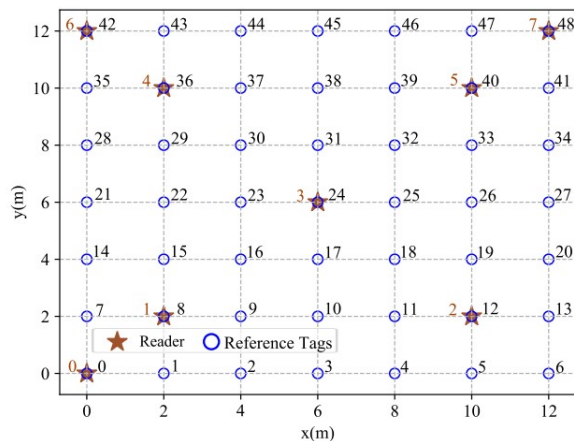


Fig. 39. Layout of readers and reference tags [71].

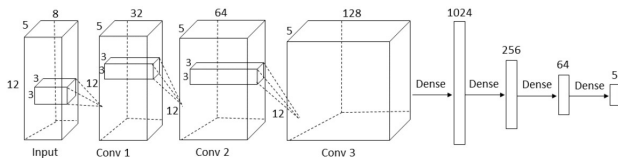


Fig. 40. Convolutional network with 3 convolutional layers and 3 fully connected layers [72].

IV. CONCLUSION

The RFID system plays an important role in the real world since it can capture and compute information automatically in the wireless communications. As presented in detail, the important blocks of RFID systems are amplifiers and antennas. For this case, this manuscript provides the comprehensive study on the recently published optimization methods used for designing RFID systems. That includes the diverse kinds of optimization methods and algorithms used for designing antennas and amplifiers. The engineers by studying this survey, will get a general idea of various algorithms and will determine a suitable method that can be applied for their design problems.

REFERENCES

- [1] Z. L. Oo, "IoT based LPG gas level detection & gas leakage accident prevention with alert system," pp. 404 – 409, 2021.
- [2] A. Tzitzis, A. Raptopoulos Chatzistefanou, T. V. Yioultis, and A. G. Dimitriou, "A real-time multi-antenna SAR-based method for 3D localization of RFID tags by a moving robot," *IEEE Journal of Radio Frequency Identification*, vol. 5, no. 2, pp. 207–221, 2021.
- [3] D. Le, S. Ahmed, L. Ukkonen, and T. Björninen, "A small all-corners-truncated circularly polarized microstrip patch antenna on textile substrate for wearable passive UHF RFID tags," *IEEE Journal of Radio Frequency Identification*, vol. 5, no. 2, pp. 106–112, 2021.
- [4] Z. Liu, Z. Fu, T. Li, I. White, R. Penty, and M. Crisp, "An ISAR-SAR based localization method using passive UHF RFID system with mobile robotic platform," in *2020 IEEE International Conference on RFID (RFID)*, 2020, pp. 1–7.
- [5] S. Qi, Y. Lu, W. Wei, and X. Chen, "Efficient data access control with fine-grained data protection in cloud-assisted IIoT," *IEEE Internet of Things Journal*, vol. 8, no. 4, pp. 2886–2899, 2021.
- [6] R. Seyfert, M. Maibaum, and S. Kroll, "RFID data storage and their key role in exploitation of metallic second life materials," *IEEE Journal of Radio Frequency Identification*, vol. 6, pp. 14–18, 2022.
- [7] J. Yu, P. Zhang, L. Chen, J. Liu, R. Zhang, K. Wang, and J. An, "Stabilizing frame slotted aloha-based IoT systems: A geometric ergodicity perspective," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 3, pp. 714–725, 2021.
- [8] Y. Zhang, K. Liu, Y. Ma, J. Wang, and S. Li, "Vehicular localization with using china electronic license plate system," *IEEE Journal of Radio Frequency Identification*, vol. 4, no. 4, pp. 322–331, 2020.
- [9] P. Avaltroni, S. Nappi, and G. Marrocco, "Antennifying orthopedic bone-plate fixtures for the wireless monitoring of local deep infections," *IEEE Sensors Journal*, vol. 21, no. 18, pp. 21 012–21 021, 2021.
- [10] M. Boada, A. Lazaro, R. Villarino, E. Gil-Dolcet, and D. Girbau, "Battery-less NFC bicycle tire pressure sensor based on a force-sensing resistor," *IEEE Access*, vol. 9, pp. 103 975–103 987, 2021.
- [11] Q. Liu, Y. Yu, D.-W. Wang, and G. Wang, "An RFID-based wireless multistate controller with quasi-isotropic radiation pattern for remote control applications," *IEEE Antennas and Wireless Propagation Letters*, vol. 20, no. 10, pp. 2023–2027, 2021.
- [12] S. Khan, W.-K. Lee, and S. O. Hwang, "A flexible gimli hardware implementation in FPGA and its application to RFID authentication protocols," *IEEE Access*, vol. 9, pp. 105 327–105 340, 2021.
- [13] D. Inserra and G. Wen, "Dual orthogonal port stacked patch antenna with vertical pins for simultaneous transmit and receive application," *IEEE Transactions on Antennas and Propagation*, vol. 69, no. 12, pp. 8908–8913, 2021.
- [14] W. Sabat, D. Klepacki, K. Kamuda, and K. Kurylo, "Analysis of electromagnetic field distribution generated in an semi-anechoic chamber in aspect of RF harvesters testing," *IEEE Access*, vol. 9, pp. 92 043–92 052, 2021.
- [15] Z. Wang, M. Xu, N. Ye, F. Xiao, R. Wang, and H. Huang, "Computer vision-assisted 3D object localization via COTS RFID devices and a monocular camera," *IEEE Transactions on Mobile Computing*, vol. 20, no. 3, pp. 893–908, 2021.
- [16] A. DiNatale, A. DiCarlofelice, and E. DiGiampaolo, "A crack mouth opening displacement gauge made with passive UHF RFID technology," *IEEE Sensors Journal*, vol. 22, no. 1, pp. 174–181, 2022.
- [17] M. Kokkonen, S. Myllymäki, J. Putaala, and H. Jantunen, "A resonator enhanced UHF RFID antenna cable for inventory and warehouse applications," *IEEE Journal of Radio Frequency Identification*, vol. 6, pp. 128–133, 2022.
- [18] A. Adeyeye, C. Lynch, A. Eid, J. Hester, and M. Tentzeris, "5.8-ghz low-power tunnel-diode-based two-way repeater for non-line-of-sight interrogation of RFIDs and wireless sensor networks," *IEEE Microwave and Wireless Components Letters*, vol. 31, no. 6, pp. 794–797, 2021.
- [19] M. Kim, U. Ha, Y. Lee, K. Lee, and H.-J. Yoo, "A 82nW chaotic-map true random number generator based on sub-ranging SAR ADC," in *ESSCIRC Conference 2016: 42nd European Solid-State Circuits Conference*, 2016, pp. 157–160.
- [20] V.-H. Duong, N. X. Hieu, H.-S. Lee, and J.-W. Lee, "A battery-assisted passive EPC Gen-2 RFID sensor tag IC with efficient battery power management and RF energy harvesting," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 11, pp. 7112–7123, 2016.
- [21] M. Nariman, F. Shirinfar, A. Papió Toda, S. Pamarti, A. Rofougaran, and F. De Flaviis, "A compact 60-GHz wireless power transfer system," *IEEE Transactions on Microwave Theory and Techniques*, vol. 64, no. 8, pp. 2664–2677, 2016.

- [22] B. Lee, M. Kiani, and M. Ghovanloo, "A smart wirelessly powered homeware for long-term high-throughput behavioral experiments," *IEEE Sensors Journal*, vol. 15, no. 9, pp. 4905–4916, 2015.
- [23] J. Lorenzo, A. Lazaro, R. Villarino, and D. Girbau, "Active backscatter transponder for FMCW radar applications," *IEEE Antennas and Wireless Propagation Letters*, vol. 14, pp. 1610–1613, 2015.
- [24] S. Khaledian, F. Farzami, H. Soury, B. Smida, and D. Erricolo, "Active two-way backscatter modulation: An analytical study," *IEEE Transactions on Wireless Communications*, vol. 18, no. 3, pp. 1874–1886, 2019.
- [25] M.-S. Kim, S.-C. Jung, J. Jeong, H. Kim, M. Seo, J. Ham, C.-S. Park, and Y. Yang, "Adaptive TX leakage canceler for the UHF RFID reader front end using a direct leaky coupling method," *IEEE Transactions on Industrial Electronics*, vol. 61, no. 4, pp. 2081–2087, 2014.
- [26] A. Bucciarelli, E. Olivetti, A. Adami, and L. Lorenzelli, "Design of experiment rational optimization of an inkjet deposition of silver on kapton," *IEEE Sensors Journal*, vol. 21, no. 23, pp. 26304–26310, 2021.
- [27] N. Cordeschi, F. De Rango, and M. Tropea, "Exploiting an optimal delay-collision tradeoff in CSMA-based high-dense wireless systems," *IEEE/ACM Transactions on Networking*, vol. 29, no. 5, pp. 2353–2366, 2021.
- [28] F. Bernardini, A. Buffi, A. Motroni, P. Nepa, B. Tellini, P. Tripicchio, and M. Unetti, "Particle swarm optimization in SAR-based method enabling real-time 3D positioning of UHF-RFID tags," *IEEE Journal of Radio Frequency Identification*, vol. 4, no. 4, pp. 300–313, 2020.
- [29] B.-Q. Zhao, H.-M. Wang, and P. Liu, "Safeguarding RFID wireless communication against proactive eavesdropping," *IEEE Internet of Things Journal*, vol. 7, no. 12, pp. 11587–11600, 2020.
- [30] X. Zhang, H.-X. Li, and H. S.-H. Chung, "Setup-independent UHF RFID sensing technique using multidimensional differential measurement," *IEEE Internet of Things Journal*, vol. 8, no. 13, pp. 10509–10517, 2021.
- [31] A. Tzitzis, S. Megalou, S. Siachalou, T. G. Emmanouil, A. Filotheou, T. V. Yioultis, and A. G. Dimitriou, "Trajectory planning of a moving robot empowers 3D localization of RFID tags with a single antenna," *IEEE Journal of Radio Frequency Identification*, vol. 4, no. 4, pp. 283–299, 2020.
- [32] H. Liu, Q. Chen, N. Pan, Y. Sun, Y. An, and D. Pan, "UAV stocktaking task-planning for industrial warehouses based on the improved hybrid differential evolution algorithm," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 1, pp. 582–591, 2022.
- [33] J. Xu, H. Sato, M. Motoyoshi, N. Suematsu, K. Yasui, and Q. Chen, "A low-loss and compact UHF RFID tag antenna for implanted denture," *IEEE Journal of Radio Frequency Identification*, vol. 6, pp. 1–7, 2022.
- [34] "The radio frequency identification system architecture," <http://http://smart-tags.com/#what-is-rfid?>, accessed: 2022-01-01.
- [35] C. Bredendiek, D. A. Funke, J. Schöpfel, V. Kloubert, B. Welp, K. Aufinger, and N. Pohl, "A 61-GHz SiGe transceiver frontend for energy and data transmission of passive RFID single-chip tags with integrated antennas," *IEEE Journal of Solid-State Circuits*, vol. 53, no. 9, pp. 2441–2453, 2018.
- [36] J. Zhu, C. Jin, and H. Liu, "Mutual inductance modeling of two coaxial solenoid antennas with large ferrite core for underground RFID applications," *IEEE Transactions on Magnetics*, vol. 57, no. 8, pp. 1–4, 2021.
- [37] G. M. Bianco, S. Amendola, and G. Marrocco, "Near-field constrained design for self-tuning UHF-RFID antennas," *IEEE Transactions on Antennas and Propagation*, vol. 68, no. 10, pp. 6906–6911, 2020.
- [38] S. Ahmed, D. Le, L. Sydänheimo, L. Ukkonen, and T. Björninen, "Wearable metasurface-enabled quasi-yagi antenna for UHF RFID reader with end-fire radiation along the forearm," *IEEE Access*, vol. 9, pp. 77229–77238, 2021.
- [39] Y. Chen, C. Hua, and Z. Shen, "Circularly polarized UHF RFID tag antenna for wireless sensing of complex permittivity of liquids," *IEEE Sensors Journal*, vol. 21, no. 23, pp. 26746–26754, 2021.
- [40] S. Khaledian, F. Farzami, D. Erricolo, and B. Smida, "A full-duplex bidirectional amplifier with low DC power consumption using tunnel diodes," *IEEE Microwave and Wireless Components Letters*, vol. 27, no. 12, pp. 1125–1127, 2017.
- [41] D. Zouache, Y. Ould Arby, F. Nouioua, and F. Ben Abdelaziz, "Multi-objective chicken swarm optimization: A novel algorithm for solving multi-objective optimization problems," *Computers Industrial Engineering*, vol. 129, pp. 377–391, 2019. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0360835219300725>
- [42] W. Shi, W. Wang, Y. Yu, S. Zhang, Y. Cao, S. Yan, and J. Gao, "Optimal deployment of phased array antennas for RFID network planning based on an improved chicken swarm optimization," *IEEE Internet of Things Journal*, vol. 8, no. 19, pp. 14572–14588, 2021.
- [43] S. Cheng, Y. Shi, and Q. Qin, *Population Diversity of Particle Swarm Optimization Algorithm on Solving Single and Multi-Objective Problems*, ser. Handbook of Research on Advancements of Swarm Intelligence Algorithms for Solving Real-World Problems. Hershey, PA, USA: IGI Global, 2020, pp. 312–344. [Online]. Available: <https://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/978-1-7998-3222-5.ch014>
- [44] B. Cao, Y. Gu, Z. Lv, S. Yang, J. Zhao, and Y. Li, "RFID reader anticollision based on distributed parallel particle swarm optimization," *IEEE Internet of Things Journal*, vol. 8, no. 5, pp. 3099–3107, 2021.
- [45] F. Passos, E. Roca, J. Sieiro, R. Fiorelli, R. Castro-López, J. M. López-Villegas, and F. V. Fernández, "A multilevel bottom-up optimization methodology for the automated synthesis of RF systems," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 39, no. 3, pp. 560–571, 2020.
- [46] L. Kouhalvandi, O. Ceylan, and S. Ozoguz, "Automated matching network modeling and optimization for power amplifier designs," in *2019 11th International Conference on Electrical and Electronics Engineering (ELECO)*, 2019, pp. 510–513.
- [47] F. Mir, L. Kouhalvandi, L. Matekovits, and E. O. Gunes, "Electromagnetic bottom-up optimization for automated antenna designs," in *2020 IEEE Asia-Pacific Microwave Conference (APMC)*, 2020, pp. 792–794.
- [48] F. Mir, L. Matekovits, L. Kouhalvandi, and E. O. Gunes, "Optimization for wideband linear array antenna through bottom-up method," in *2020 IEEE International Conference on Electrical Engineering and Photonics (EExPolytech)*, 2020, pp. 51–54.
- [49] R. González-Echevarría, E. Roca, R. Castro-López, F. V. Fernández, J. Sieiro, J. M. López-Villegas, and N. Vidal, "An automated design methodology of RF circuits by using pareto-optimal fronts of EM-simulated inductors," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 36, no. 1, pp. 15–26, 2017.
- [50] M. H. Bakr, J. W. Bandler, K. Madsen, and J. Søndergaard, "An introduction to the space mapping technique," *Optimization and Engineering*, vol. 2, no. 4, pp. 369–384, Dec 2001. [Online]. Available: <https://doi.org/10.1023/A:1016086220943>
- [51] S. Fidler, M. Boben, and A. Leonardis, "A bottom-up and top-down optimization framework for learning a compositional hierarchy of object classes," in *2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 2009, pp. 3–3.
- [52] L. Kouhalvandi, O. Ceylan, and S. Ozoguz, "Automated top-down pruning optimization approach in RF power amplifier designs," *Analog Integrated Circuits and Signal Processing*, vol. 106, no. 3, pp. 525–534, Mar 2021. [Online]. Available: <https://doi.org/10.1007/s10470-020-01730-w>
- [53] L. Kouhalvandi, L. Matekovits, and I. Peter, "Deep learning assisted automatic methodology for implanted MIMO antenna designs on large ground plane," *Electronics*, vol. 11, no. 1, 2022. [Online]. Available: <https://www.mdpi.com/2079-9292/11/1/147>
- [54] S. M. Aslam and S. Samreen, "Gesture recognition algorithm for visually blind touch interaction optimization using crow search method," *IEEE Access*, vol. 8, pp. 127560–127568, 2020.
- [55] M. Hussein and F. Özyurt, "A new technique for sentiment analysis system based on deep learning using Chi-Square feature selection methods," *Balkan Journal of Electrical and Computer Engineering*, vol. 9, no. 4, pp. 320–326, 2021.
- [56] H. Yin, J. Zhai, P. Chen, and C. Yu, "Directed graph navigated digital predistortion of mmWave power amplifiers for 6G hopping applications," *IEEE Microwave and Wireless Components Letters*, vol. 31, no. 11, pp. 1235–1238, 2021.
- [57] N. L. K. Nguyen, N. S. Killeen, D. P. Nguyen, A. N. Stameroff, and A.-V. Pham, "A wideband gain-enhancement technique for distributed amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 68, no. 9, pp. 3697–3708, 2020.
- [58] L. Kouhalvandi, M. Piroola, and S. Ozoguz, "Automated two-step power amplifier design with pre-constructed artificial neural network," in *2020 43rd International Conference on Telecommunications and Signal Processing (TSP)*, 2020, pp. 617–620.
- [59] S. Yarman, "Design of ultra wideband power transfer networks," *New York, NY, USA: Wiley*, 2010.
- [60] F. Mir, L. Kouhalvandi, L. Matekovits, and E. O. Gunes, "Automated optimization for broadband flat-gain antenna designs with artificial neural network," *IET Microwaves, Antennas & Propagation*, vol. 15, no. 12, pp. 1537–1544, 2021. [Online]. Available: <https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/mia2.12137>
- [61] J. Nan, H. Xie, M. Gao, Y. Song, and W. Yang, "Design of UWB antenna based on improved deep belief network and extreme learning machine surrogate models," *IEEE Access*, vol. 9, pp. 126541–126549, 2021.

- [62] J. Gao, Y. Tian, and X. Chen, "Antenna optimization based on co-training algorithm of gaussian process and support vector machine," *IEEE Access*, vol. 8, pp. 211 380–211 390, 2020.
- [63] W. Su, J. Zhu, H. Liao, and M. M. Tentzeris, "Wearable antennas for cross-body communication and human activity recognition," *IEEE Access*, vol. 8, pp. 58 575–58 584, 2020.
- [64] L. Kouhalvandi, O. Ceylan, and S. Ozoguz, "Automated RF power amplifier optimization and design: From lumped elements to distributed elements," in *2019 27th Telecommunications Forum (TELFOR)*, 2019, pp. 1–4.
- [65] J. Jin, C. Zhang, F. Feng, W. Na, J. Ma, and Q. Zhang, "Deep neural network technique for high-dimensional microwave modeling and applications to parameter extraction of microwave filters," *IEEE Transactions on Microwave Theory and Techniques*, vol. 67, no. 10, pp. 4140–4155, 2019.
- [66] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–44, 05 2015.
- [67] L. Li, X. Yu, Z. Liu, Z. Zhao, K. Zhang, and S. Zhou, "RFID dynamic performance measurement system embedded in multiscale deep learning," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–12, 2021.
- [68] L. Kouhalvandi, O. Ceylan, and S. Ozoguz, "Automated deep neural learning-based optimization for high performance high power amplifier designs," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 67, no. 12, pp. 4420–4433, 2020.
- [69] —, "Multi-objective efficiency and phase distortion optimizations for automated design of power amplifiers through deep neural networks," in *2021 IEEE MTT-S International Microwave Symposium (IMS)*, 2021, pp. 233–236.
- [70] P. Yan, S. Choudhury, and R. Wei, "A machine learning auxiliary approach for the distributed dense RFID readers arrangement algorithm," *IEEE Access*, vol. 8, pp. 42 270–42 284, 2020.
- [71] C. Peng, H. Jiang, and L. Qu, "Deep convolutional neural network for passive RFID tag localization via joint RSSI and PDOA fingerprint features," *IEEE Access*, vol. 9, pp. 15 441–15 451, 2021.
- [72] X. Li, Y. Zhang, I. Marsic, A. Sarcevic, and R. Burd, "Deep learning for RFID-based activity recognition," vol. 2016, 11 2016.



Lida Kouhalvandi, IEEE member & PhD, joined the Department of Electrical and Electronics Engineering at Dogus University as an assistant professor in October 2021. She received her PhD in Electronics Engineering in 2021 from the Istanbul Technical University, Istanbul, Turkey. She received her MSc in Electronics Engineering in 2015 from the Istanbul Technical University, Istanbul, Turkey, and her BSc in Electronics Engineering in 2011 from the Azad University of Tabriz, Tabriz, Iran. In recognition of her research, she received the Doctoral Fellowship

at Department of Electronics and Telecommunications, Politecnico di Torino, Turin, Italy from 2019 to 2020 and also she joined to Politecnico di Torino, Turin, Italy as a Research fellowship from February 2021 up to July 2021. Dr. Kouhalvandi's research interests as a radio frequency and analog engineer are power amplifier, antenna, analog designs, and implantable medical devices. She also has experience in computer-aided designs and optimization algorithms through machine learning. she received 'Best Paper Award' from EExPolytech-2021: Electrical Engineering and Photonics conference in 2021. Additionally, her PhD thesis accepted for the presentation at PhD Forum of the IEEE/ACM Design Automation Conference (DAC) in 2021.

Performance Analysis of Three Level Three Switch Vienna-Type Rectifier based on Direct Power Control

Doğan Çelik

Abstract—The Vienna-type rectifier is broadly employed in various implementations thanks to realizable three-level operation, simple structure and controllable DC-link voltage. This paper proposes an adaptive proportional integral (PI) with anti-windup based direct power control (DPC) to improve dynamic response during load change, DC-link voltage step change and start up and to reduce current tracking errors. An adaptive PI with anti-windup DC-link voltage control is utilized to enhance dynamic response of the DC-link voltage and the instantaneous power because the PI regulator affected by the load change and system parameter variations. The PI based conventional power control is used to indicate the availability and accuracy of the proposed control strategy. The proposed control strategy performs outstanding performance and copes with vigorous disturbances in contrast to the conventional strategy. A detailed theoretical analysis of the proposed approach has been performed. Extensive case studies such as current tracking signals, DC-link voltage step change and load power increase and decrease carried out by PSIM software are performed to validate the excellent behavior of the proposed control strategy.

Index Terms—DC-link control; Direct power control; PI with anti-windup; Vienna-type rectifier.


I. INTRODUCTION

VIENNA-TYPE rectifier provides, simple circuit structure, compact size, low voltage stress and achieves a switching voltage that is only half the output voltage. Conventional uncontrolled diode-bridge rectifiers (UDBRs) provide low cost and high reliability. However, the U-DBRs have some limitations uncontrolled power factor, higher total harmonic distortion (THD) and lower efficiency. Therefore, six-switch two levels PWM has been prevailed one of the options owing to the disadvantages of the UDBRs [1]. However, compared to the conventional converters, the Vienna-type rectifiers achieves higher efficiency, less voltage stress, used less power switches, having unity power factor correction, higher reliability and lower cost. It can be utilized

in the areas of electric vehicle (EVs) charging, motor drives, uninterruptible power system, hybrid AC/DC energy systems and wind power generation. Besides, the stable and constant DC output voltage is provided for DC/DC power converters in regards to EVs charging [2-6]. The control of the DC-link voltage, power and power factor is considerable important for Vienna-type rectifiers [7].

The operating principle and control of Vienna type rectifiers differ from the traditional rectifiers. The three-level and three switches Vienna-type rectifiers result in unbalance DC capacitor voltages [8]. Voltage and current closed-loop control is commonly utilized for Vienna-type rectifiers. The current control loop is mainly controlled by hysteresis current control (HCC). On the other hand, the proportional integral (PI) control generally has been used by many researchers for the voltage outer loop. However, the slow dynamic response, the linear summation of error and the voltage overrun are some shortcomings [2]. The rectifiers are usually influenced by the external disturbance and changing operation conditions. Various control approaches have been conducted for Vienna-type rectifiers such as HCC method [9, 10], carrier based-discontinuous pulse width modulation (PWM) [11], a lagging reactive power compensation method [12] and a model predictive control (MPC) [13-15]. In these control algorithms, classical PI controller has been employed to regulate the DC-link voltage. However, direct power control (DPC) achieves superior performance for Vienna-type rectifiers compared to the above-mentioned control methods [16]. In [17], authors have presented a DPC based MPC for AC/DC PWM rectifier, but it is affected by variation of the system parameters. On the other hand, sliding mode control (SMC) based control algorithms provision better dynamics and robustness, and also can lessen the overshoots for various disturbance conditions in contrast to the classical PI controller. In [18], Yang et al. have proposed a SMC based control strategy for voltage source rectifier. In [19], Hang et al. have discussed the PI based space vector PWM strategy for Vienna-type rectifier. In [20], authors have addressed deadbeat based MPC for Vienna-type rectifier. Furthermore, to mitigate source current harmonics, a proportional resonant based control approach is discussed for a three-level Vienna rectifier [21].

DOĞAN ÇELİK, is with Department of Electrical and Electronic Engineering University of Van Yuzuncu Yıl University, Van, Turkey, (e-mail: dogancelik@yyu.edu.tr).

 <https://orcid.org/0000-0002-8348-130X>

Manuscript received February 17, 2022; accepted April 19, 2022.

DOI: [10.17694/bajece.1072287](https://doi.org/10.17694/bajece.1072287)

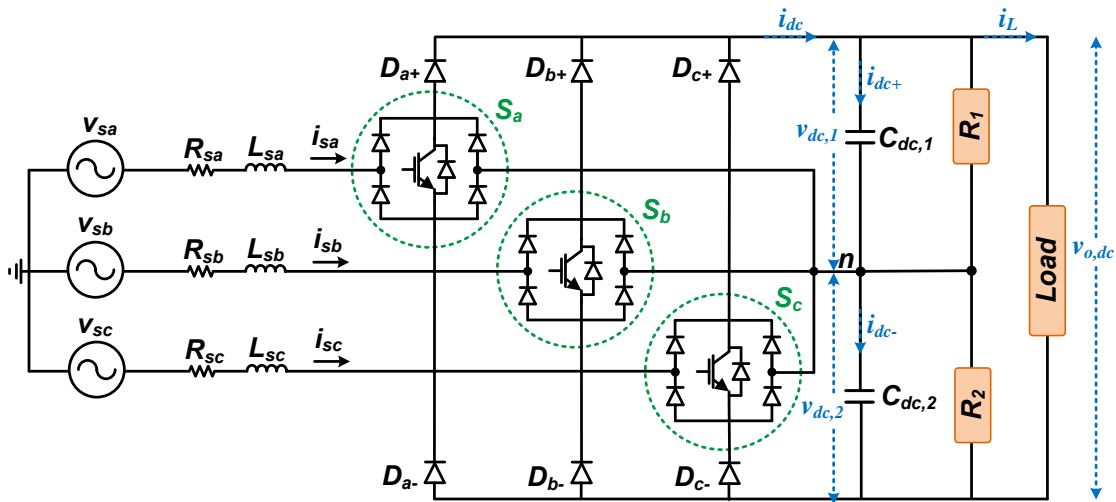


Fig.1 Schematic diagram of the Vienna rectifier topology with dual DC loads.

In [22], He et al. have addressed a hybrid discontinuous PWM for three-level to two-level conversion based Vienna rectifier. A hybrid control scheme including MPC and PI is handled to compensate reactive power and regulate DC-link voltage [23].

This paper proposes a control strategy with a simple and fast dynamic response for three switches three levels Vienna-type rectifier. The DPC has been used in the inner current loop and adaptive PI with anti-windup DC-link control has been adopted in the outer voltage loop, which dealt with slow dynamic response and voltage overshoot in the PI control. Various case results carried out by PSIM software has been conducted to show good robustness and fast tracking performance of the proposed control strategy.

This paper is divided as: The topology of Vienna-type rectifier is discussed in Section II. Section III presents the proposed control strategy. Performance evaluation of the conventional and proposed control strategies are conducted in Section IV. Section V provides main findings of this paper.

II. DESCRIPTION OPERATIONAL MODE OF VIENNA-TYPE RECTIFIER

The circuit topology of the Vienna-type rectifier represented in Figure 1 includes a three bidirectional switching units (S_a , S_b and S_c) and main diode bridge. The switching units consist of active switches and diode rectifiers, which are controlled to provide steady DC voltage, sinusoidal AC current and neutral point voltage balance [24]. $C_{dc,1}$ and $C_{dc,2}$ divided by two parts as the upper and lower the DC-link capacitors, respectively. The DC-side is tied to the load and the AC-side of the Vienna-type rectifier is tied to the utility grid. $v_{dc,1}$ and $v_{dc,2}$ represent upper and lower the DC-link voltages, respectively.

The operational states of the Vienna rectifier are presented in Figure 2. The polarity of line currents and the ON/OFF states of the switches at any instant of operation have stated the rectifier pole voltages with supposing a continuous conduction mode. With the positive line current (i_{sa}) with the

controlled OFF switch S_a , the voltage between the rectifier pole (A) and the DC-link neutral point (n) is $v_{dc}/2$ as seen in Fig.2a. With the positive line current with the controlled ON switch S_a , the line to neutral voltage v_{an} is 0 (see Fig. 2b). With the negative line current (i_{sa}) with the controlled OFF or ON switch, the line to neutral voltage v_{an} can be $-v_{dc}/2$ or zero. Similarly, this operating procedure can be applied to phases B and C [9] (see Fig. 2c and 2d). Besides, the switching state of Vienna-type rectifier is given in Table I.

The grid current dynamics can be expressed for the AC-side of a Vienna-type rectifier.

$$L_{s,x} \frac{di_{s,x}}{dt} = v_{s,x} - R_{sx} i_{s,x} - (v_{n,x} + v_{n0}) \quad (1)$$

where x represents a-b-c phases. L_s and R_s denote the input inductance and resistance, respectively. It is supposed that $C_{dc,1} = C_{dc,2} = C_{dc}$, and neutral-point voltage is determined with the difference between v_{dc1} and v_{dc2} .

$$\begin{cases} C_{dc,1} \frac{dv_{dc,1}}{dt} = i_{dc+} - \frac{v_{dc,1} + v_{dc,2}}{R_1} \\ C_{dc,2} \frac{dv_{dc,2}}{dt} = i_{dc-} - \frac{v_{dc,1} + v_{dc,2}}{R_2} \end{cases} \quad (2)$$

$v_{x,n}$ is specified by the direction of input AC current and the switching states of bidirectional switches because of the operational characteristics of Vienna-type rectifier [24]. $v_{x,n}$ is written by;

$$v_{x,n} = (1 - S_x) \left(\frac{\text{sign}(i_x) + 1}{2} v_{dc,1} + \frac{\text{sign}(i_x) - 1}{2} v_{dc,2} \right) \quad (3)$$

where $\text{sign}(\cdot)$ defines the sign function.

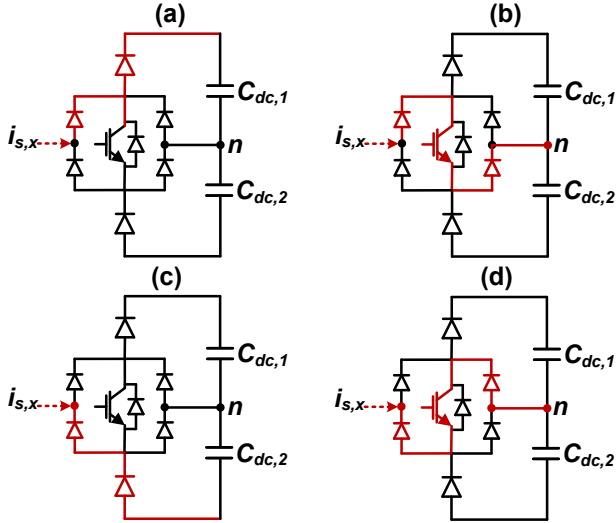


Fig. 2. Current directions for phase units ($x = a, b, c$): a) positive line current with the controlled OFF switch, b) positive line current with the controlled ON switch, c) negative line current with the controlled OFF switch and d) negative line current with the controlled ON switch.

TABLE I
SWITCHING STATE OF VIENNA-TYPE RECTIFIER

S_a	S_b	S_c	v_{an}	V_{bn}	V_{cn}
0	0	0	$+v_{dc}/2$	$-v_{dc}/2$	$-v_{dc}/2$
0	0	1	$+v_{dc}/2$	$-v_{dc}/2$	0
0	1	0	$+v_{dc}/2$	0	$-v_{dc}/2$
0	1	1	$+v_{dc}/2$	0	0
1	0	0	0	$-v_{dc}/2$	$-v_{dc}/2$
1	0	1	0	$-v_{dc}/2$	0
1	1	0	0	0	$-v_{dc}/2$
1	1	1	0	0	0

III. THE PROPOSED CONTROL STRATEGY

The proposed control strategy consists of two parts as DC-link voltage and current control loops as depicted in Figure 3. The proposed control strategy has fulfilled the following requirements: providing fast dynamic response under load and DC voltage variations, reducing current tracking errors, low current harmonics and constant switching frequency operation.

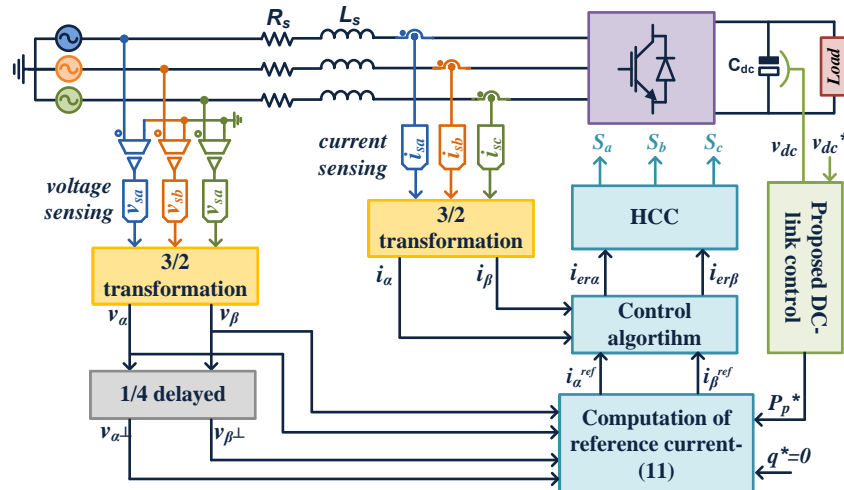


Fig. 3 The entire proposed control strategy.

A. Adaptive PI with anti-windup DC-link control

Dynamic performances of the DC-link have a considerable significance in the functioning of a rectifier. The dynamic nature of the rectifier is great important for the DC-link regulator to less affect to system parameters, load change and main voltage. In this paper, an adaptive PI with anti-windup control is employed as the DC-link voltage outer loop controller of Vienna-type rectifier to deal with instability and variations of the DC-link voltage. The output of the adaptive PI with anti-windup is used to acquire the active power reference. Output power at the DC-side is equal to the AC-side input power with neglecting power losses [25]. The dynamics of the DC-link voltage is expresses by (4) as;

$$C_{dc} \frac{dv_{dc}}{dt} = i_{dc} - i_L = \frac{P_p^*}{v_{dc}} - i_L \quad (4)$$

where i_L denotes load current. Figure 4 describes the proposed adaptive PI with anti-windup DC-link control and the system dynamics. The control diagram of the DC-link voltage can be designed based on (5). The integral state \dot{q} is written as;

$$\dot{q} = \begin{cases} \Delta v_{dc} & u = u^* \\ \Delta v_{dc} k_i k_{aw} (u = u^*) & u \neq u^* \end{cases} \quad (5)$$

where $u = u^*$ and $u \neq u^*$ represents linear region and saturation region, respectively. Δv_{dc} is voltage tracking error. Anti-windup gain k_{aw} is commonly chosen as $1/k_p$ [26-31]. Based on Figure 4, the closed-loop function for the DC-link can be obtained as (6);

$$u(s) = \frac{k_p s + k_i}{s - k_{aw}} \Delta v_{dc} - \frac{k_{aw}}{s - k_{aw}} u^* \quad (6)$$

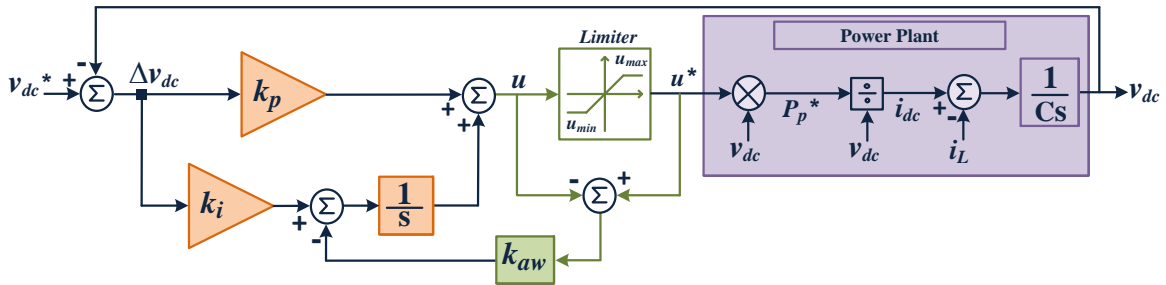


Fig. 4 Adaptive PI with anti-windup control diagram of the DC-link voltage.

B. Reference current generation

The control aim for Vienna-type rectifiers is to accurately and quickly track the references of the DC-link voltage and power/current signals. x and x_{\perp} are variable in $\alpha\beta$ stationary coordinate and lagging 90° of x is x_{\perp} . Original vector and its delayed value can be utilized to reduce the calculation burden. According to [32-35], $x_{\alpha\beta}^p$ and $x_{\alpha\beta}^n$ sequences are defined by (7);

$$\begin{pmatrix} x \\ x_{\perp} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -j & j \end{pmatrix} \begin{pmatrix} x_{\alpha\beta}^p \\ x_{\alpha\beta}^n \end{pmatrix} \quad (7)$$

The inverse of (7) is given by (8);

$$\begin{pmatrix} x_{\alpha\beta}^p \\ x_{\alpha\beta}^n \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & j \\ 1 & -j \end{pmatrix} \begin{pmatrix} x \\ x_{\perp} \end{pmatrix} \quad (9)$$

The active power reference P_p^* is obtained by the adaptive PI with anti-windup DC-link regulation. Reference reactive power is chosen as zero for both controllers to provide unity power factor. The reference current signals can be obtained by original voltage vectors with their 90° lagging signals.

$$\begin{pmatrix} i_{\alpha}^* \\ i_{\beta}^* \\ i_{\alpha\perp}^* \\ i_{\beta\perp}^* \end{pmatrix} = \frac{P_p^*}{\Delta_p} \begin{pmatrix} v_{\beta\perp} \\ -v_{\alpha\perp} \\ -v_{\beta} \\ v_{\alpha} \end{pmatrix} \quad (9)$$

where

$$\Delta_p = v_{\alpha} v_{\beta\perp} - v_{\beta} v_{\alpha\perp} \quad \Delta_p \neq 0 \quad (10)$$

The proposed reference current signals are obtained in $\alpha\beta$ -stationary coordinates as;

$$i^{ref} = i_{\alpha} + j i_{\beta} = \frac{P_p^* (v_{\beta\perp} - j v_{\alpha\perp})}{v_{\alpha} v_{\beta\perp} - v_{\beta} v_{\alpha\perp}} \quad (11)$$

On the other hand, the conventional reference current

signals can be created in $\alpha\beta$ -stationary coordinates as [36];

$$i^{ref} = i_{\alpha} + j i_{\beta} = \frac{P_c^* (v_{\alpha} + j v_{\beta})}{v_{\alpha} v_{\alpha} + v_{\beta} v_{\beta}} \quad (12)$$

where P_c^* is created based on the conventional PI control.

IV. RESULTS AND DISCUSSIONS

Performance Analysis of Three Switches Three Levels Vienna-Type Rectifier has been simulated in PSIM software environment to demonstrate the effectiveness of the proposed control strategy. The parameters for the proposed system are given in Table II. Two control strategies are applied in Vienna-type rectifier system to achieve validity of the proposed control strategy. Switching signals for two control strategies are created based on the HCC.

- One strategy is conventional control that includes conventional PI DC-link control and conventional reference current generation.

- Another strategy is proposed control that includes adaptive PI with anti-windup DC-link control and DPC.

TABLE II
SIMULATION PARAMETERS

Parameters	Values
Source line voltage v_a, v_b, v_c (rms)	220V
Source frequency	50Hz
DC-link capacitance	$C_{dc,1}=C_{dc,l}=1000\mu\text{F}$
Load	$5\Omega/15\Omega$
Filter inductance	1mH

The steady-state, dynamic response and current tracking errors waveforms of the conventional and proposed control strategies are depicted in Figure 5. It is indicated that the proposed control strategy can accomplish good tracking of current references with nearly 8ms while the conventional control has slower dynamic response with about 30ms (see Figure 5a and 5b). Besides, current tracking errors are effectively minimized by the proposed control.

Figure 6 depicts the comparison of the results between conventional and the proposed control under increasing dynamic load power (sudden change of R_L value from 10 to 6.67Ω). While the DC-link voltage is restored with the proposed control to the given value within nearly 10ms, the conventional control has slower dynamic response with about 50ms to provide to the given value.

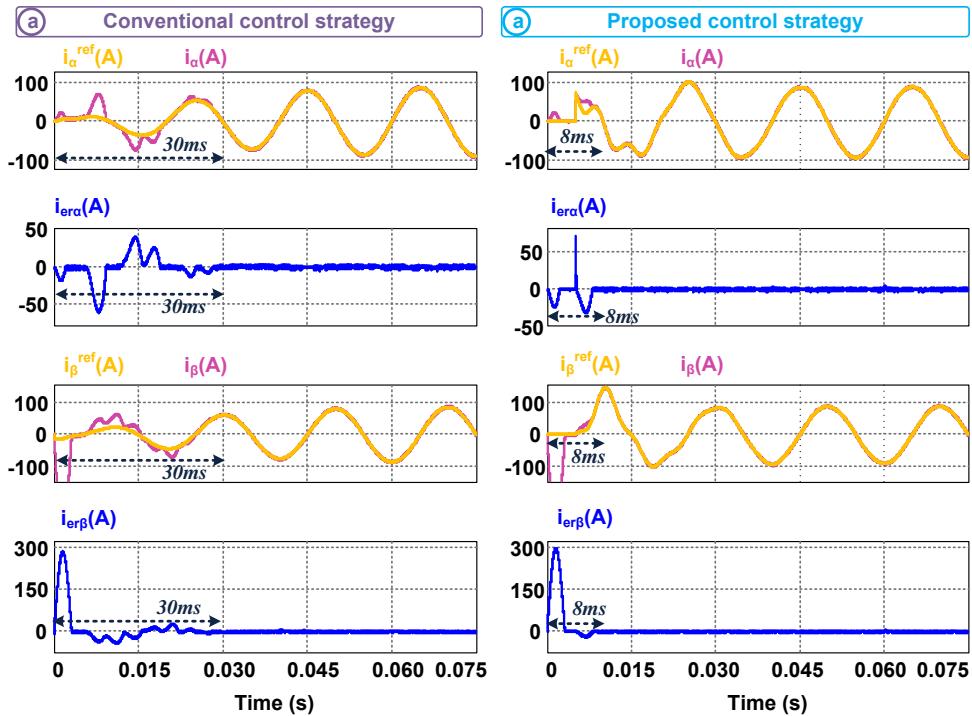


Fig. 5 Current tracking signals for a) conventional control and b) proposed control

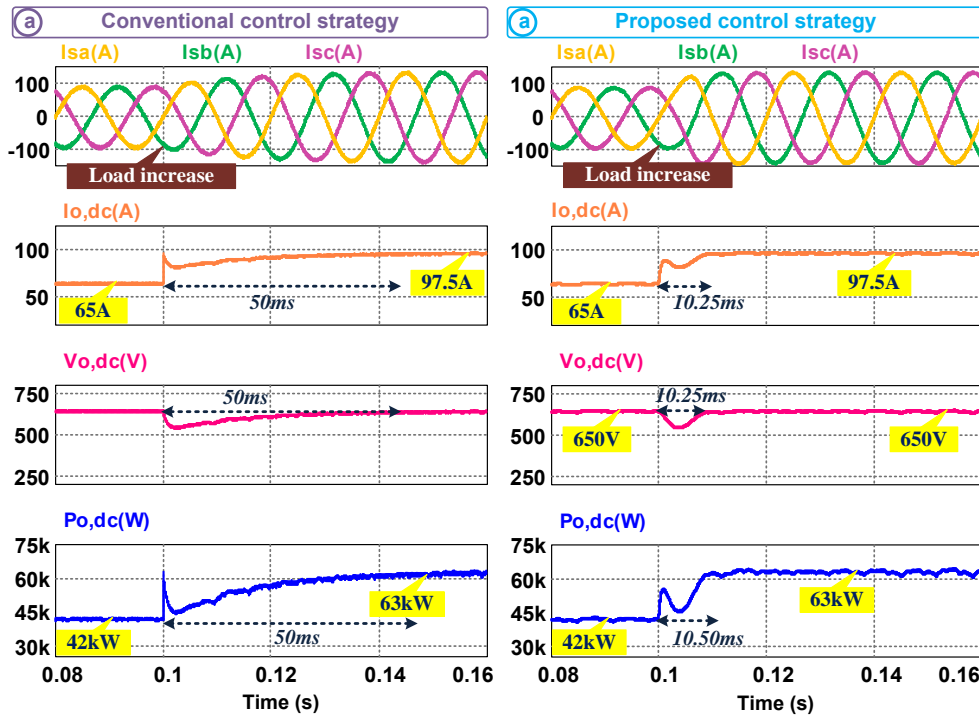


Fig. 6 Load power increase for a) conventional control and b) proposed control

The effectiveness of the proposed control strategy is also examined under load power decrease (sudden change of R_L value from 10 to 15Ω) as depicted in Figure 7. With the proposed control, the output current, DC-link voltage and output power soon achieve a steady state within 25ms while with the conventional control, these signals has been reached a steady state within 45ms. The results indicate that the proposed control provisions strong robustness against sudden load disturbances in contrast to the conventional control.

Figure 8 reveals the control of the DC link voltage ($V_{o,dc}$) is achieved at 650V, 700V and 625V after the step change occurs in reference DC-link voltage (V_{dc}^*). The fast regulation of the DC-link voltage is effectively accomplished with 25ms for the proposed control under this operation while the dynamic response with the conventional control is slower with 45ms as shown in Figure 8a. While the current THD of the conventional control is 1.7%, the current THD is only 0.9% with using the proposed control.

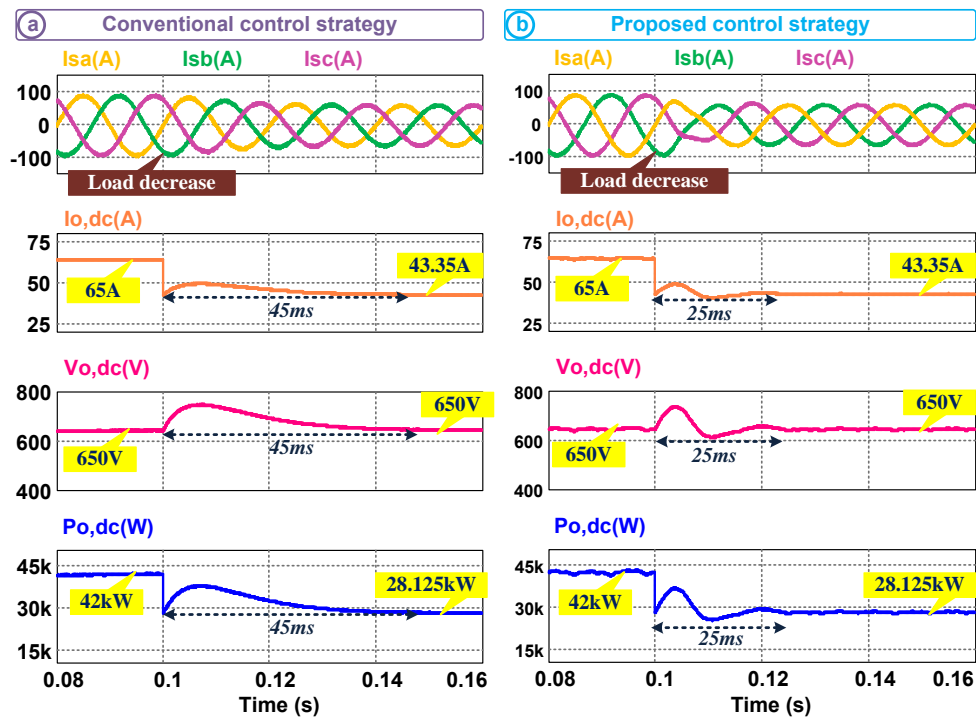


Fig. 7 Load power decrease for a) conventional control and b) proposed control

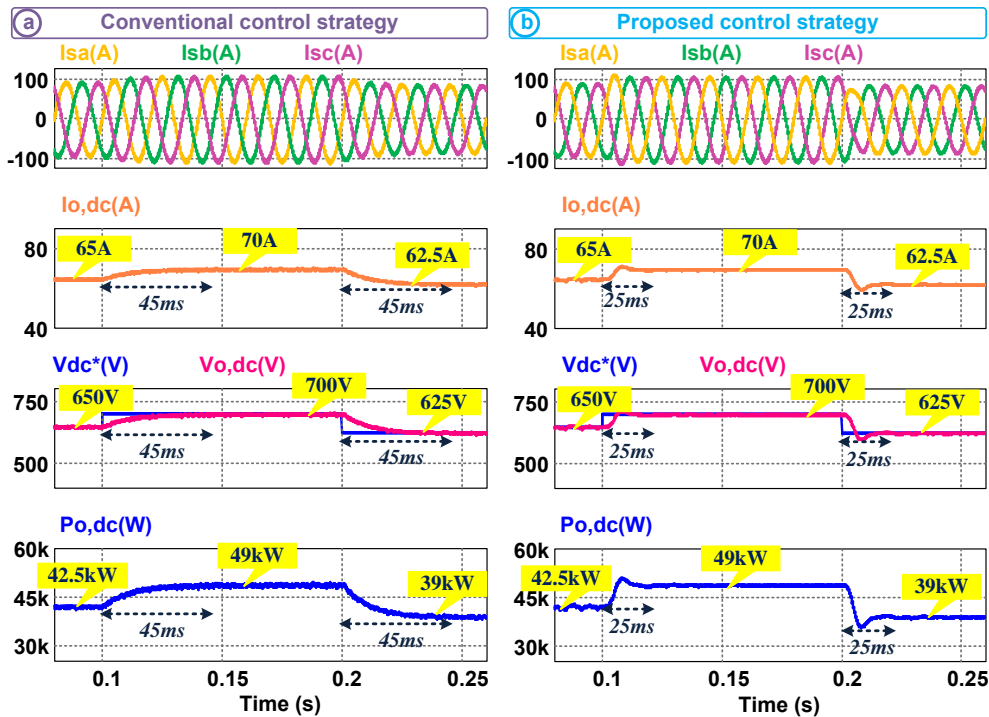


Fig. 8. DC-link voltage step change for a) conventional control and b) proposed control

V. CONCLUSION

In this paper, to achieve fast dynamic response and low current tracking errors, adaptive PI with anti-windup based DPC is proposed. The DPC is employed to generate current signals and adaptive PI with anti-windup control is utilized to acquire the desired DC-link voltage. The above analysis and results have indicated that the proposed control enhances the fast response speed and reduces current tracking errors. The

proposed control can accomplish better tracking of current references and the DC-link voltage step change within nearly 8ms and 25ms, respectively although the conventional control strategy provides 30ms and 45ms for them. Besides, the DC-link voltage has been kept under all cases. The proposed control achieves good robustness against sudden load disturbances compared to the conventional control. The proposed control strategy has carried out the operation of Vienna-type rectifier under various case studies, effectively.

However, although adaptive PI with anti-windup have a broad range of gain alteration; the error can occur with large gain.

The proposed control can be employed in various applications such as wind power generation systems, the charging station of an EV and boost-type unidirectional converter topologies. Besides, the future scope of this study is to improve a control strategy for fast charging EVs.

ACKNOWLEDGMENT

The authors are grateful to Van Yuzuncu Yil University Scientific Research Projects Coordination Unit (Van, Turkey) for financial support of this study (Project number: FYD-2021-9636).

REFERENCES

- [1] H. Zhao, T. Q. Zheng, Y. Li, J. Du, P. Shi, "Control and Analysis of Vienna Rectifier Used as the Generator-Side Converter of PMSG-based Wind Power Generation Systems." *Journal of Power Electronics*, vol. 17, no. 1, 2017, pp. 212-221.
- [2] X. Feng, Y. Sun, X. Cui, W. Ma, Y. Wang, "A compound control strategy of three-phase Vienna rectifier under unbalanced grid voltage." *IET Power Electronics*, vol. 14, no. 16, 2021, pp. 2574-2584.
- [3] S. Bayhan, H. Komurcugil, N. Guler, "An enhanced finite control set model predictive control method with self-balancing capacitor voltages for three-level T-type rectifiers." *IET Power Electronics*, 2022, pp. 1-11.
- [4] Y. Fu, N. Cui, J. Song, Z. Chen, C. Fu, C. Zhang, "A hybrid control strategy based on lagging reactive power compensation for Vienna-type rectifier." *IEEE Transactions on Transportation Electrification*, vol. 7, no. 2, 2021, pp. 825-837.
- [5] W. Ding, C. Zhang, F. Gao, B. Duan, H. Qiu, "A zero-sequence component injection modulation method with compensation for current harmonic mitigation of a Vienna rectifier." *IEEE Transactions on Power Electronics*, vol. 34, no. 1, 2019, pp. 801-814.
- [6] M. Ateş, S. Laribi, "New results on the global asymptotic stability of certain nonlinear RLC circuits." *Turkish Journal of Electrical Engineering & Computer Sciences*, vol. 26, no. 1, 2018, pp. 434-441.
- [7] A. Karafil, H. Özbay, "Power control of single phase active rectifier." *Balkan Journal of Electrical and Computer Engineering*, vol. 7, no. 3, 2019, pp. 332-336.
- [8] J. Adhikari, I. V. Prasanna, S. K. Panda, "Reduction of input current harmonic distortions and balancing of output voltages of the Vienna rectifier under supply voltage disturbances." *IEEE Transactions on Power Electronics*, vol. 32, no. 7, 2017, pp. 5802-5812.
- [9] B. Wang, G. Venkataramanan, A. Bendre, "Unity power factor control for three-phase three-level rectifiers without current sensors." *IEEE Transactions on Industry Applications*, vol. 43, no. 5, 2007, pp. 1341-1348.
- [10] M. İnci, "Performance analysis of T-type inverter based on improved hysteresis current controller." *Balkan Journal of Electrical and Computer Engineering*, vol. 7, no. 2, 2019, pp. 149-155.
- [11] J. Wang, S. Ji, S. Liu, H. Jiang, W. Jiang, "A discontinuous PWM strategy to control neutral point voltage for Vienna rectifier with improved PWM sequence." *IEEE Journal of Emerging and Selected Topics in Power Electronics*, 2021, pp. 1-10.
- [12] Y. Fu, N. Cui, J. Song, Z. Chen, C. Fu, C. Zhang, "A hybrid control strategy based on lagging reactive power compensation for vienna-type rectifier." *IEEE Transactions on Transportation Electrification*, vol. 7, no. 2, 2021, pp. 825-837.
- [13] B. Xu, K. Liu, X. Ran, "Computationally efficient optimal switching sequence model predictive control for three-phase Vienna rectifier under balanced and unbalanced DC links." *IEEE Transactions on Power Electronics*, vol. 36, no. 11, 2021, pp. 12268-12280.
- [14] S. Xie, Y. Sun, M. Su, J. Lin, Q. Guang, "Optimal switching sequence model predictive control for three-phase Vienna rectifiers." *IET Electric Power Applications*, vol. 12, no. 7, 2018, pp. 1006-1013.
- [15] M. Gökdağ, O. Gülbudak, "Model predictive control of an indirect matrix converter with active damping capability." *Balkan Journal of Electrical and Computer Engineering*, vol. 8, no. 1, 2020, pp. 31-39.
- [16] Y. Zhang, C. Qu, "Direct power control of a pulse width modulation rectifier using space vector modulation under unbalanced grid voltages." *IEEE Transactions on Power Electronics*, vol. 30, no. 10, 2015, pp. 5892-5901.
- [17] X. Ran, B. Xu, K. Liu, J. Zhang, "An improved low-complexity model predictive direct power control with reduced power ripples under unbalanced grid conditions." *IEEE Transactions on Power Electronics*, vol. 37, no. 5, 2022, pp. 5224- 5234.
- [18] H. Yang, Y. Zhang, J. Liang, J. Gao, P. D. Walker, N. Zhang, "Sliding-mode observer based voltage-sensorless model predictive power control of PWM rectifier under unbalanced grid conditions." *IEEE Transactions on Industrial Electronics*, vol. 65, no. 7, 2017, pp. 5550-5560.
- [19] L. Hang, M. Zhang, B. Li, L. Huang, S. Liu, "Space vector modulation strategy for VIENNA rectifier and load unbalanced ability." *IET Power Electronics*, vol. 6, no. 7, 2013, pp. 1399-1405.
- [20] Y. Lu, Z. Liu, S. Meng, J. Ji, J. Lyu, "Deadbeat predictive power control for Vienna rectifier under unbalanced power grid condition." *Energy Reports*, vol. 7, 2021, pp. 257-266.
- [21] B. Zhang, C. Zhang, X. Xing, X. Li, Z. Chen, "Novel three-layer discontinuous PWM method for mitigating resonant current and zero-crossing distortion in Vienna rectifier with an LCL Filter." *IEEE Transactions on Power Electronics*, vol. 36, no. 12, 2021, pp. 14478-14490.
- [22] Z. He, H. Ding, Z. Chen, Z. Qi, X. Zhuyu, Z. Dandi, S. Jiannan, "A hybrid DPWM for Vienna rectifiers based on the three-level to two-level conversion." *IEEE Transactions on Industrial Electronics*, 2021, pp. 1-11.
- [23] X. Li, Y. Sun, H. Wang, M. Su, S. Huang, "A hybrid control scheme for three-phase Vienna rectifiers." *IEEE Transactions on Power Electronics*, vol. 33, no. 1, 2018, pp. 629-640.
- [24] B. Xu, X. Ran, "Computationally efficient optimal switching sequence model predictive control for three-phase Vienna rectifier under balanced and unbalanced DC links." *IEEE Transactions on Power Electronics*, vol. 36, no. 4, 2021, pp. 12268-12281.
- [25] Y. Zhang, J. Liu, H. Yang, J. Gao, "Direct power control of pulsewidth modulated rectifiers without dc voltage oscillations under unbalanced grid conditions." *IEEE Transactions on Industrial Electronics*, vol. 65, no. 10, 2018, pp. 7900-7910.
- [26] S. Ouchen, M. Benbouzid, F. Blaabjerg, A. Betka, H. Steinhart, "Direct power control of shunt active power filter using space vector modulation based on supertwisting sliding mode control." *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 9, no. 3, 3243-3253.
- [27] J. W. Choi, S. C. Lee, "Antiwindup strategy for PI-type speed controller." *IEEE Transactions on Industrial Electronics*, vol. 56, no. 6, 2009, pp. 2039-2046.
- [28] H. B. Shin, "New antiwindup PI controller for variable-speed motor drives." *IEEE Transactions on industrial electronics*, vol. 45, no. 3, 1998, pp. 445-450.
- [29] T. L. Chiah, C. L. Hoo, E. C. Y. Chung, "Study of anti-windup pi controllers with different coupling-decoupling tuning gains in motor speed." *Asian Journal of Control*, 2021, pp. 1-10.
- [30] M. Ates, "Circuit theory approach to stability and passivity analysis of nonlinear dynamical systems." *International Journal of Circuit Theory and Applications*, vol. 50, no. 1, 2022, pp. 214-225.
- [31] M. Ates, N. Kadam, "Novel stability and passivity analysis for three types of nonlinear LRC circuits." *An International Journal of Optimization and Control: Theories & Applications (IJOCTA)*, vol. 11, no. 2, 2021, pp. 227-237.
- [32] N. Jin, S. Hu, C. Gan, Z. Ling, "Finite states model predictive control for fault-tolerant operation of a three-phase bidirectional AC/DC converter under unbalanced grid voltages." *IEEE Transactions on Industrial Electronics*, vol. 65, no. 1, 2017, pp. 819-829.
- [33] J. Svensson, M. Bongiorno, A. Sannino, "Practical implementation of delayed signal cancellation method for phase-sequence separation." *IEEE Transactions on Power Delivery*, vol. 22, no. 1, 2006, pp. 18-26.
- [34] Y. Zhang, C. Qu, "Model predictive direct power control of PWM rectifiers under unbalanced network conditions." *IEEE Transactions on Industrial Electronics*, vol. 62, no. 7, 2015, pp. 4011-4022.
- [35] Y. Zhang, J. Jiao, J. Liu, "Direct power control of PWM rectifiers with online inductance identification under unbalanced and distorted network conditions." *IEEE Transactions on Power Electronics*, vol. 34, no. 12, 2019, pp. 12524-12537.

- [36] H. Zhang, C. Zhang, X. Xing, C. Liu, X. Li, B. Zhang, "Three-layer double vectors model predictive control strategy for current harmonic reduction and neutral point voltage balance in Vienna rectifier." IEEE Transactions on Transportation Electrification, 2022, pp. 1-12.

BIOGRAPHIES



DOĞAN ÇELİK, received the B.Sc. degree in Electrical and Electronics Engineering, Cukurova University, Adana, Turkey in 2013. He completed M.Sc. program at Department of Electrical and Electronics Engineering, Yuzuncu Yil University, Van, Turkey in 2016. He received the Ph.D. degree from Van Yuzuncu Yil University, Van, Turkey in 2020. His research

interest areas are microgrid applications, distributed generation, distributed generation/integration of renewable energy systems, power quality, active power filters, electric vehicles, power system control.

A Weakly Supervised Clustering Method for Cancer Subgroup Identification

Duygu Ozcelik and Oznur Tastan

Abstract—Identifying subgroups of cancer patients is important as it opens up possibilities for targeted therapeutics. A widely applied approach is to group patients with unsupervised clustering techniques based on molecular data of tumor samples. The patient clusters are found to be of interest if they can be associated with a clinical outcome variable such as the survival of patients. However, these clinical variables of interest do not participate in the clustering decisions. We propose an approach, WSURFC (Weakly Supervised Random Forest Clustering), where the clustering process is weakly supervised with a clinical variable of interest. The supervision step is handled by learning a similarity metric with features that are selected to predict this clinical variable. More specifically, WSURFC involves a random forest classifier-training step to predict the clinical variable, in this case, the survival class. Subsequently, the internal nodes are used to derive a random forest similarity metric among the pairs of samples. In this way, the clustering step utilizes the nonlinear subspace of the original features learned in the classification step. We first demonstrate WSURFC on handwritten digit datasets, where WSURFC can capture salient structural similarities of digit pairs. Next, we apply WSURFC to find breast cancer subtypes using mRNA, protein, and microRNA expressions as features. Our results on breast cancer show that WSURFC could identify interesting patient subgroups more effectively than the widely adopted methods.

Index Terms—Clustering, cancer subtype identification, patient subgroup identification.

I. INTRODUCTION


A major hurdle in devising more effective cancer therapies is the accurate stratification of patients into subgroups [1]. This stems from the fact that most cancer types are heterogeneous at the molecular level; seemingly similar tumors that are classified into the same cancer type may have distinct molecular profiles resulting in distinct clinical trajectories [2]. The availability of large sets of patient molecular data has opened up opportunities to redefine the subtypes of cancers [3].

The widely adopted approach for grouping cancer patients using patient molecular data is to apply unsupervised clustering techniques such as k-means, hierarchical clustering or non-negative matrix factorization (NMF) on the genomic data

Duygu Ozcelik conducted this research at the Department of Computer Engineering, Bilkent University, Ankara, TURKEY. e-mail: duyguozcelik89@gmail.com

 <https://orcid.org/0000-0001-8980-6200>

Oznur Tastan is with the Faculty of Engineering and Natural Sciences, Sabanci University, Istanbul, TURKEY. Coresponding author e-mail: otastan@sabanciuniv.edu

 <https://orcid.org/0000-0001-7058-5372>

of patients, sometimes combined with consensus clustering [4]–[11]. There are also more advanced methods that use probabilistic modeling of multi-omics data to project the data to a lower dimension (reviewed and benchmarked in [12] and [13]) such as PARADIGM [14], iCluster [15], multi-omics perturbation based approaches such as PINSplus [16] or methods that integrates biological pathways and multi-omics data, such as PAMOGK [17]. All of these approaches take an unsupervised approach, and the patient clusters are deemed interesting if they are found to be associated with a clinical variable of interest, such as the survival rate of the patients. Therefore, the clinical variable of interest does not participate in the clustering decisions to guide the clusters.

One body of method that incorporates clinical variable of interest makes use of the association of the features with the clinical variable of interest in the feature selection step. Bair et al. [18] test the null hypothesis of no association between the feature and the outcome variable and applies feature selection based on the test-statistic. Next, with the remaining feature, they perform clustering using a conventional clustering algorithm such as k-means or hierarchical clustering. Koestler et al. [19] propose a method called “semi-supervised recursively partitioned mixture models” with the same rationale. It also calculates a score for each feature and measures the association between the features and the outcome variable of interest. Next, clustering is carried out using only the features with the largest scores. The difference between Bair et al. and Koestler et al. is that the latter applies a recursively partitioned mixture models algorithm [20] instead of a standard clustering algorithm. Although these approaches are simple, they are limited in the sense that they only take account of the univariate relationship of features with the clinical variable.

We utilize weakly supervised learning and propose a new approach, WSURFC, where the clinical variable of interest guides the clustering process. The weak supervision is achieved by learning a similarity metric by predicting a clinical variable, in our study, the survival class as long and short survivors. With guidance from the clinical label, WSURFC learns a non-linear similarity metric among patients, which is then used to cluster the patients. We first demonstrate the methodology on an unrelated, but easier-to-expect problem, digit classification. Then we apply it to breast cancer subgroup identification.

II. METHODS

This section presents a detailed description of our proposed method and the datasets used. Let \mathcal{D} represent the set of n

Algorithm 1 WSURFC: Weakly Supervised Random Forest Clustering

Input: \mathcal{D} the patient set, n number of patients, p , number of features, $\mathbf{X} \in \mathbb{R}^{n \times p}$ feature matrix, \mathbf{y} class labels associated with the clinical variable of interest for the patient set, \mathcal{F} the random forest classifier, d_l and d_u : the lower bound and upper bounds of the range where the depth will be sampled, respectively, \mathbf{r} Random Forest parameters, \mathbf{c} clustering parameters.

Output: Partitioning of C_i 's where $\mathcal{D} = \cup_{i=1}^k C_i$, k is the number of clusters and $C_i \cap C_j = \emptyset$ for all $i, j \in 1, \dots, k$.

1. $\mathcal{F} \leftarrow \text{RFClassifier}(\mathbf{X}, \mathbf{y}, \mathbf{r})$
2. $\mathbf{S} \leftarrow \text{RFSimilarity}(\mathcal{F}, d_l, d_u)$
2. Convert \mathbf{S} to a distance matrix, \mathbf{D}
3. $C \leftarrow \text{Cluster}(\mathbf{D}, \mathbf{c})$
4. return C

patients. We denote the feature vector derived from patient molecular data with $\mathbf{x}^{(i)} \in \mathbb{R}^p$ and the clinical variable of interest for patient i is denoted with $y^{(i)}$. In this work, we assume that the clinical variable of interest is dichotomized; however, it can be generalized to continuous outputs by using a random forest regression or random forest survival model instead of a classifier. We will represent the $n \times p$ feature matrix with \mathbf{X} and \mathbf{y} is the $n \times 1$ label vector. We want to find a partitioning \mathcal{C} such that: \mathcal{D} is grouped into a number of disjoint subsets C_i 's where, $\mathcal{D} = \cup_{i=1}^k C_i$ and where the clustering is guided by \mathbf{y} .

A. WSURFC: Weakly Supervised Random Forest Clustering

- **Step 1:** Given the feature vector of patient samples, \mathbf{X} and the clinical variable of interest, \mathbf{y} , train a random forest classifier to classify the clinical variable. We denote the random forest model with \mathcal{F} .
- **Step 2:** Calculate the patient similarities based on co-occurrence in the feature subspaces formed by the trees in \mathcal{F} . To calculate the similarity of patients i and j , draw a random depth, d , within a predefined depth range. Sort down i and j in the forest of trees and check whether i and j fall onto the same internal node at this randomly drawn depth d . Calculate the pairwise patient similarity based on the fraction of times the patients fall on the same internal node. Form the patient similarity matrix S and convert the similarity matrix to a distance matrix.
- **Step 3:** Use this distance matrix for clustering patients.

The detailed procedure is summarized in Algorithm 1 in detail and demonstrated in Figure 1.

1) *Step 1: Random Forest Classification:* WSURFC uses a random forest classifier to predict the clinical variable to obtain the feature subspaces. Random forest is an ensemble method that learns many decision trees and aggregates their results [21]. Each tree is independently trained using a bootstrap sample of the training examples. In growing a tree, at each split, m features are randomly selected from the global set of features, and the one that maximizes an impurity criterion, in this case, the Gini index, is selected. The input samples

Algorithm 2 Random Forest Subspace Similarity (RFSim)

Input: D set of n samples, B number of trees in the random forest, \mathcal{F} random forest model, Z_b b -th bootstrap sample by which tree T_b is trained with, $z_{i,j}$ number of bootstrap samples where i and j are in the bag.

Output: \mathbf{S} : $n \times n$ similarity matrix

1. For each i, j pair in D
 - i. For all bootstrap samples Z_b where i, j are both in Z_b
 - (a) Get tree T_b of B
 - (b) Get height h_b of T_b
 - (c) Sample d from $[h_b \times d_s, h_b \times d_e]$ uniformly at random
 - (d) Traverse i on T_b until depth d is reached and find the internal node p_i on which i falls at that depth
 - (e) Traverse j on T_b until depth d is reached and find the internal node p_j on which j falls
 - (f) **if** $p_i == p_j$ **then**
 $\mathbf{S}[i, j] \leftarrow \mathbf{S}[i, j] + 1$
 - ii. $\mathbf{S}[i, j] \leftarrow \frac{\mathbf{S}[i, j]}{z_{i, j}}$
2. return \mathbf{S}

that arrive at a particular node are further split based on the value of the selected best feature. A path from the root to a node includes a subset of features in a tree. These subsets of features and their combination forms a feature subspace as depicted in Figures 1. WSURFC assesses the similarities of the input samples under these formed feature subspaces based on whether they fall in the same subspace or not.

2) *Step 2: Calculation of Random Forest Subspace Similarity:* Using the random forest ensemble, we calculate a similarity metric, which we refer to as the *RFSim*. By sorting down the example pairs onto random depths in the tree and checking how often they arrive at the same internal node in the tree, we construct a similarity matrix of the patients. Different random depths provide different views of the samples under different feature combinations. Checking whether a pair of patients would fall on the same internal node translates into checking if they are in the same subspace created by a nonlinear combination of a subset of features. For the pairs that end up at the same node, their co-occurrence count is incremented by 1. This is repeated for all the trees, and the similarities are finally normalized with the number of bootstrap samples where both samples are in the bag. The steps of this calculation are presented in Algorithm 2.

The similarity metric is similar to random forest proximity [21] with a key difference. Random forest proximity is calculated based on how often the example pairs fall onto the same leaf node; thus, the feature combinations that predict the labels are used. On the other hand, we calculate the similarity based on internal nodes; these feature representations may not be strong enough to predict the label. However, they can still be descriptive enough to reveal similarities of the examples, as we illustrate in our experiments. We observe that a depth chosen from the mid-level of the tree is useful, as discussed

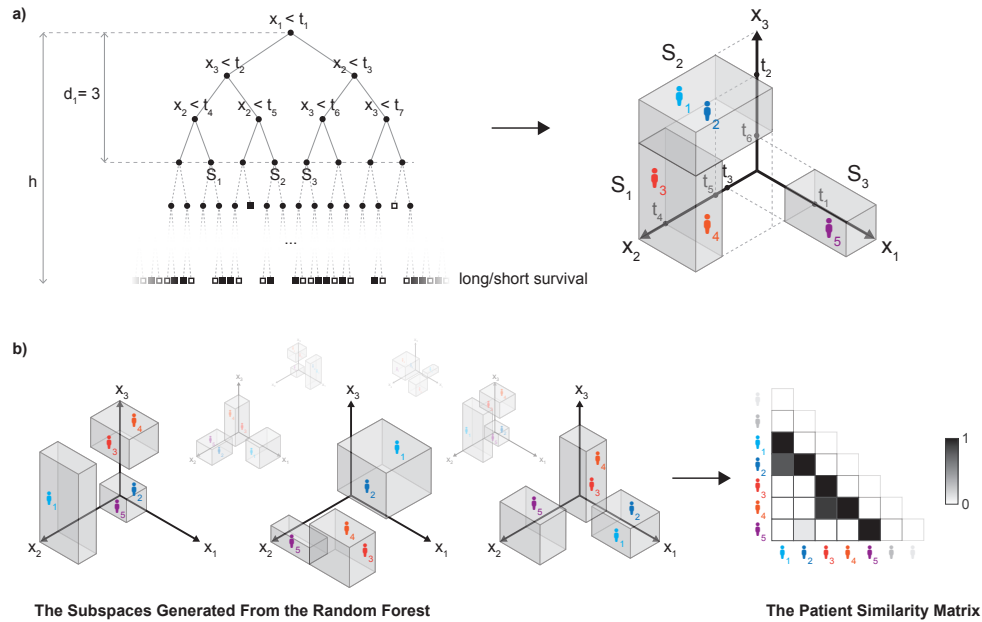


Fig. 1: Illustration of WSURFC algorithm. a) A tree in the random forest is trained to classify patients into long survivor or short survivor classes. t_1, t_2, t_3, t_4, t_5 are features of the patients. h is the height of the tree. At the random depth $d_1 = 3$, patients 3 and 4 shown with orange color fall into same node S_1 . Patients 1 and 2, shown in blue, fall on node S_2 , and patient 5 falls on node S_3 . At this level, three subspaces S_1, S_2 and S_3 arise. These subspaces are used to calculate the similarity of patient pairs. b) Different depth values to generate different partitions of subspaces are selected uniformly at random. The similarity of the two patients is calculated based on how many times they fall in the same subspace for the trees in which these two patients are both in the bootstrap samples. Patients 3 and 4 are in the same subspace three times, while patients 2 and 5 are in the same subspace for once. Therefore, the similarity value of patients 3 and 4 is closer to 1, and it is indicated with a darker color.

in the Results section.

3) *Step 3: The Final Clustering*: In the last step of the algorithm, we convert the similarity matrix to a distance matrix by subtracting the values from 1 and input this distance matrix to the clustering algorithm. For clustering, we use the hierarchical clustering algorithm with average linkage, but other clustering algorithms can be used at this stage.

B. Datasets

We first use MNIST handwritten digit dataset [22] to demonstrate WSURFC. The dataset contains images of 28×28 handwritten pixel digits. We sample 5000 examples from the entire set. Next, we apply our algorithm to breast cancer molecular data. We use data that contain solid primary tumors mRNA, microRNA, and protein expression levels that are made available through the Cancer Genome Atlas (TCGA) project [23]. The data is retrieved through the UCSC (University of California, Santa Cruz) Cancer Genomics Browser. The mRNA expression data contain 1196 patients and 20531 transcripts. miRNA expression data comprise information from 1194 patients and 1046 features. Protein expression data hold information on 747 patients and 131 proteins.

III. RESULTS

To demonstrate the method's effectiveness, we perform experiments on a problem that is easier to inspect, the well-known MNIST handwritten digit dataset. Subsequently, results

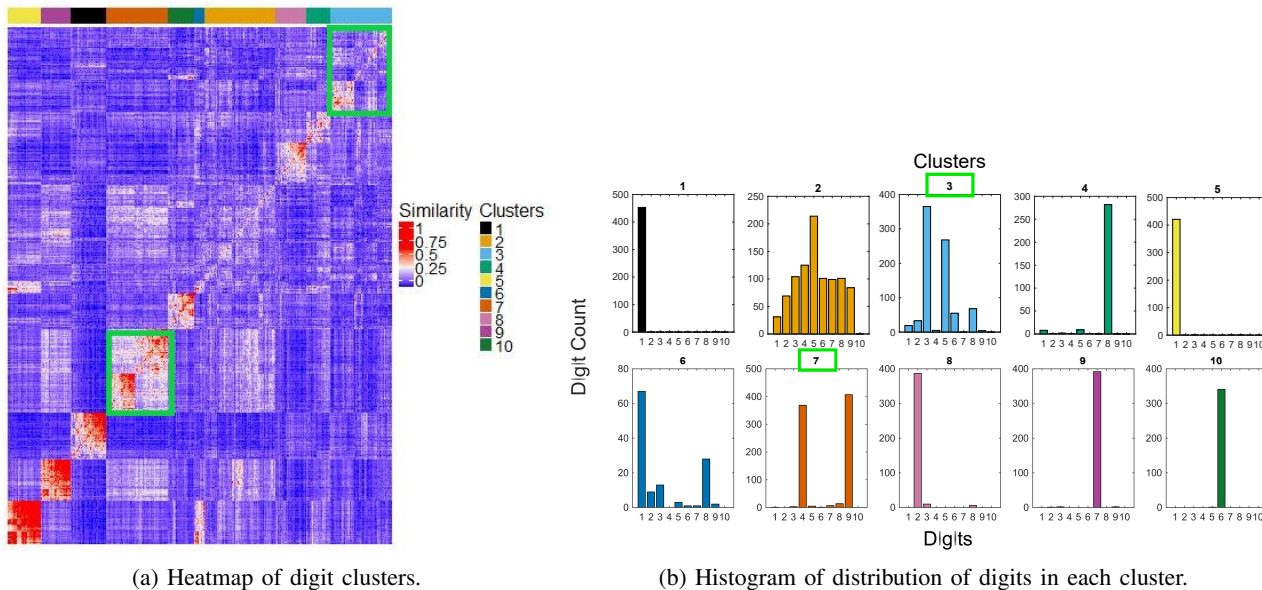
of applying WSURFC on the breast cancer patient dataset are presented.

A. Results on MNIST Digit Dataset

The random forest classifier is generated with 200 trees and trained with digit labels as class labels. In constructing the similarity matrix, we sample from $[\frac{h}{3}, \frac{2h}{3}]$ interval depth. The digits are then clustered into 10 clusters by inputting the corresponding similarity matrix to the hierarchical clustering algorithm.

Figure 2 shows a heatmap of the digit similarities and the clusters. The histogram shows the digit label distribution in each cluster. The aim of clustering here is not to arrive at clusters that represent purely single digits. Because we use subspaces not close to the leaves; however, we hope to reveal structurally similar digits. WSURFC finds these similarities that are not input. For example, Cluster 7 contains mostly digit pairs 4 and 9. Both of these digits have very similar structures. Similarly, cluster 3 reveals that 3 and 5 are similar, and additionally, 8 exhibits similarities to this digit pair. These results indicate that although the classifier is trained with 10 digit labels, WS-RFClust can uncover structural similarities between individual digits. Supplementary Figure 1 shows the silhouette width of the examples of these clusterings.

1) *The effect of Sampling from Interval Nodes at Different Depths*: In applying WSURFC, the depth levels in each tree



(a) Heatmap of digit clusters.

(b) Histogram of distribution of digits in each cluster.

Fig. 2: Clustering results of handwritten digit dataset. a) Colors on the heatmap represent similarities computed for sample pairs (reds indicate high similarity, blue indicates low). The bars on top indicate different clustering. b) Each subplot that bears the same color on the histogram displays the digit content of the clusters based on their true class labels. The x-axis of a histogram represents digits, and the y-axis represents the number of observed samples in each digit. The two interesting clusters, 3 and 7, are marked with green boxes.

are sampled randomly but within a predefined depth range. To understand the effect of the sampling depth, we experiment with different depth intervals on the digit dataset: let h be the height of a tree in the forest. We experiment with selecting d from the interval lower part of the tree that is from $(0, \frac{h}{3}]$, the middle part from $[\frac{h}{3}, \frac{2h}{3}]$ and third interval is from $[\frac{2h}{3}, h]$. To speed up calculations, for training random forest classifiers, we sample 1500 digits in these experiments.

Supplementary Figure 2a displays the results where the intervals are selected from the interval nodes closer to the root. In cluster 8, we observe that the digits 4 and 9 are in the same cluster; these are digit pairs with very similar shapes. In cluster 2, in Supplementary Figure 2b, where the intervals are sampled in the medium part of the trees, the grouping of digits 4 and 9 is clearer (cluster 9). Similarly, cluster 2 reveals that 3 and 5 are similar and that digit 8 is similar to these digits.

B. Application of WSURFC to Breast Cancer

We apply WSURFC to breast cancer patient data with three different input types, mRNA, miRNA, and protein expressions (RPPA). In each case, we dichotomize the survival time of patients into two classes. Patients with survival time shorter than the 25% quartile are labeled as low survivors, whereas patients with survival times longer than the 75% quartile are labeled as high survivors. These labels constitute the class labels for the classification step. The number of patients is different in each case as the number of available patient samples for each data type is different. We apply different feature selection criteria including, ttest, ROC, Entropy, Chernoff, and Wilcoxon statistical tests to reduce the number of features. We experiment with different feature set size values and select the

one that produces the best 5-fold cross-validation accuracy. We apply 5-fold cross-validation to the training data 10 times and form decisions based on the average accuracy over 10 runs. We apply WSURFC by sampling uniformly at random from depths from the interval $[\frac{h}{3}, \frac{2h}{3}]$, where h is the height of the trees. We also run the widely adapted method NMF-Consensus clustering on each of these datasets for comparison.

To inspect clusters in each case, we use other available clinical data in breast cancer. We compare clusters in terms of survival, tumor stage distributions, and PAM50 subtypes. We used the Kaplan-Meier curves and the log-rank test to check if clusters' separations in terms of survival distributions. χ^2 test of independence is used to test the association of the clusters with the tumor stages and PAM50 subtypes.

1) *WSURFC with mRNA Expression Data*: There are 1196 patient samples with mRNA expression data. The number of long survivor patients is 299, while the number of short survivors is 300. We apply WSURFC on these 1196 patients. Let k denote the number of clusters, we try clustering with $k = 2, 3, 4, 5, 6$. Finally, we use the trained model to cluster all the samples. The best clustering resulted in $k = 5$ clusters. Figure 3a shows the Kaplan-Meier survival curves for each of the clusters; we apply the log-rank test on the survival distributions of the clusters. For $k = 5$, the p-value of the test is $4.5e-05$, indicating that survival distributions of clusters are distinct at significance level 0.05. The silhouette width plot for this clustering is shown in Figure 3b. Supplementary Figures 3 and 4 show silhouette width and survival plots for all k values, respectively.

Figure 4a demonstrates the Kaplan-Meier survival plots of the clusters when consensus NMF is applied to the mRNA data. Smallest p-value is achieved when $k = 6$ is p -value=

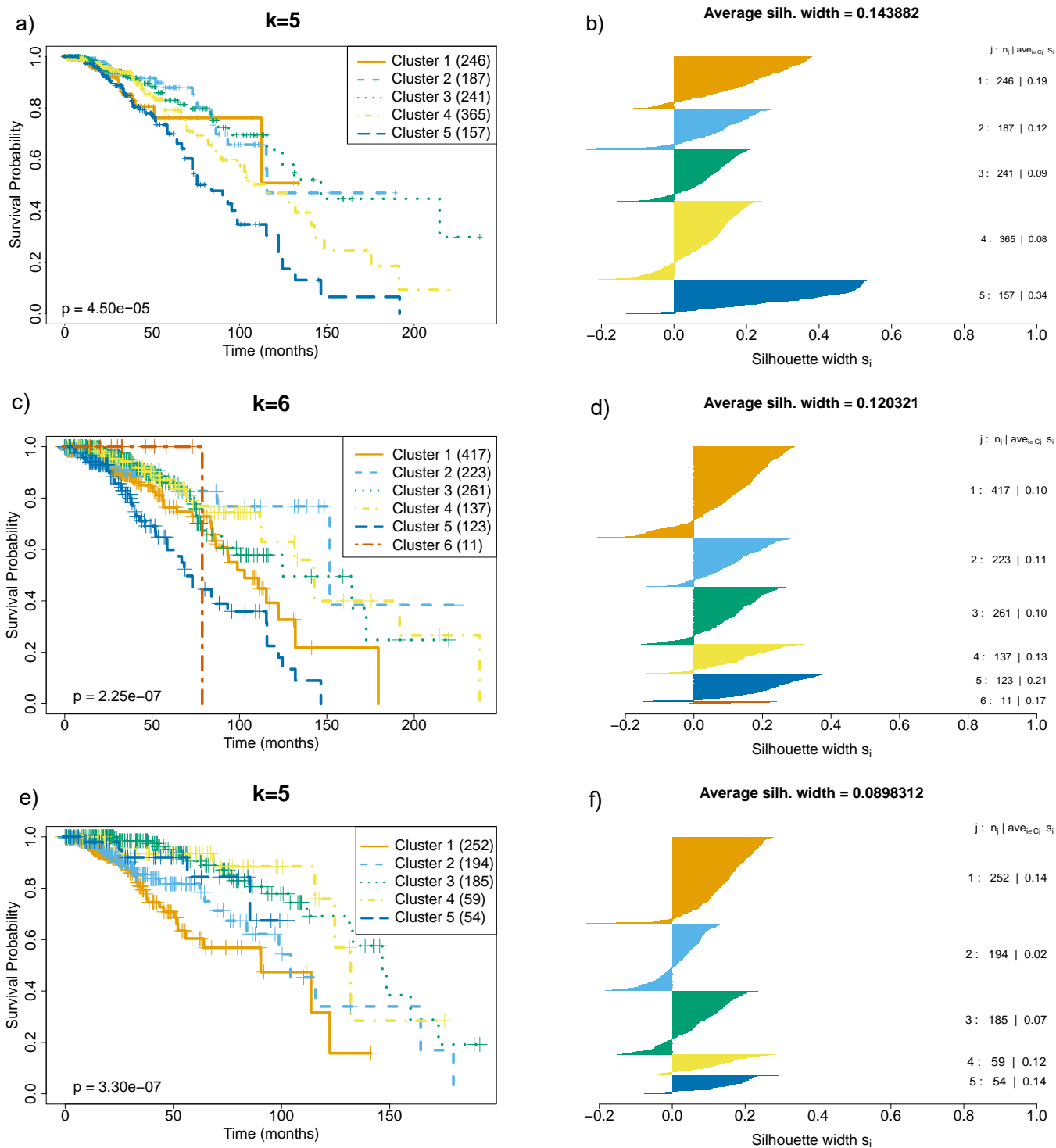


Fig. 3: In survival plots of each data type (a, c, and e), the x-axis shows the survival time in months, and the y-axis shows survival probability at a given time. In the silhouette graph of each data type (b,d,f), the x-axis is the ruler that shows the width of each cluster. j is cluster id, n_j is number of patients in cluster C_j and S_i is the silhouette width of C_j . The y-axis shows $j : n_j | \text{ave}_{i \in C_j} S_i$ for each cluster. Average silhouette width is the overall average of all clusters. a) Survival plot for mRNA clusters. b) Silhouette width plot for mRNA clusters. c) Survival plot of microRNA clusters. d) Silhouette width plot for miRNA clusters. e) Survival plot of RPPA clusters. f) Silhouette width plot for RPPA clusters.

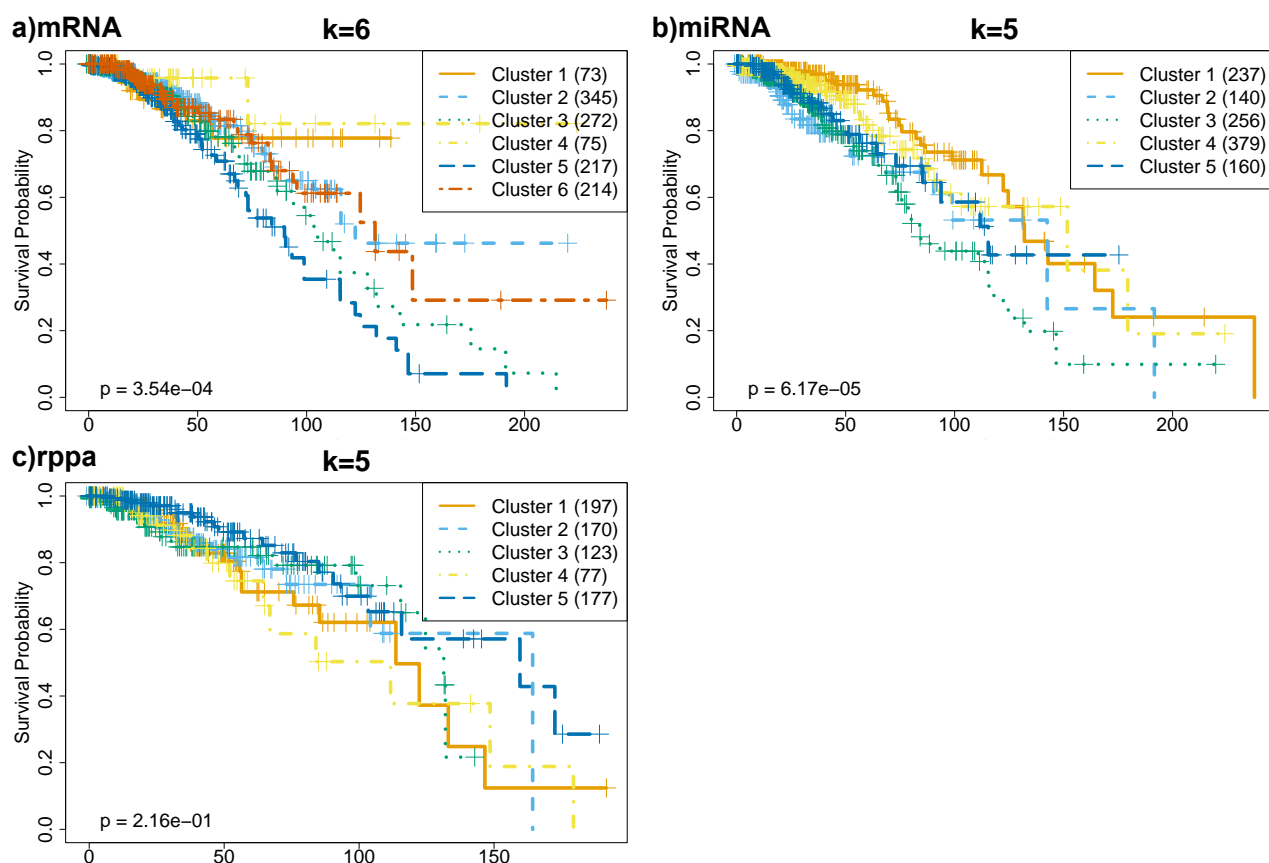


Fig. 4: a) Survival plot of consensus NMF method on mRNA data. Smallest p-value is achieved at $k = 6$. b) Survival plot of consensus NMF method on microRNA data. Smallest p-value is achieved at $k = 5$. c) Survival plot of consensus NMF method on RPPA data. Smallest p-value is achieved at $k = 5$

$3e-04$. This is 100 x fold larger than the p-value we obtain in WSURFC and indicates that better clusters are obtained with WSURFC. We test the null hypothesis that there is no association between the tumor stages and the clusters at the significance level 0.05 and reject the null hypothesis with $p = 0.03$. In this test, we exclude stages Stage IB, Stage II, Stage III, Stage Tis, Stage X, Stage IV; because only a few patients belong to these stages. We next check whether the identified clusters are related to the PAM50 subtypes. We conclude WSURFC clusters have a strong association with the intrinsic molecular subtypes. The corresponding test is $p < 2.2e-16$.

WSURFC with miRNA Expression Data: microRNA expression matrix contains 1172 available patients and 1046 miRNA expression level features. We follow the same steps in training the mRNA expression data. We select only the low and high survivor 587 patients from miRNA expression data and select 200 features with the t-test. We divide these patients into 476 training and 116 test examples. Test examples are classified with a random forest. Then, all the patients (1172) that are available in the dataset are input to train the model, and WSURFC constructs a similarity matrix of patients. We apply hierarchical clustering for $k = 2, 3, 4, 5, 6$. Best clustering is achieved with $k = 6$. Figure 3b shows the Kaplan-Meier plot for $k = 6$, which yields a very low p-value of $2.25e-07$

and the silhouette width plot for this clustering is shown in Figure 3d. Supplementary Figures 5 and 6 show silhouette width and survival plots for all k values, respectively.

We applied Consensus NMF to the microRNA dataset to compare the clustering performance of WSURFC with Consensus NMF. We run the consensus NMF algorithm dataset with 1172 samples containing all the patients. Figure 4 b) demonstrates Kaplan-Meier survival plot for the best clustering, $k = 5$ with p-value ($p = 6.17e-05$). This value is larger than p-value which we obtain in WSURFC as $p = 2.25e-07$ in $k=5$. Therefore, we conclude that WSURFC provides a better separation of clusters.

We next checked the association of the clusters found with the tumor stages, the χ^2 test results with $p = 0.002 < 0.05$, therefore we reject the null hypothesis in favor of the alternative hypothesis, which states that the tumor stages are associated with WSURFC subtypes in miRNA dataset. Note that we excluded stages Stage IB, Stage II, Stage III, Stage IIIB, Stage Tis, Stage X, Stage IV due to small numbers of patients belonging to those stages. Finally, we tabulate the data into PAM50 cluster ids and WSURFC subtypes and apply the χ^2 test of independence for the clustering results with $k = 6$. The resulting $p < 2.2e-16$ of test is considerably smaller than the significance level 0.05.

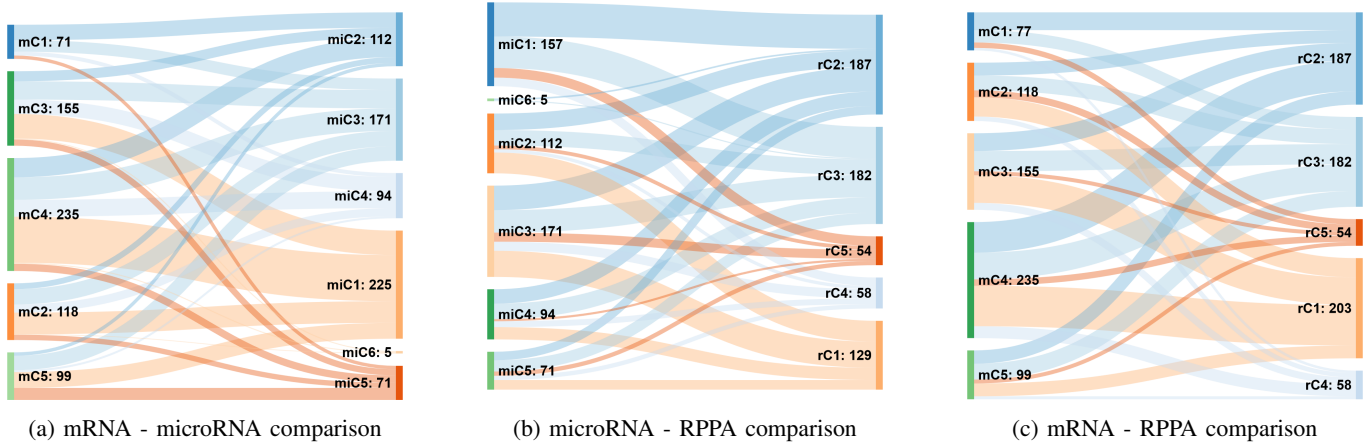


Fig. 5: Overlap between clusters obtained with different molecular data as inputs. The bars in the left and right columns are results of two different clusterings; the flow between a pair of clusters indicates the overlapping patients in those clusters. The diagrams are created using the SankeyMATIC tool.

WSURFC with RPPA Expression Data: As the final input, we experiment with protein expression. Protein expression data is collected on 744 available patients and 131 features. There are only 131 features in the RPPA dataset; therefore, we use all the features without any feature selection. The protein expression data contains 373 low and high survivor patients. We utilize 299 of them as the training set and 74 of them as the test set. We apply hierarchical clustering for $k = 2, 3, 4, 5, 6$. The resulting clusters are compared with respect to survival rate, tumor stage, and PAM50 subtypes.

Supplementary Figure 7 indicates silhouette width graph of clustered patients in the protein expression dataset. Supplementary Figures 7 and 8 show silhouette width and survival plots for all k values, respectively. The best clustering is achieved when $k = 5$ with a p -value of $3.30e-07$ when $k=5$. Supplementary Figures 7 and 8 show silhouette width and survival plots for all k values, respectively.

We apply Consensus NMF to the protein expression dataset to compare the clustering performance of WSURFC with Consensus NMF. We run the consensus NMF algorithm dataset with 744 samples containing all the patients. Figure 4c) demonstrates Kaplan-Meier survival plots for each k value when consensus NMF is applied. p -value is $p = 0.055$ when $k=5$, overall p -value range is between $0.01 - 0.2$. Consensus NMF results are not confidently below $\alpha = 0.05$. Therefore we conclude that Consensus NMF clusters are not significantly different in terms of survival rate. WSURFC outperforms Consensus NMF in terms of survival rate differentiation between subgroups.

We exclude stages Stage IB, Stage II, Stage III, Stage IIIB, Stage Tis, Stage X, Stage IV; because there are only a few patients who belong in these stages. The χ^2 test of independence yields to p -value= 0.02; therefore, we reject the null hypothesis at a significance level of 0.05. We conclude that tumor stages are randomly distributed in the WSURFC subtypes. To compare PAM50 subtypes and WSURFC subtypes, we apply the χ^2 test of independence for $k = 5$. $p < 2.2e-16$ of the test is considerably smaller than 0.05.

Therefore WS-RClust clusters have a strong correlation with the PAM50 molecular subtypes.

C. Overlap among clusters

When used as input, each molecular input data yields a different clustering of patients. We analyzed the overlap among these clustering results. Figure 5 shows the pairwise overlap between the clustering results. We conducted this analysis with the patients whose profiles contain data pertinent to these molecular types. Some of the clusters have a large overlap, such as C4 which is obtained with mRNA and C1 of miRNA data, as indicated by the large flow size (Figure 5a). Similarly, the miRNA C1 cluster is mainly composed of patients that belong to the RPPA C2 and C3 clusters (Figure 5b). mRNA C4 cluster with RPPA C1 cluster and mRNA C3 with RPPA C1 cluster share a considerable number of patients (Figure 5c). These results also indicate that different molecular data types can bring different views into clusterings, and a multi-view approach could be useful to integrate the clusters. We leave this direction as a future work.

IV. CONCLUSION

Inaccurate grouping of patients hinders the development of effective targeted therapies. Identifying patient subgroups with similar molecular profiles can reveal the unique molecular characteristics that shape them and open up possibilities for targeted therapeutics. Traditionally, unsupervised clustering analysis is applied to the genomic data of tumor samples, and the patient clusters are considered interesting if they can be associated with a clinical outcome variable such as the survival rate of patients [5], [12], [24], [25]. We propose a weakly supervised clustering framework (WSURFC) in place of this unsupervised framework. In this approach, the clustering partitions are weakly guided with the clinical outcome of interest. We achieve this by using the similarity of patients under subsets of features created in a random forest ensemble which is trained with a label of interest.

We apply WSURFC to handwritten digit datasets to understand the effect of several parameters. To understand how the sampling from different levels of the tree would affect clustering, we vary the interval range from which we sample random depths. We observe that if the depths are close to the tree height, the resulting partitions are found to be close to the leaves, and therefore, these clusters correspond to the classes. If we choose depths near the root, the structure information is lost. Thus, we conclude that sampling from a medium-range is critically important to attain the best trade-off between predictive accuracy and speed.

A widely adopted technique in cancer is using the Consensus NMF clustering approach. We apply WSURFC to TCGA breast cancer miRNA, mRNA, and protein expression datasets separately to identify breast cancer subtypes. We also run the Consensus NMF approach on the same datasets to see if we can capture better subgroupings of patients. We vary the number of clusters and analyze these clusters in terms of internal cluster validity metrics, such as silhouette width, and external clinical data, such as tumor stage and PAM50 classification. When the data are clustered into 5 or 6 subgroups, the resulting survival rates of subgroups significantly differ from each other and are better in terms of survival separation compared to the consensus NMF approach. Regardless of the input expression data type, the method performs well, and the resulting clusters are found to be associated with the tumor stages and PAM50 subtypes. Although the different clusterings obtained with different molecular data types have large overlap, they are different. A multi-view clustering approach, where each view is obtained with a molecular data type, could be interesting.

In this study, we have limited our analysis to breast cancer, but the approach presented herein can be applied to any cancer type and any clinical variable of interest. This work assumes that the target variable y is discrete; however, the approach can easily be extended to the cases where y is a continuous variable, replacing the random forest classifier with a regressor. Alternatively, the target variable can be cast as a survival variable, and a random survival forest can be adapted.

ACKNOWLEDGMENT

The results shown here are in part based upon data generated by the TCGA Research Network: <https://www.cancer.gov/tcga>.

REFERENCES

- [1] L. Hood and S. H. Friend, "Predictive, personalized, preventive, participatory (p4) cancer medicine," *Nature reviews Clinical oncology*, vol. 8, no. 3, p. 184, 2011.
- [2] I. Dago-Jack and A. T. Shaw, "Tumour heterogeneity and resistance to cancer therapies," *Nature reviews Clinical oncology*, vol. 15, no. 2, pp. 81–94, 2018.
- [3] D. Koboldt, R. Fulton, M. McLellan, H. Schmidt, J. Kalicki-Verizer, J. McMichael, L. Fulton, D. Dooling, L. Ding, E. Mardis *et al.*, "Comprehensive molecular portraits of human breast tumours," *Nature*, vol. 490, no. 7418, pp. 61–70, 2012.
- [4] P. S. B. Joel S. Parker, "Supervised risk predictor of breast cancer based on intrinsic subtypes," *Journal of Clinical Oncology*, vol. 27, no. 8, p. 1160–1167, 2009.
- [5] R. G. Verhaak, K. A. Hoadley, E. Purdom, V. Wang, Y. Qi, M. D. Wilkerson, C. R. Miller, L. Ding, T. Golub, J. P. Mesirov *et al.*, "Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in *pdgfra*, *idh1*, *egfr*, and *nf1*," *Cancer cell*, vol. 17, no. 1, pp. 98–110, 2010.

- [6] The Cancer Genome Atlas Network, "Comprehensive molecular portraits of human breast tumours," *Nature*, vol. 490, pp. 61–70, 2012.
- [7] A. Ally, M. Balasundaram, R. Carlsen, E. Chuah, A. Clarke, N. Dhalla, R. A. Holt, S. J. Jones, D. Lee, Y. Ma *et al.*, "Comprehensive and integrative genomic characterization of hepatocellular carcinoma," *Cell*, vol. 169, no. 7, pp. 1327–1341, 2017.
- [8] The Cancer Genome Atlas Network, "Integrated genomic analyses of ovarian carcinoma," *Nature*, vol. 474, pp. 609–615, 2011.
- [9] K. A. Hoadley, C. Yau, D. M. Wolf, A. D. Cherniack, D. Tamborero, S. Ng, M. D. Leiserson, B. Niu, M. D. McLellan, V. Uzunangelov *et al.*, "Multiplatform analysis of 12 cancer types reveals molecular classification within and across tissues of origin," *Cell*, vol. 158, no. 4, pp. 929–944, 2014.
- [10] F. Chen, D. S. Chandrashekar, S. Varambally, and C. J. Creighton, "Pan-cancer molecular subtypes revealed by mass-spectrometry-based proteomic characterization of more than 500 human cancers," *Nature communications*, vol. 10, no. 1, pp. 1–15, 2019.
- [11] M. Ceccarelli, F. P. Barthel, T. M. Malta, T. S. Sabedot, S. R. Salama, B. A. Murray, O. Morozova, Y. Newton, A. Radenbaugh, S. M. Pagnotta *et al.*, "Molecular profiling reveals biologically discrete subsets and pathways of progression in diffuse glioma," *Cell*, vol. 164, no. 3, pp. 550–563, 2016.
- [12] N. Rappoport and R. Shamir, "Multi-omic and multi-view clustering algorithms: review and cancer benchmark," *Nucleic acids research*, vol. 46, no. 20, pp. 10546–10562, 2018.
- [13] L. Cantini, P. Zakeri, C. Hernandez, A. Naldi, D. Thieffry, E. Remy, and A. Baudot, "Benchmarking joint multi-omics dimensionality reduction approaches for the study of cancer," *Nature communications*, vol. 12, no. 1, pp. 1–12, 2021.
- [14] C. J. Vaske, S. C. Benz, J. Z. Sanborn, D. Earl, C. Szeto, J. Zhu, D. Haussler, and J. M. Stuart, "Inference of patient-specific pathway activities from multi-dimensional cancer genomics data using paradigm," *Bioinformatics*, vol. 26, no. 12, pp. i237–i245, 2010.
- [15] R. Shen, A. B. Olshen, and M. Ladanyi, "Integrative clustering of multiple genomic data types using a joint latent variable model with application to breast and lung cancer subtype analysis," *Bioinformatics*, vol. 25, no. 22, pp. 2906–2912, 2009.
- [16] H. Nguyen, S. Shrestha, S. Draghici, and T. Nguyen, "Pinsplus: a tool for tumor subtype discovery in integrated genomic data," *Bioinformatics*, vol. 35, no. 16, pp. 2843–2846, 2019.
- [17] Y. I. Tepeli, A. B. Ünal, F. M. Akdemir, and O. Tastan, "Pamogk: a pathway graph kernel-based multimomics approach for patient clustering," *Bioinformatics*, vol. 36, no. 21, pp. 5237–5246, 2020.
- [18] E. Bair and R. Tibshirani, "Semi-supervised methods to predict patient survival from gene expression data," *PLOS Biology*, vol. 2, no. 4, April 2004.
- [19] D. C. Koestler *et al.*, "Semi-supervised recursively partitioned mixture models for identifying cancer subtypes," *Bioinformatics*, vol. 26, no. 20, pp. 2578–85, 2010.
- [20] E. A. Houseman *et al.*, "Model-based clustering of dna methylation array data: a recursive-partitioning algorithm for high-dimensional data arising as a mixture of beta distributions," *BMC Bioinformatics*, vol. 9, p. 365, 2008.
- [21] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001.
- [22] Y. Lecun and C. Cortes, "The MNIST database of handwritten digits." [Online]. Available: <http://yann.lecun.com/exdb/mnist/>
- [23] National Cancer Institute. (2011) The cancer genome atlas. [Online]. Available: <http://cancergenome.nih.gov/>
- [24] M. Hofree, J. P. Shen, H. Carter, A. Gross, and T. Ideker, "Network-based stratification of tumor mutations," *Nature methods*, vol. 10, no. 11, pp. 1108–1115, 2013.
- [25] N. K. Speicher and N. Pfeifer, "Integrating different data types by regularized unsupervised multiple kernel learning with application to cancer subtype discovery," *Bioinformatics*, vol. 31, no. 12, pp. i268–i275, 2015.



Duygu Ozcelik received her B.Sc. degree in Computer Science from Bilkent University of Ankara, Turkey (2011) and her M.Sc. in Computer Science from Bilkent University of Ankara, Turkey (2016). She has been working as a Senior Software Engineer at Havelsan since 2015.



Oznur Tastan obtained her B.Sc. degree in Biological Sciences and Bioengineering from Sabanci University (2004) and her M.Sc. and Ph.D. degrees in Computer Science from Carnegie Mellon University (2007 and 2011), respectively. She worked as a postdoctoral researcher at Microsoft Research New England. She worked as an assistant professor at Computer Engineering Department of Bilkent University from 2012 to 2017. Since 2017, she has been an assistant professor at the Faculty of Engineering and Natural Sciences of Sabanci University,

affiliated with the Computer Science and Engineering and Molecular Biology Genetics and Bioengineering programs.

Estimation by ANN of Luminous Efficacy of Lamps Used for Lighting

Zuleyha Ok Davarci, Mustafa Sahin and Onur Akar*


Abstract—The increase in energy production costs, and the increasing demand for energy have made energy saving necessary. Energy saving in lighting is able to be actualized by providing economically the same level of light without decreasing the illumination quality. A good lighting is possible by directing required amount of light to spaces desired to be lit. For a good lighting, efficiency factor is a significant measure. By adjusting the luminous efficacy in lamps, the electric energy to be consumed by the lamps is able to be minimized. The electric energy consumed for producing luminous energy will be less at the extent of greatness of luminous efficacy. In this study, the luminous efficacy of lamps of different brands were estimated by the use of data regarding voltage rating, power rating, and total luminous rating of lamps. According to the results of the analysis made by the developed ANN model, the luminous efficacy of lamps of different brands were estimated with an accuracy rate of above 98%. Regarding the lighting systems to be developed, the selection of lamps will be made based on their ability to provide economically the same level of light without decreasing the lighting quality, and to provide the required luminous efficacy for increasing the luminous efficacy.

Index Terms—Luminous Efficacy, Energy saving, Lighting quality, Artificial Neural Network (ANN)


I. INTRODUCTION

ELECTRIC energy has been one of the essential factors required in each phase of our lives [1,2]. In our daily lives, electric energy is used in each field such as lighting, heating, communication health, traffic, workplaces, schools, hospitals, water distribution, security systems, energy production, and transportation [3]. When we consider the challenges faced in transmitting the electricity to consumers along with the high cost of electric energy production techniques, it is able to be observed that economical and efficient use of electric energy is required [4,5]. While 20% of the total electric energy consumption is used for lighting in general in developing


ZULEYHA OK DAVARCI, is with Department of Electrical and Electronics Engineering, Faculty of Technology, Afyon Kocatepe University, Afyon, Turkey, (e-mail: zdavarci@usr.aku.edu.tr).

 <https://orcid.org/0000-0002-5736-4193>

MUSTAFA SAHİN, is with Department of Electronics and Automation, University of Health Sciences, Istanbul, Turkey, (e-mail: mustafa.sahin4@sbu.edu.tr).

 <https://orcid.org/0000-0003-2284-8507>

ONUR AKAR, Department of Electronics and Automation, Marmara University, Istanbul, Turkey, (e-mail: onur.akar@marmara.edu.tr).

 <https://orcid.org/0000-0001-9695-886X>

countries, about 19% of the total electric energy consumption is used for interior and exterior lighting across the globe [6, 7]. The technology is renewed rapidly by the increasing population of the world, and accordingly the amount of energy required for indoor and outdoor lighting increases. Lighting is one of the most significant factors required for people to maintain their lives and affairs at many locations such as buildings, tunnels, hospitals, workplaces, establishments, factories, roads etc. [8,9]. In buildings, actualization of lighting without going to extremes, and making the design as to obtain sufficient light in a manner that will minimize the electric energy costs are the main objectives. For the efficient use of energy in lighting, it is required to make the lighting design as conforming to the intended purpose in the installation phase of the buildings [10]. By doing so, the fields of electricity and architecture are being merged, and environment-friendly new buildings are being created [11]. During the design of lighting systems, it is required to consider their compatibility with the purpose of work to be performed along with energy saving, and the presence of suitable lighting mechanisms. While creating the suitable lighting conditions, criteria such as the direction of light along with its colour, and the sharpness and softness of shades in the light level obtained are also important. It is required to consider these [12]. Moreover, the minimization of the preparation phase and usage period of the lighting systems, and regular maintenance and cleaning of lighting elements are also required. On the other hand, energy saving in lighting is able to be ensured when lamps with high efficiency factor are used, in a well designed lighting system, with a correct maintenance method [13]. Lamps with high luminous efficacy should be preferred in lighting. Because the electric consumed for procurement of equal luminous energy will be less at the extent of greatness of efficiency factor. The lamps with low luminous efficacy will use more electricity from the network. Briefly, in lighting, energy saving is able to be ensured with fixtures of high luminous efficacy [14]. As the type of lamp used on the fixture affects the electrical power and luminous flux, the efficiency will also vary. As the type of lamp used on each fixture, and as the electrical power they consume and their luminous flux are different, the luminous efficacy (efficiency factor) will also be different. For this reason, as the characteristics of lamps such as fluorescent lamp, incandescent lamp, halogen lamp, high-pressure sodium vapor lamp, or LED lamp are different, the luminous efficacy will be directly affected. Moreover, as the additional pieces used on the fixture such as lens, reflector or diffuser keep the light under control,

Manuscript received Nov. 13, 2021; accepted Feb 15, 2022.
DOI: 10.17694/bajece.1022960

they affect the luminous efficacy [15]. There are many studies in literature with respect to lighting systems, new lighting lamps and equipments. Considering the similar studies in literature, O. Olgun et al. evaluated the luminous and electrical performances of lighting fixtures with passive infrared sensor. As the result of their study, they determined that the most efficient fixtures with PIR sensor were ones with LED chip according to the results of electrical and luminous analyses. And they specified that fixtures with E27 lamp base on which LED lamps were used followed the fixtures with PIR sensor. They concluded that the results of economic analyses were parallel with the results of electrical and luminous analyses [16]. And K.M.M. Hasan et al. examined the LED lamps, compact fluorescent lamps, fluorescent lamps, and incandescent lamps, and performed detailed analyses with respect to these lamps. And as the result of their analyses, they obtained the highest efficiency value by the LED lamps. They determined that incandescent fixtures had higher power factor, and that total harmonic distortion (THD) values of LED lamps were higher compared to fluorescent and incandescent lamps [17]. S.Akın examined the fluorescent lamps, and the locations where they were being used in architecture. In the study, the types of fluorescent lamps used today were classified, and it was determined that they were more economic compared to incandescent lamps due to their high luminous efficacy and longer life. As such lamps are highly various in terms of their characteristics of size, shape, power, efficiency, colour etc., they concluded that they were being used in various lighting systems, and that they were able to meet various requirements [18]. Ç.Perdahcı performed energy analysis, through computer simulation program, of a lighting system formed by the use of conventional fixtures at a metal processing industrial plant after replacing the fixtures with fixtures using LED lamp by preserving the level of light. Design of lighting project of the plant was performed by the use of LED fixtures, and fluorescent fixtures operating with electronic ballast. The power consumed by the fixtures used, and value of luminous flux and luminous efficiency of the fixtures were calculated separately for each scenario. The values of luminous flux were compared with the amount of power consumed by two types of lighting fixtures, and electrical power consumption was analyzed. According to the results of simulation of lighting system's design performed for evaluating the efficiency of fluorescent fixture and LED fixture, the luminous efficacy of fluorescent fixture was found as 67.02 lm/W, and the luminous efficacy of fixture with LED lamp was as 141.94 lm/W [19]. And M.T. Gençoglu et al. emphasized the importance of energy saving and energy efficiency, and specified the benefits of energy saving in lighting. They determined the things required to be performed for a correct and efficient lighting, and the things required to be considered while obtaining energy saving. Moreover, they suggested the preference of lamps with high luminous efficacy, with long lifetime, and with low value of luminous flux that arise in time [20]. Ö. Güler et al. showed that the photometric values of fixtures significantly affected the installation and operating costs of fixtures as the result of calculations obtained with a design actualized by forming a sample route model. Moreover, they specified that the values required to be considered in terms of installation and operating costs during

the selection of lamps were economic life, colour characteristics and luminous efficacy. It was specified that the luminous efficacy and especially lives of normal types of metal halogen lamps were lower compared to other lamps. But it was also specified that the new types of metal halogen lamps operated with electronic units recently reached to very good values in terms of lifetime and luminous efficacy [21]. S. Gün Çam examined different lighting techniques that may be used for lighting of building façades. Buildings located in various countries were determined as samples, and lighting systems were examined as per the determined techniques by addressing the buildings in Istanbul lit with LED. It was concluded that LED lamps were being preferred due to their high luminous efficacy, and long lifetime [22]. M. Winterbottom et al. specified that lighting with inefficient fluorescent lamps decreased the visual performance by causing headache. And they specified that, in order to prevent that, automatic blackout lamps should be present in the classes to be built provided that they are non-vibrant. Moreover, it was specified that the values obtained in the beginning vanish in time due to environmental conditions, decrease of luminous efficacy and lifetimes of lamps, and fouling factor [23]. Ç. Perdahcı et al. examined the LED technology, and the value of this technology in terms of energy saving by examining the efficient lighting, energy saving in lighting, and the control systems along with energy efficiency. Moreover, it was pointed out that it is required for the luminous efficacy to have high efficiency classification during selection of lamp [24]. M. Ş. Küçüköğlü examined the effective use energy in lighting by considering that it would be possible through mitigating the formation burden of lighting system, and through minimizing the usage period of artificial lighting systems. Moreover, it was specified that lamps play a significant role in energy consumption especially by their efficiency values. It was specified that, at plants that are being used for a long period continuously along the days or at nights, the selection of lamps, from among the ones having the required quality in terms of colour characteristics, with highest efficiency value is extremely important in terms of effective use of luminous energy [25]. B. Taşkan performed analyses for the determination of required spans in order to prevent the formation of flicker effect in road lighting due to speed, and evaluated the results. He specified that the luminous efficacy of each lamp or fixture was different [26].

Following the examination of all these studies, the luminous efficacy of lamps were defined in general, and they were compared with the luminous efficacy of different lamps. The benefit of luminous efficacy, and its effect in lighting were emphasized.

The change that makes this study different from the others is the estimation of the luminous efficacy of a lamp by ANN model. In here, the ANN model was defined by deeming the power, voltage, flux and brand of lamps as input values, and by deeming the luminous efficacy as output value. It was concluded that the luminous efficacy estimated also by the use test data was compliant at a rate of 98%. By the use of ANN model, the lamp which will provide the luminous efficiency suitable for the environment, as per specific values without the requirement of making calculation each time, will be

determined much easier in this manner, and energy saving in lighting will be ensured.

II. MATERIAL AND METHODS

Examining only one lamp for observing the effect on the lighting of lamps used in lighting systems is not being sufficient. Examining the effects of luminous efficacy on lighting by the use of numerous lamps will enable use to get better results. Moreover, the use of numerical methods during such calculations enables us to reach to more accurate data and solution methods. Working with ANN is being more suitable in numerical and experimental studies in terms of both cost and use of time. In the study performed, the luminous efficacy of lamps were estimated by the use of machine learning method. For the development of the model, the ANN model in the MATLAB (Neural Network Toolbox) software was used.

A. Luminous Efficacy (Efficiency Factor)

The luminous efficacy, as specified in Equation (1), is defined as the value obtained by dividing the luminous flux radiated from a lamp to the power used by the source.

$$e = \frac{\Phi}{P} \quad (1)$$

In the Equation;

- e indicates the luminous efficacy (lm/ W);
- Φ indicates the luminous flux (lumen);
- P indicates the power of lamp (W).

Wind This expression, that expresses the lamp's luminous efficacy in lm/w, should be the most basic criterion required to be considered primarily regarding efficiency. The luminous efficacy is a physical quantity that indicates how well the visible light radiated by a light source is. Factors such as change of outer environment's temperature, differences in lighting manner of the lamp, changes in mains voltage, utilization time affect the efficiency of lamp. The luminous efficacy is one of the most important criteria characterizing a lamp. The electric energy consumed for producing threshold luminous energy will be less at the extent of greatness of luminous efficacy. In theory, the greatest luminous efficacy is obtained in case of a radiation with a single wavelength of $\lambda_0=555$ nm. In this case, it is $V_{\kappa}=1$ for the whole radiation, and $e=683$ lm/ W is found. But in practice, it could be reached only to the level of 200 lm/W. The cause of this is the loss of temperature on lamp, and the wavelength of radiation being 589 nm (in case of 589 nm, then $V_{\kappa}=0.76$) Lamps with high luminous efficacy (lumen/Watt) should be preferred in lighting. Because the electric consumed for procurement of equal luminous energy will be less at the extent of greatness of luminous efficacy. The lamps with low luminous efficacy will use more electricity from the network. And this implies that energy saving is able to be ensured through fixtures with high luminous efficacy in lighting [27-

29]. As the type of lamp used on the fixture affects the electrical power and luminous flux, the efficiency will also vary. As the type of lamp used on each fixture, and as the electrical power they consume and their luminous flux are different, the luminous efficacy will also be different. For this reason, as the characteristics of lamps such as fluorescent lamp, incandescent lamp, halogen lamp, high-pressure sodium vapor lamp, or LED lamp are different, the luminous efficacy will be directly affected. Moreover, as the additional pieces used on the fixture such as lens, reflector or diffuser keep the light under control, they affect the luminous efficacy [30,31]. Each source of light has different luminous efficacy as per the manner and nature of radiation. The parameters such as temperature of outer environment, ballast characteristics, changes in mains voltage, and utilization period cause changes in luminous efficacy.

In this study, brands, and voltage, power, luminous flux values of lamps, and luminous efficacy values of such values were used as data. Regarding the data used in the study, market search was performed for the label values of different types of lamps, and the data was obtained from the catalogues. The luminous efficacy values were formed as being calculated with Equation (1) by the use of catalogue values of lamps subjected to examination. In Fig.1-4, the graph of change of luminous efficacy, used in the ANN training, as per the flux, power, voltage and brand of lamps is shown.

B. Artificial Neural Network (ANN)

ANN has a characteristic similar to neurons in human brain, and it consists of complex systems formed by coherence of artificial neurons with a different connection geometry. By these complex systems, it is being tried to solve the problems that cannot be solved by classic methods [30,32]. ANN is a well accepted method also in the solution of poorly conditioned problems along with complex problems. And the algorithms used are based on the solution of complex differential equations, and powerful computers and time are required for the obtainment of accurate results [33]. Compared to very complex mathematical methods and algorithms, ANN may learn the key information examples from multi-dimensional database [34]. In addition to these, the error tolerance of ANN is high as it can process noisy and deficient data. The weights adjusted each time with each input from outside are representative of actualization of learning. Mathematically, the weight values form the most determinant points of the geometrical figure which tries to best represent all the input values, and which is represented by the regression curve where the sum of distances of all input values is minimum. By this means, the most accurate answer is able to be provided for the value entered [35]. Another factor playing an effective role on learning in ANN is the number of layers used on ANN. Even if the number of layers change among models, a ANN model, consisting of 3 layers, is sufficient even for the most complex problems.

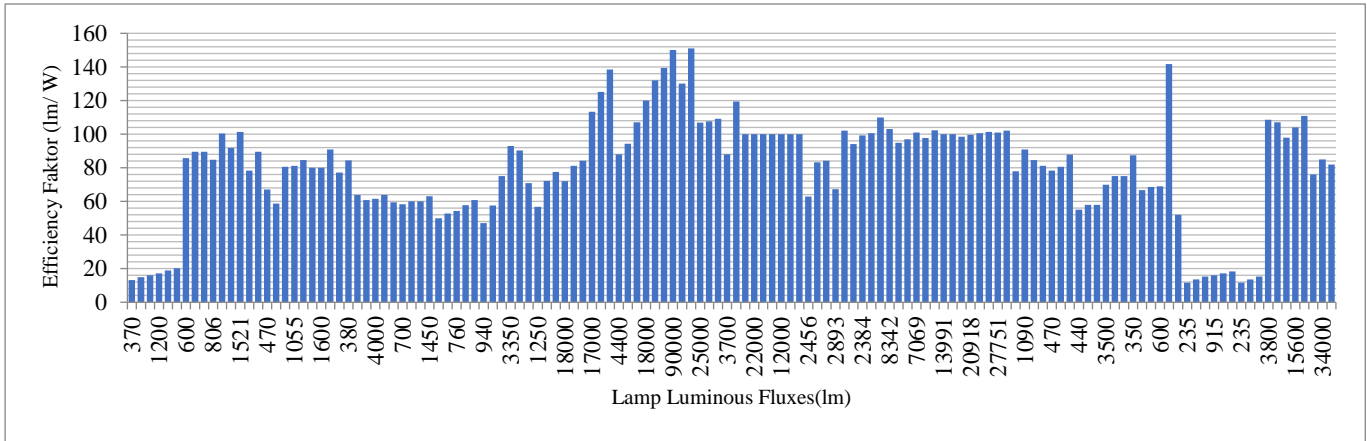


Fig.1. Graph of Change of luminous efficacy as per Luminous Flux of Lamps

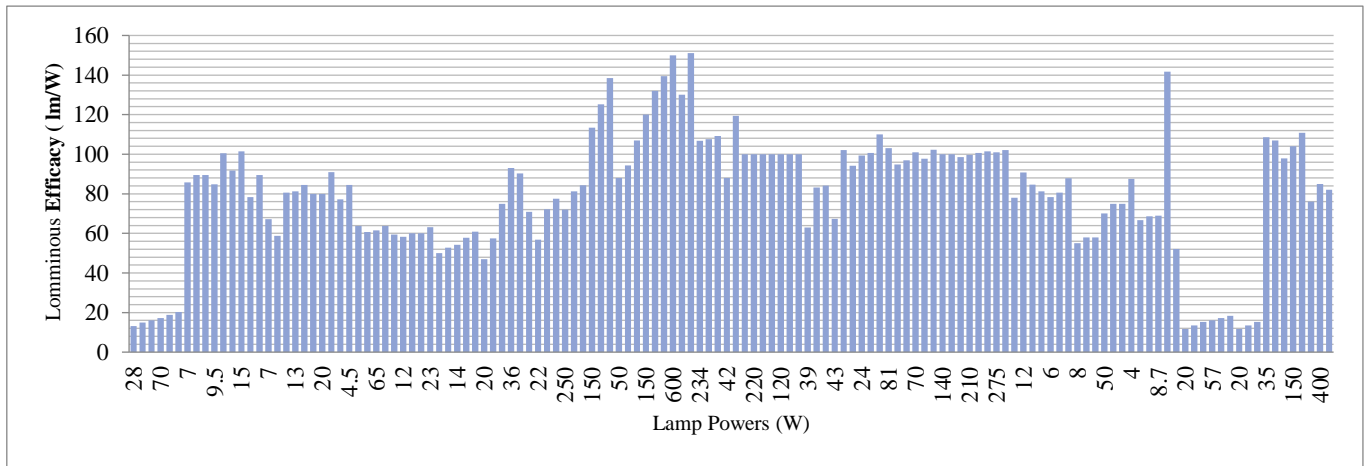


Fig.2. Graph of Change of luminous efficacy as per Power of Lamps

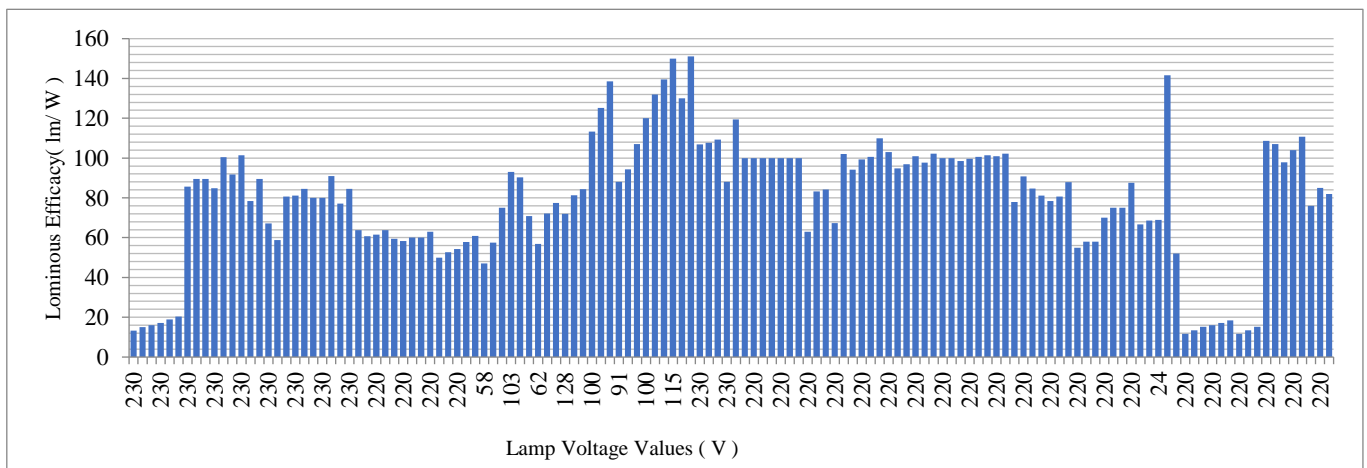


Fig.3. Graph of Change of luminous efficacy as per Voltage of Lamps

This doesn't mean that the problems may not be solved by YSA models with layers below 3 or more than 3. In this 3

layered ANN, there are input layer, hidden layer, and output layer. If it is a 4 layered structure, the 1st layer is called the input

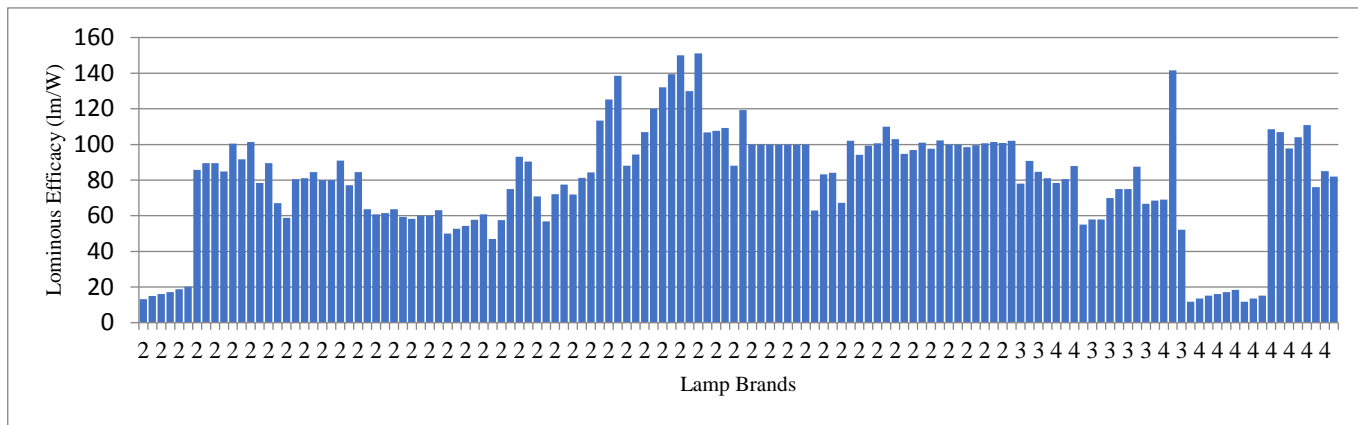


Fig.4. Graph of Change of luminous efficacy as per Brands of Lamps

layer, the 4th layer is called the output layer, and the layers between them are called the hidden layers. In ANN, number of iterations also varies among problems, and as per the self-sensitivity of the problem. High number of cycles will decrease the network performance, and will extend the time of reaching the result, and low number of cycles will cause generation of rough results as to hinder reaching the result. For this reason, while applying ANN for each problem, number of iterations is being tried to be determined by the trial-and-error method [36].

In order to make estimation with ANN, first it is required to train the ANN to be formed. In this study, the variables used for ANN training are input (weight) values, and output (target) values. Brands, and values of voltage, power, total luminous flux of lamps were used as ANN input values. Luminous efficacy values (lumen/Watt) of different lamps were assigned as ANN output values. Input and output variables, used in the

estimation of luminous efficacy by ANN, are as observed in Table 1.

In Fig.5, the input (weight) and output (target) values of ANN model, formed for the estimation of luminous efficacy, are observed. In ANN, it is required to transform the verbal

TABLE I
INPUT (WEIGHT) AND OUTPUT (TARGET) DATA USED IN ANN TRAINING, AND THEIR INTERVAL OF CHANGE

Inputs	Weight and Target	Max.	Min.
X ₁	Brands of lamps	4	2
X ₂	Voltage values of lamps	230	12
X ₃	Power values of lamps	2000	3
X ₄	Total luminous flux values of lamps	180000	235
Y ₁	Luminous efficacy values of lamps	151.11	11.75

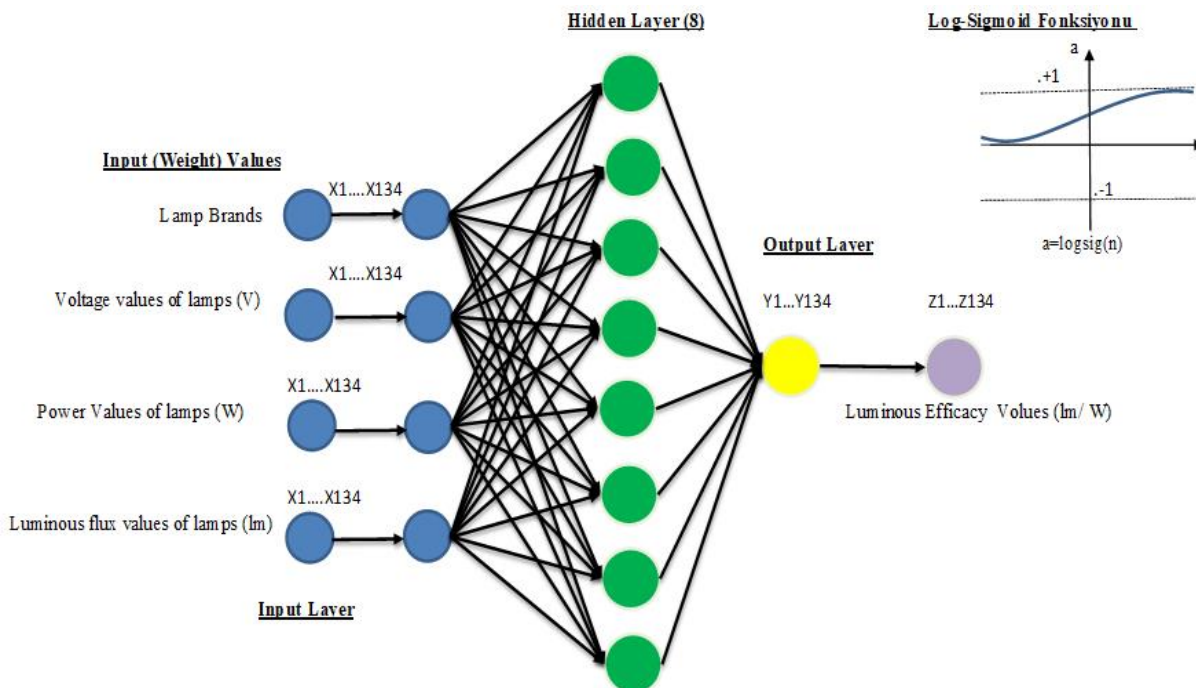


Fig.5. The weight and target values of ANN model for the estimation of luminous efficacy

data to numeric data. Without the transformation, it will not be possible for us to introduce these values to ANN.

In ANN, operation is performed between values of “0” and “1”. For this reason, identification was made by assigning values other than “0” and “1” for the brands. The values assigned are as shown in the below Table 2.

TABLE II
VALUES ASSIGNED TO BRANDS OF LAMPS USED

Brand:	Values
Philips	2
Nade	3
Ostram	4

In ANN, values much greater or much smaller than the normal are able to be observed between the weight and target data. When these are mistakenly entered in the training set, they cause the rise of excessively great or small values during the calculation of net inputs, and they are able to mislead the network. The scaling of all the training data at a specific interval, namely mostly at the interval of 0-1, will assist both the reduction on the same scale of information from different environments, and the elimination of the effect of incorrectly entered excessively great and small values. Thus, before the ANN training phase, the learning performance of ANN is able to be increased by performing some preliminary operations.

The training of ANN is able to be made more efficient by subjecting the weight and target data, to be used in the training of network, to specific operations. This operation is called the normalization operation. The normalization operation is applied on raw training data. Because raw training data is not directly used in ANN training [37]. In this study, the training data was reduced to the interval of (0 and 1) by the use of min.-max. normalization method. Equation 2 was used to reduce the training data to this interval.

$$X' = \frac{(X_i - X_{min})}{(X_{max} - X_{min})} \tag{2}$$

In the Equation;

- X' is the normalized value,
- X_i is the input value of ANN,
- X_{min} is the smallest value of input value,
- X_{max} is the greatest value of input value [38,39].

In this study, following the normalization operation of data used for the training of ANN, the real values were found by making inverse transformation. Change graphs of input and output values, subjected to normalization, used in the training of ANN are shown below. The graph of change of the value of luminous efficacy subjected to normalization in comparison with luminous flux values of lamps subjected to normalization is as observed in Fig.6.

The minimum and maximum values of data, used in ANN training, and subjected to normalization, are as observed in Table 3.

TABLE III
MIN. AND MAX. VALUES OF DATA USED IN THE ANN TRAINING, AND SUBJECTED TO NORMALIZATION

Inputs	Weight and Target	Max.	Min.
X_1	Brands of lamps	1	0
X_2	Voltage values of lamps	0	0
X_3	Power values of lamps	1	0
X_4	Total luminous flux values of lamps	1	0
Y_1	Luminous efficacy values of lamps	1	0

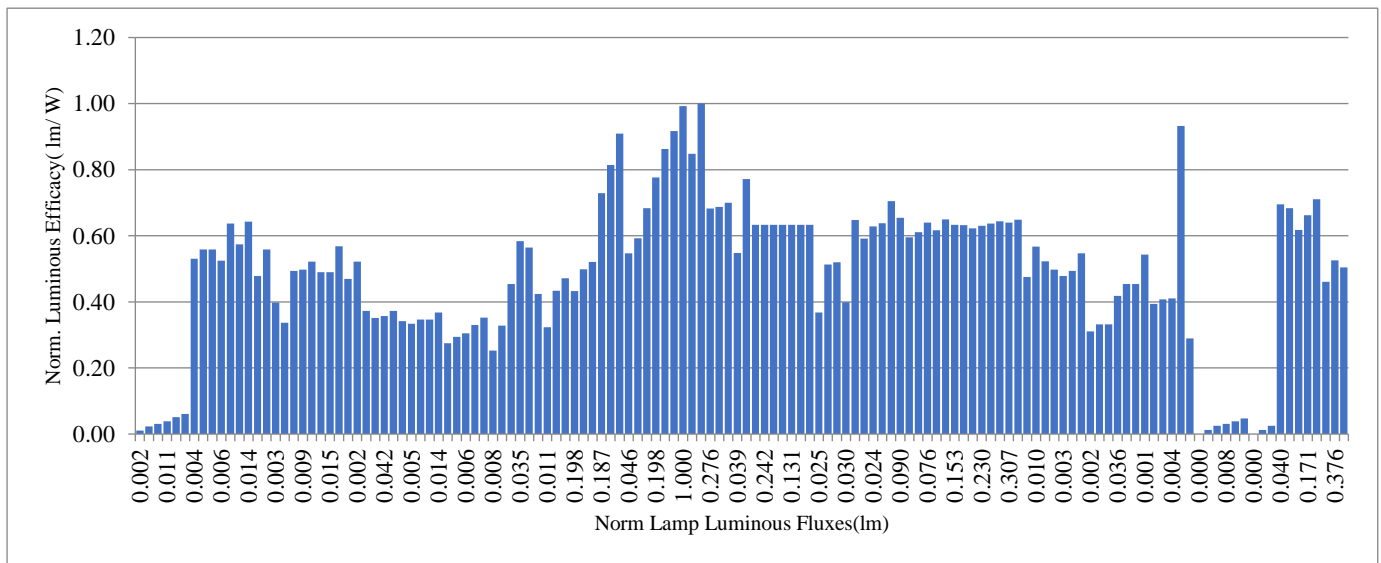


Fig.6. Graph of change of luminous efficacy subjected to normalization in comparison with luminous fluxes of lamps subjected to normalization

The graph of change of the value of luminous efficacy subjected to normalization in comparison with power values of lamps subjected to normalization is as observed in Fig.7.

The graph of change of the value of luminous efficacy subjected to normalization in comparison with voltage values of lamps subjected to normalization is as observed in Fig.8.

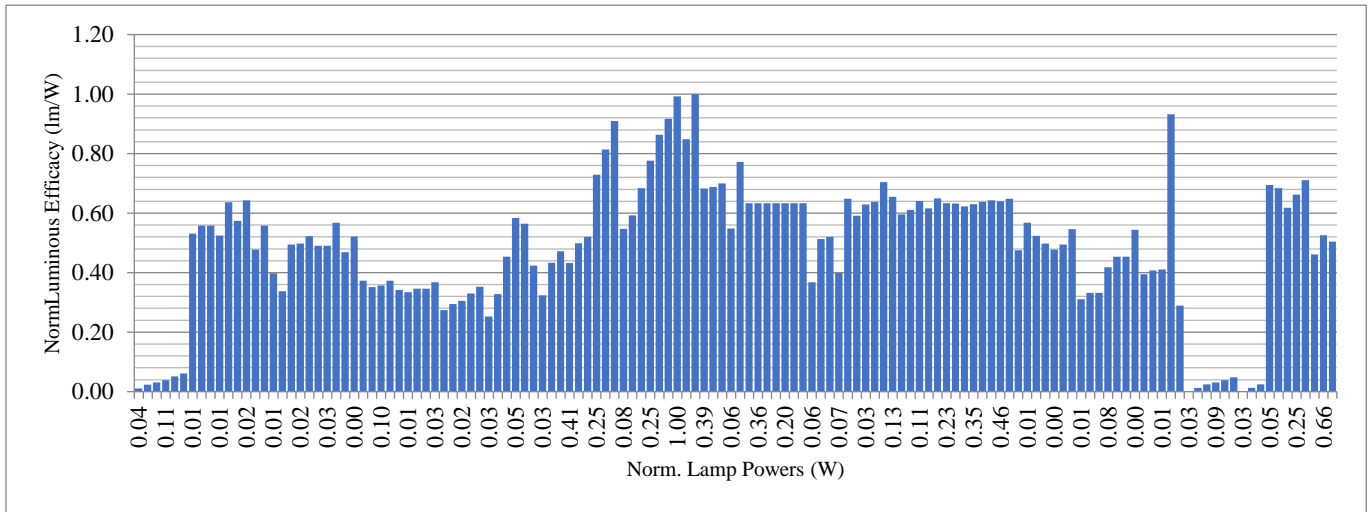


Fig.7. Graph of change of luminous efficacy subjected to normalization in comparison with powers of lamps subjected to normalization

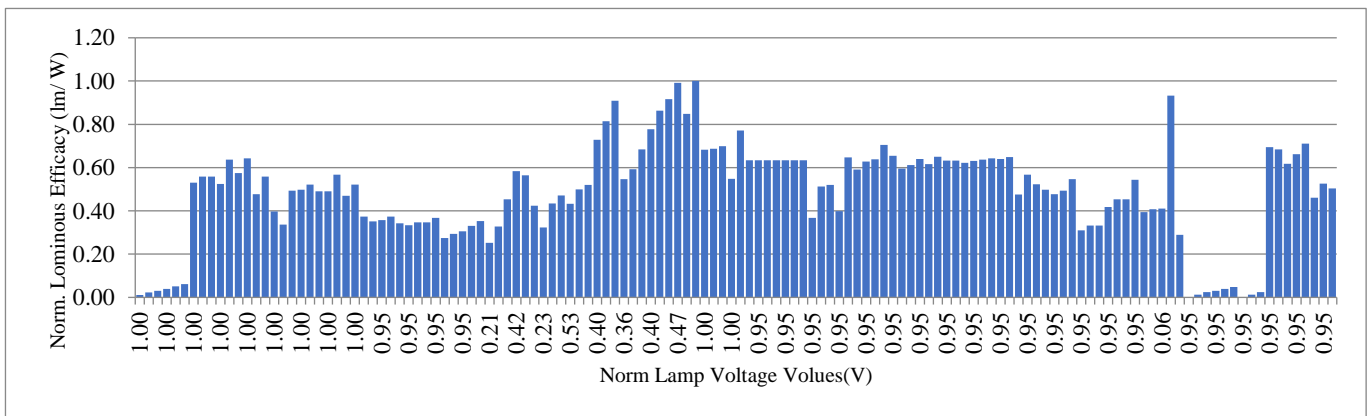


Fig.8. Graph of change of luminous efficacy subjected to normalization in comparison with voltages of lamps subjected to normalization

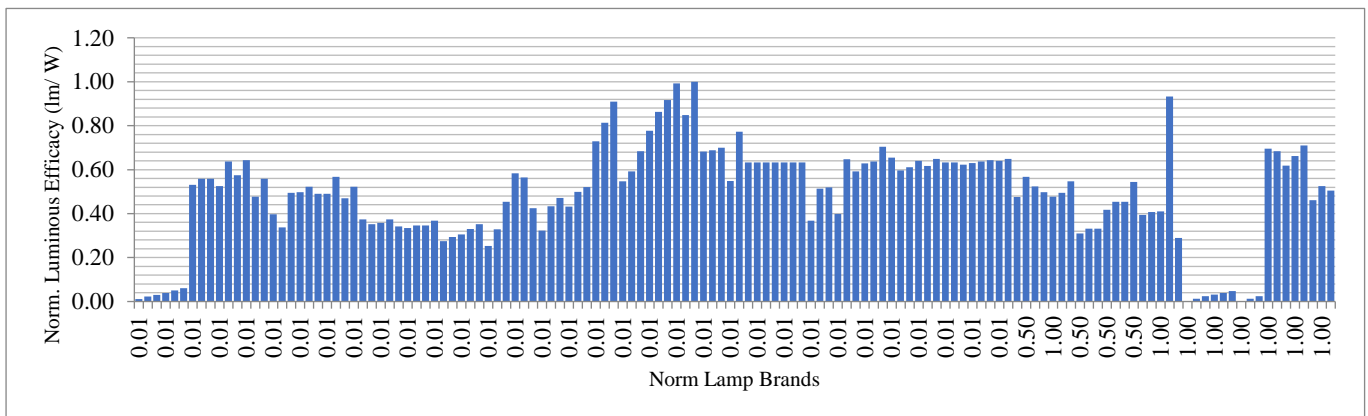
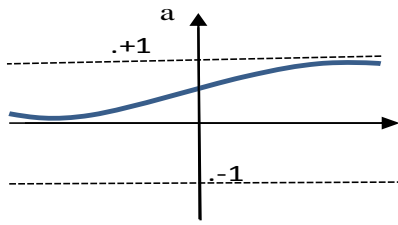


Fig.9. Graph of change of luminous efficacy subjected to normalization in comparison with brands of lamps subjected to normalization

The graph of change of the value of luminous efficacy subjected to normalization in comparison with brands of lamps subjected to normalization is as observed in Fig.9.

III. FINDINGS AND DISCUSSION

In the defined model, feed-forward backpropagation was selected as ANN model for the estimation of luminous efficacy. On the network formation screen, Training function TRAINLM, Adaption learning function, namely LEARNGDM was preferred. And MSE (mean squared error) was used as performance function. And the LOGSIG (Log-Sigmoid Transfer Function) function, observed in Fig.10, was used as transfer function.



a=logsig(n)
Fig.10. Log-Sigmoid Function

The functions and data used in ANN training are as observed in Table 4.

TABLE IV
FUNCTIONS AND DATA USED IN ANN TRAINING

Network Properties	
Network Type	: Feed-Forward backpropagation
Training function	: TRAINLM
Adaption learning function	: LEARNGDM
Performance function	: MSE
Number of layers	: 2
Properties for	: Layer 1
Transfer Function	: LOGSIG

The view of neural network of the defined ANN model is given in Fig.11. In the ANN, there are 4 input values, and 1 output value.

In Fig.12, training, validation, test, and ALL values of the estimation are observed. ANN training value of input values was actualized with a regression of 0.99771. The set up 4x1

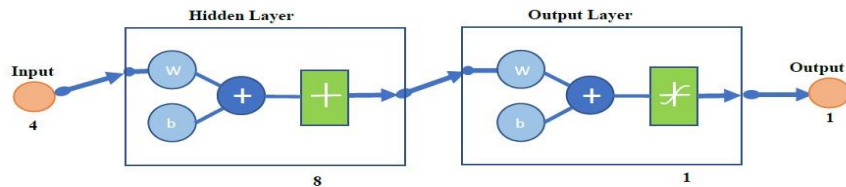


Fig.11. Schematic representation of the defined ANN

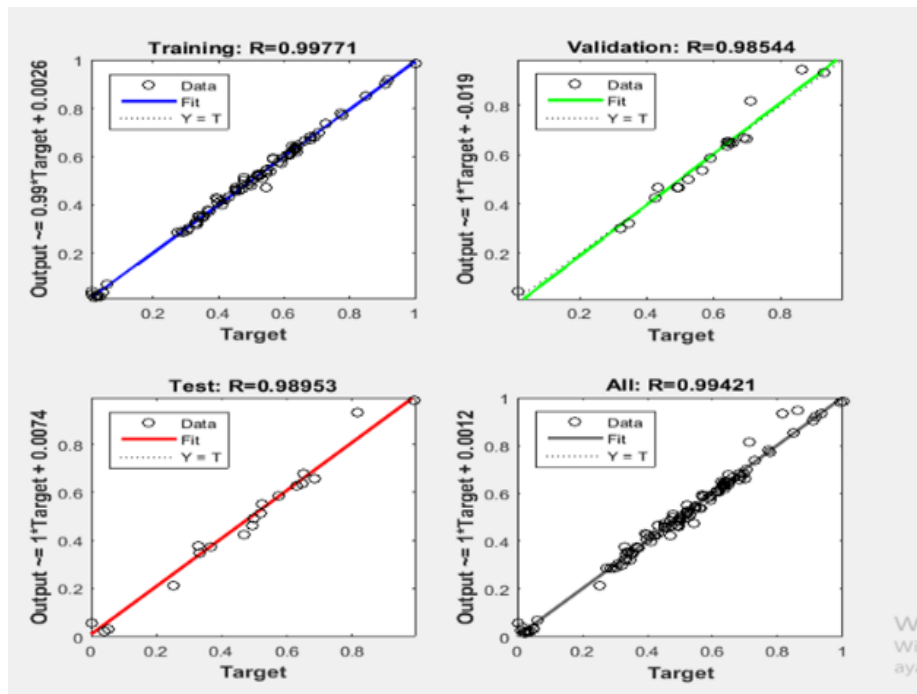


Fig.12. Training, Validation, Test, and All Values in the Estimation Made

ANN estimated validation value was actualized by 0.98544. Test regression value of values, determined as target, was actualized with a high validity as 0.98953. All input regression value was found as 0.99421. When we consider Fig.12, we observe that the training values were found to be correct at a rate of 99.771%. Validation value is also important.

When we consider it, it is observed that it was found correct in ALL, namely in total, at a rate of 99.54%. As the correlation value is equal to one as it is observed, it is observed that there is perfect similarity between network output and target output. As observed in Fig.13, it is observed that the closest value was obtained at about 100th iteration. In the model developed, the best performance value was determined as 0.0013004. In here,

it was intended to estimate the lamps' luminous efficacy by the ANN model developed through the use of specific input variables of lamps.

The graph, indicating that the luminous efficacy values estimated by ANN, after entering weight and target values in ANN, and the currently known luminous efficacy of lamps tally with each other, is observed in Fig.14.

In Fig.15, the 3D graph, indicating that the luminous efficacy provided ANN and the estimated luminous efficacy tally with each other at a rate of 99.771%, is observed.

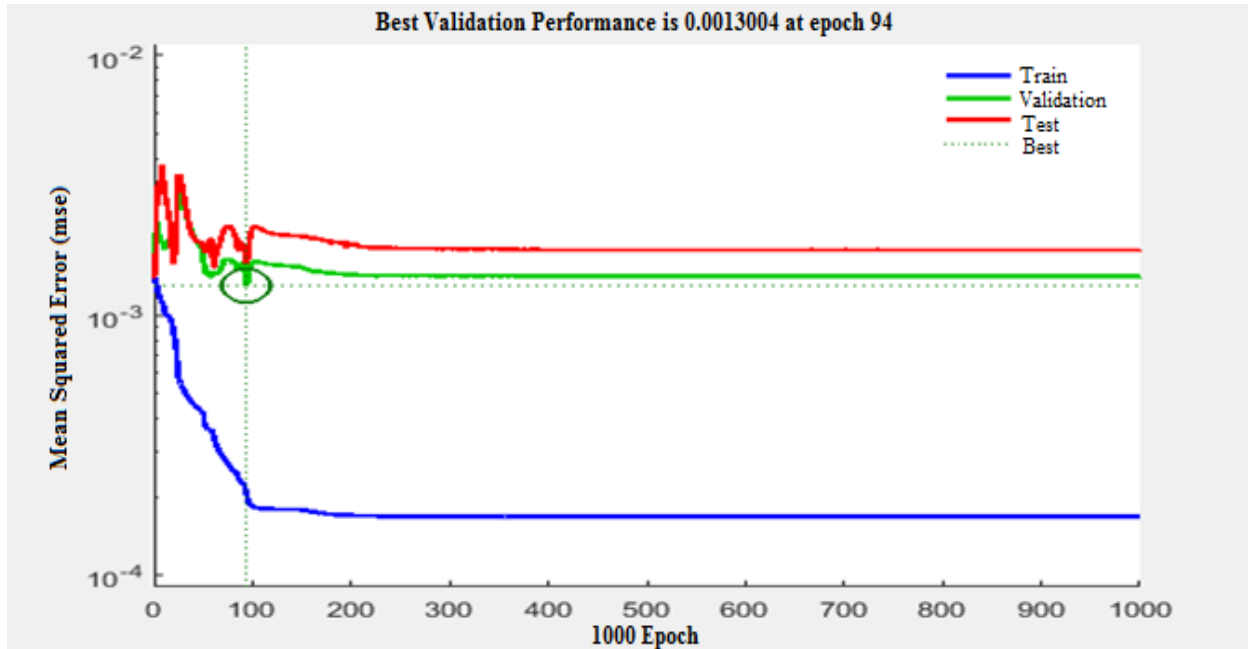


Fig.13. Graph of the Estimation Made

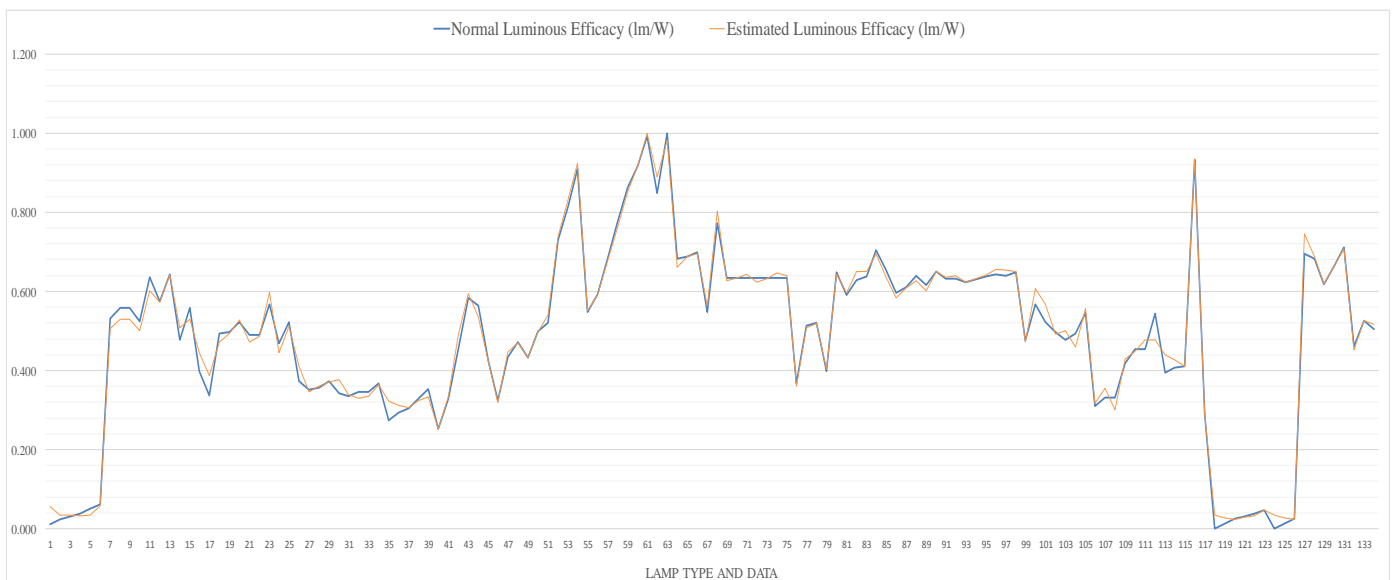


Fig.14. Graph of current values of luminous efficacy, and values of luminous efficacy estimated by ANN

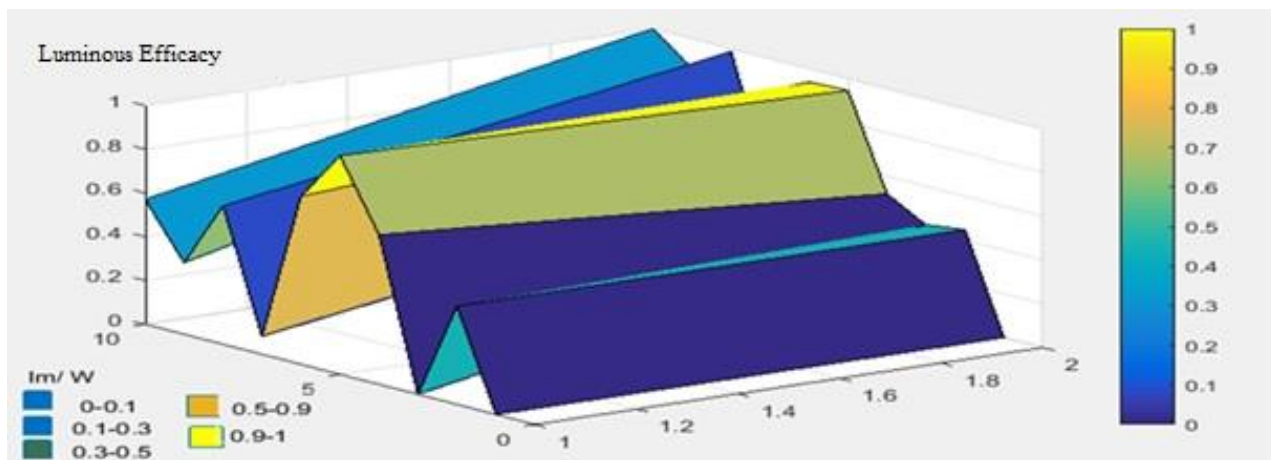


Fig.15. 3D graph indicating that the test values of luminous efficacy, and the estimated values of luminous efficacy tally with each

VI. CONCLUSION

In this study, it was intended to estimate the lamps' luminous efficacy by the ANN model developed through the use of specific input variables of lamps. During the development of ANN model, 536 units of data (4 different characteristics) were evaluated according to 4 different input variables. In the evaluation of ANN model developed, brands, and values of voltage, power, total luminous flux of lamps were used as input values. Luminous efficacy values of different lamps were determined as ANN output values. The set up 4x1 ANN estimated validation value was actualized by 0.98544. Test regression value of values, determined as target, was actualized with a high validity as 0.98953. All input regression value was found as 0.99421. In the model developed, the best performance value was determined as 0.0013004. According to the results of analysis performed by ANN, in the determination of luminous efficacy, the values were estimated correctly at a rate of 99.771% in comparison with input variables. In this manner, for the calculation of luminous efficacy, estimation is made by ANN by entering specific values instead of making calculation one by one. The luminous efficacy of the lamp used will be able to be estimated, and accordingly, utilization of more suitable lamp will be ensured. In addition, both cost saving and energy saving will be ensured due to use of lamp of sufficient power without the utilization of a lamp with higher power. The energy consumed will be less at the extent of greatness of luminous efficacy. By the study, it was observed that the real values, and the values obtained through estimation were very close, and that they were nearly close to 1. It is being considered that this will serve as a model for the future similar studies.

REFERENCES

- [1] R. Kralikova, M. Andrejiova, E. Wessely, "Energy Saving Techniques and Strategies for Illumination in Industry", *Procedia Engineering*, 100, 187-195, 2015.
- [2] Y. Oğuz, M. Şahin, Y. Güven, "Importance of Solar Lighting systems in Terms of Environmental Pollution", *Balkan Journal of Electrical & Computer Engineering*, Volume 3, Issue, 208 - 212, 30.12.2015.
- [3] F. Teng, M. Aunedi, G. Strbac, "Benefits of flexibility from smart electrified transportation and heating in the future UK electricity system", *Applied Energy*, 167, 420-431, 2016.
- [4] J. Bernatt, S. Gawron, T. Glinka, "Energy-Saving Electric Drives", In 2018 International Symposium on Electrical Machines (SME) 1-5, IEEE, 2018.
- [5] Z. K. Küçük, N. Eren, "Light Pollution and Smart Outdoor Lighting", *Balkan Journal of Electrical & Computer Engineering*, Vol. 9, No. 2, April 2021.
- [6] G. Shahzad, H. Yang, A. W. Ahmad, C. Lee, "Energy-Efficient Intelligent Street Lighting System Using Traffic-Adaptive Control.", *IEEE Sensors Journal*, vol. 16, no. 13, July 2016.
- [7] R. Simanjuntak, P. Dupuis, L. Canale, N. I. Sinisuka, G. Zissis, "Power quality of energy saving lamps under wide voltage variations", 2014 IEEE Industry Application Society Annual Meeting, 2014, pp. 1-7.
- [8] A. A. Baloch, P. H. Shaikh, F. Shaikh, Z. H. Leghari, N. H. Mirjat, M. Uqaili, "Simulation tools application for artificial lighting in buildings", *Renewable and Sustainable Energy Reviews*, 82, 3007-3026, 2018.
- [9] M. S. Cengiz, "Effects of Luminaire Angle on Illumination Performance in Tunnel Lighting", *Balkan Journal of Electrical & Computer Engineering*, Vol. 7, No. 3, July 2019.
- [10] IEA, Policies for Energy-Efficiency lighting, Energy Efficiency Policy Profiles, International Energy Agency, 2014.
- [11] G. Tekbiyık, "Sürdürülebilir mimarlıkta yenilenebilir enerji kaynaklarının kullanımı, kamu binalarında uygulama yöntemleri ve örneklerinin incelenmesi", Master's thesis, Fatih Sultan Mehmet Vakıf Üniversitesi, Mühendislik ve Fen Bilimleri Enstitüsü, 2018.
- [12] D. T. De Graaf, Dessouky M., F. O. H. Müller, "Sustainable lighting of museum buildings", *Renewable Energy*, 67, 30-34, 2014.
- [13] Koninklijke Philips Electronics N.V., "The LED Lighting Revolution, A Summary Of The Global Savings Potential.", 2012.
- [14] Ş. Sirel, "Aydınlatma Tasarımında Temel Kurallar", YFU Yayınları, İstanbul, 3-5, 1996.
- [15] Cubuk LED "Etkinlik Faktörü Nedir", 2021. <https://cubukled.com/etkinlik-faktoru-isiksal-verim-nedir/>, 16.06.2021
- [16] O. Olgun, İ. Nakir, "Türkiye'deki Pasif Kızılötesi Sensörlü Aydınlatma Armatürlerinin Işıksal ve Elektriksel Performanslarının Değerlendirilmesi", *Araştırma Makalesi, Avrupa Bilim ve Teknoloji Dergisi Sayı 16, S. 765-775, Ağustos 2019.*
- [17] K. M. M. Hasan, M. S. Rahman, M. A. Rafiq, "Experimental Analysis of Harmonic and Power Factor for Various Lighting Loads", 2017 3rd International Conference on Electrical Information and Communication Technology (EICT), 7-9 Aralık 2017, Khulna, 2017.
- [18] S. Akın, "Flüoresan lambaların gelişim süreci ve mimaride kullanımı", Yüksek Lisans Tezi, Yıldız Teknik Üniversitesi, Fen Bilimleri Enstitüsü, 2000.
- [19] C. Perdahçı, "Metal İşleme Tesis Aydınlatmasında Led Lamba ve Floresan Lamba Karşılaştırılması", *Fırat Üniv. Müh. Bil. Dergisi Science and Eng. J of Fırat Univ.* 30(3), 105-113, 2018.
- [20] M. T. Gençoğlu, E. Özbay, "Aydınlatmada Enerji Verimliliği Yöntemleri", *Fırat Üniversitesi, mühendislik Fakültesi*, syf.1-5, 2007.

- [21] Ö. Güler, Ş. Onaygil, "Yol Aydınlatması Tesisatlarında Armatür Fotometrik Değerlerinin Önemi", İstanbul Teknik Üniversitesi, Enerji Enstitüsü, Syf. 1-11, İstanbul, 2015.
- [22] S. Çam Gün, L. Dokuzer Öztürk, "Yapı Yüzü Aydınlatma Teknikleri ve Uygulama Örnekleri", VIII. Ulusal Aydınlatma Sempozyumu, Syf.142-158, TMMOB Elektrik Mühendisleri Odası, İzmir, 2015.
- [23] M. Winterbottom, A. Wilkins, "Lighting and discomfort in the classroom", Journal of Environmental Psychology, 29(1), 63-75, 2009.
- [24] C. Perdahçı, U. Hanlı, "Verimli Aydınlatma Yöntemleri", Kocaeli Üniversitesi Mühendislik Fakültesi, Syf.1-5.
- [25] M. Ş. Küçükdoğan, "Aydınlatmada Etkin Enerji Kullanımı", İTÜ Mimarlık Fakültesi, Syf.1-5, İstanbul, 2019.
- [26] B. Taşkan, "Flicker Etkisi Esaslı Yol Aydınlatmasında Optimum Direk Açıklığının Belirlenmesi", Bilecik Şeyh Edebali Üniversitesi, Fen Bilimleri Enstitüsü Elektrik-Elektronik Mühendisliği, Syf. 1-79, Yüksek Lisans Tezi-2016.
- [27] A. Ünal, S. Özenç, "Aydınlatma Tasarımı ve Proje Uygulamaları", Birsen Yayınevi, İstanbul 2004.
- [28] M. Özkaya, "Aydınlatma Tekniği", Birsen yayınevi, İstanbul,2000.
- [29] R. Ganslandt., H. Hofmann, "Handbook of Lighting Design.", Erco Edition, Germany, 1992.
- [30] M. Şahin, "Yapay Sinir Ağları ile Dahili Ortamlardaki Aydınlık Düzeyinin Analizi", Marmara Üniversitesi Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, İstanbul, 2010.
- [31] S. Dursun, U. K. Terzi, O. Akar, T. Sonmezocak, "Comparative Analysis of Lighting Elements Effects on Electric System", European Journal of Technique, Vol.11, No.2, 2021.
- [32] A. Sözen, M. A. Akçayol, "Modelling (Using Artificial Neural-Networks) the Performance Parameters of a Solar-Driven Ejector-Absorption Cycle", Applied Energy, 79(3): 309-325, 2004.
- [33] A. Arslan, R. Ince, "The neural network approximation to the size effect in fracture of cementitious materials", Engineering Fracture Mechanics, 54(2): 249-261,1996.
- [34] M. Şahin, "Yapay Sinir Ağları ile Aydınlik Kalitesi Kontrolü", Article, AKÜ FEMÜBİD 13 (2013) 025201 (1-10),2013.
- [35] Ç. Öztemel, "Yapay Sinir Ağları" Papatya Yayıncılık 49-52, 2003.
- [36] L. Fausett, "Fundamentals of Neural Network, Architectures, Algorithms, And Applications", 1994.
- [37] H. M. Beale, T. M. Hagan, B. H. Demuth, "Neural Network Toolbox User's Guide", Revised for Version 8.0 (Release 2012b), 2012.
- [38] L. R. Medsker, "The future of artificial neural networks could be bright", Computers/Control engineering 10, 28-9,1997.
- [39] I. Villaverde, M. Granna, "Neuro-evolutionary mobile robot egomotion estimation with a 3D ToF camera", Neural Computing and Applications, 20, 345- 354, 2011.

BIOGRAPHIES



ZULEYHA OK DAVARCI was born in Lübeck, Germany. She graduated from Kırıkkale University Electrical and Electronics Engineering in 2002. She is doing his master's degree in Afyon Kocatepe University Faculty of Technology, Department of Electrical and Electronics Engineering. His research areas include Lighting and systems, ANN and its applications.



MUSTAFA ŞAHİN received the B. Eng. in Electrical Technology from Marmara University in 2006, M.Sc. and PhD from Marmara University, Turkey in 2010 and in 2014 respectively. Dr. Şahin worked an Assistant Professor in Faculty of Engineering and Departement Electrical &

Electronics Engineering at Erzincan Binali Yıldırım University and after he worked an Assistant Professor of at Department Electronics Engineering at Afyon Kocatepe University. He currently is an Assisrant Professor of Biomedical Device Technology Program at University of Health Sciences. His research areas include Lighting and systems, Power Systems, ANN and its applications.



ONUR AKAR was born in 1981 in Giresun.He received his undergraduate, graduate and doctorate degrees from Marmara University in 2005, 2011 and 2020, respectively. He worked as a lecturer at Istanbul Gedik University between 2010-2020. He served as the Head of Electricity Program at Istanbul Gedik University between 2012-2015. He served as the Head of the Department of Electricity and Energy between 2021-2022 as an assistant professor in the Department of Electricity and Energy of the same university. He is still working as an Assistant Professor Marmara University, Vocational School of Technical Sciences. His research interests include Control Systems, Renewable Energy Systems, Power Systems and Lighting Systems.

Coefficient Diagram Method Based Decentralized Controller for Fractional Order TITO Systems

Miray Günay Bulut, Furkan Nur Deniz *


Abstract— Fractional calculus has gained increasing attention from researchers because of providing accurate modelling and flexible controller design in control applications. More research to design controllers for Fractional Order Two-Input Two-Output (FOTITO) systems, which inherently have certain difficulties, is needed when the studies on such control applications are reviewed. In this study, Coefficient Diagram Method (CDM) based decentralized controllers are designed for FOTITO systems. To this end, integer order approximate models of FOTITO systems are obtained and decoupled into two subsystems by using simplified and inverted decoupling configurations. The resulting higher-order approximate subsystem transfer functions are reduced by a model reduction method to facilitate CDM-based decentralized controllers design. Then, CDM-based decentralized controllers are designed by using each subsystem, which enable to control the FOTITO system. Simulation results for two different FOTITO systems, one of which is time delayed, show that the proposed approach exhibits successful performance.

Index Terms— Coefficient Diagram Method, Decentralized Controller, Fractional Order Systems


I. INTRODUCTION

INDUSTRIAL SYSTEMS generally consist of multiple-input multiple-output (MIMO) processes which contain multivariable and multiple loops. MIMO systems have complicated loops that cause difficulties on the control of such systems. These complicated loops, where an input variable affects all output variables, can lead to unexpected interaction problems [1].

MİRAY GÜNAY BULUT, is with Department of Electrical and Electronics Engineering University of Adiyaman, Adiyaman, Turkey, (e-mail: mgunay@adiyaman.edu.tr).

 <https://orcid.org/0000-0001-6479-707X>

***FURKAN NUR DENİZ**, is with Department of Electrical and Electronics Engineering University of Inonu, Malatya, Turkey, (e-mail: furkan.deniz@inonu.edu.tr).

 <https://orcid.org/0000-0002-2524-7152>

Manuscript received August 20, 2021; accepted April 13, 2022.
DOI: [10.17694/bajece.984815](https://doi.org/10.17694/bajece.984815)

The interaction problems arising in these systems can adversely affect control performance of the system. For this reason, it is necessary to consider the interaction problem when designing controllers for MIMO systems. One of the techniques to eliminate the interaction problem and simplify the controller design is the decoupling technique [2], [3]. Decoupling techniques provide that the MIMO systems are decomposed into single-input single-output (SISO) subsystems by extra decouplers included into the systems, and they enable the design of decentralized controllers for each loop by considering the interaction problems [4]. Various methods such as ideal, simplified and inverted decoupling have been given for the decoupling of systems in the literature [5]–[7]. Among these methods, simplified and inverted decoupling methods are more preferred in applications due to their simplicity [3], [8], [9].

Decoupling methods and formulations are generally applied to two-input two-output (TITO) systems, which are the simplest type of the MIMO systems. Thus, various control strategies developed for TITO systems, which are relatively less complex and easily configurable, can also be generalized for MIMO systems [10]. There are many studies on the design of decentralized controllers for TITO systems using decoupling methods. For example, decentralized controller design based on Characteristic Ratio Assignment (CRA) for TITO systems, which use conventional decoupling method, was presented in [11]. A decentralized PID controller technique was suggested for TITO systems which used reduced models of diagonal elements according to Nyquist plots fitting [12]. An internal model control strategy was proposed for TITO systems by using conventional and inverted decoupling methods [13]. Stability regions of decentralized PI controllers for TITO systems were obtained to calculate all stabilizing controller parameters [8].

Fractional order differential equations more accurately represent the model of a dynamic system [14]. In this direction, various fractional order models have been developed for TITO systems [15]–[17]. Although the fractional order calculus has the advantage of yielding more accurate modeling for dynamic systems, analysis and control of such systems can be quite difficult due to their complexity. Some studies on the fractional order TITO (FOTITO) control systems have been presented in the literature. For example, simplified, inverted and ideal decoupling techniques were expanded for FOTITO systems considering properness and frequency dependent RGA (relative gain array) [16]. An FOTITO model, which has more accurate

time response, for a prototype of a hydraulic canal was proposed by using experimental data [18]. All stabilizing parameters of decentralized PI controllers for an FOTITO thermo-electric temperature process model, which uses inverted and simplified decoupling, were presented [9]. A fractional order IMC (Internal Model Control) methodology, which exhibits robust performance for set point tracking and disturbance rejection, was suggested for FOTITO systems [15]. In another study, fixed low order decentralized controllers were given for FOTITO systems by using optimization according to CRA (Characteristic Ratio Assignment) method [17]. There is a need for more research and development of control algorithms for FOTITO control systems which allow to obtain better and more accurate mathematical models.

In current study, decentralized controllers are designed for FOTITO systems by using the Coefficient Diagram Method (CDM) [19]. Integer order approximate models of FOTITO systems are obtained by using the M-SBL fitting approximation method [20]. Then, the subsystem models, which are configured with simplified and inverted decoupling methods, are used to design decentralized controllers. Since these decoupled approximate models become higher order models, they are transformed to more practicable models suitable for CDM by using a model reduction method [21]. CDM has generally been used to design controllers for SISO systems [19], [22]–[25]. Since industrial systems consist of MIMO systems, CDM based controller design has been extended to MIMO systems. In one of Manabe's studies on MIMO systems [26], CDM was used to control a dual-control surface missile. In the same study, Manabe additionally decomposed MIMO systems into single-input multiple-output (SIMO) systems and performed controller design by applying the CDM design procedure to each SIMO system. Similarly, in another study [27], Manabe applied the CDM in the control of a fighter with dual control surfaces, which consists of a MIMO system. Also, decentralized controller design [28], decentralized PI controller [29], [30] and PID controller [31] designs by using CDM were presented for TITO systems. In this study, unlike the others, the systems are fractional order systems which bring design difficulties due to its nature.

The main advantage of CDM is the ability to predefine specifications of time response such as settling time and overshoot during the design process [25]. Thus, CDM offers the possibility to guarantee a simple and robust controller design [25]. The basic properties of CDM are briefly as follows; i) There is a little or no overshoot. ii) Two degrees of freedom (2DOF) control system is used. iii) It has robust performance to changes in parameters and limited uncertainties in the control system. iv) It provides convenience to the designer thanks to the predefined settling time [22]–[25]. In this study, CDM based decentralized controllers are designed for two different FOTITO systems, and simulation results validate that the proposed approach provides good performance.

II. FRACTIONAL ORDER TITO SYSTEMS AND DECOUPLING METHODS

A FOTITO system can be expressed as a transfer function matrix given below.

$$G(s) = \begin{bmatrix} G_{11}(s) & G_{12}(s) \\ G_{21}(s) & G_{22}(s) \end{bmatrix} \quad (1)$$

where each element of transfer function matrix $G_{ij}(s) = \sum_{m=0}^p u_m s^{\beta_m} / \sum_{n=0}^q v_n s^{\alpha_n}$ can be expressed as fractional-order transfer functions for $\alpha_i, \beta_j \in R$, $i, j = 1, 2$.

$G_{ij}(s) = K_{ij} e^{-\theta_j s} / (T_{ij} s^{\alpha_{ij}} + 1)$ can be expressed as a fractional order system with time delay. The time delay should be considered while applying the decoupling techniques to FOTITO systems. In this context, if $\theta_{11} \leq \theta_{12}$ and $\theta_{22} \leq \theta_{21}$, the decoupler becomes causal and can be realized easily [32].

When the system is an FOTITO process, decoupling methods can be applied to facilitate the design of the controllers by considering the interaction problem [16]. The most well-known decoupling methods are ideal, simplified and inverted decoupling methods [33]. Researchers should decide on the selection of the decoupling method for the system by considering the advantages and disadvantages of these methods. The advantage of ideal decoupling is that the controller transfer matrix is very easy to determine, the disadvantage is that it has realization problems due to the complexity of the decoupler elements [34]. In simplified decoupling, the values of the diagonal elements of the decoupler matrix are assumed to be 1, so it is easy to determine the decoupling elements. Weischedel et al. [6] observed that simplified decoupling is more robust compared to ideal decoupling in their study. The inverted decoupling has a combination of the positive properties of ideal and simplified decouplings [7]. In this study, the simplified and inverted decoupling methods are applied to FOTITO systems. Fig. 1 shows the block diagrams of simplified and inverted decoupling methods.

The decoupling transfer function matrix $D(s)$ and the diagonal transfer function matrix $T(s)$ can be written as follows [33],

$$D(s) = \begin{bmatrix} D_{11}(s) & D_{12}(s) \\ D_{21}(s) & D_{22}(s) \end{bmatrix} \quad (2)$$

$$T(s) = \begin{bmatrix} T_1(s) & 0 \\ 0 & T_2(s) \end{bmatrix} = G(s)D(s) \quad (3)$$

The decoupler transfer function matrix of the simplified and inverted decouplings are the same and expressed as follows:

$$D(s) = G(s)^{-1} T(s) = \begin{bmatrix} 1 & -\frac{G_{12}(s)}{G_{11}(s)} \\ -\frac{G_{21}(s)}{G_{22}(s)} & 1 \end{bmatrix} \quad (4)$$

The diagonal transfer matrix $T(s)$ of the simplified decoupling is calculated by using Equations (1), (3) and (4) as follows:

$$T(s) = \begin{bmatrix} G_{11}(s) - \frac{G_{12}(s)G_{21}(s)}{G_{22}(s)} & 0 \\ 0 & G_{22}(s) - \frac{G_{12}(s)G_{21}(s)}{G_{11}(s)} \end{bmatrix} \quad (5)$$

The diagonal transfer matrix $T(s)$ of the inverted decoupling is calculated as:

$$T(s) = \begin{bmatrix} T_1(s) & 0 \\ 0 & T_2(s) \end{bmatrix} = \begin{bmatrix} G_{11}(s) & 0 \\ 0 & G_{22}(s) \end{bmatrix} \quad (6)$$

The diagonal transfer matrix of the controller $C(s)$ is defined as

$$C(s) = \begin{bmatrix} C_1(s) & 0 \\ 0 & C_2(s) \end{bmatrix} \quad (7)$$

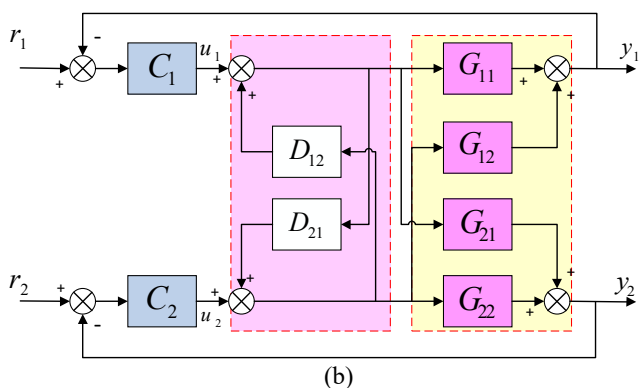
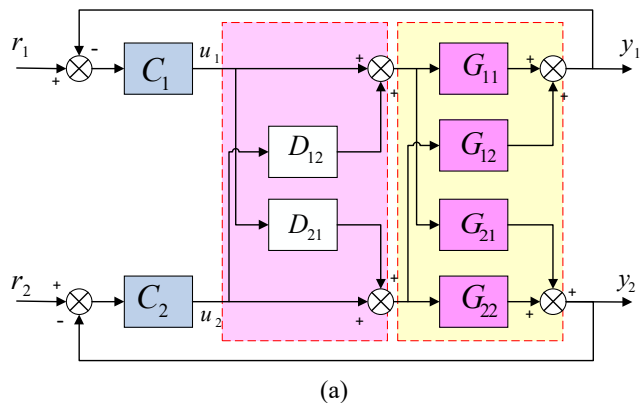


Fig.1. a) Simplified decoupling, b) Inverted decoupling

FOTITO system is decoupled into two SISO subsystems by using simplified and inverted decoupling, and the transfer function of each subsystem is denoted as $T_i(s)$, and its controller is denoted as $C_i(s)$ for $(i=1,2)$.

Block diagram of each subsystem is given in Figure 2.

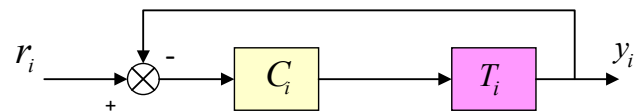


Fig.2. Feedback control system for subsystems.

III. COEFFICIENT DIAGRAM METHOD

Coefficient Diagram Method developed by Manabe in 1991 [19], is an algebraic method that is used to design a high order controller. CDM is a polynomial approach where the numerator and denominator polynomials of the transfer function of the system are treated independently. The advantage of the CDM based control system is to predefine settling time and maximum overshoot at the beginning of the design task [25].

CDM based configuration illustrated in Fig. 3 has a two-degree of freedom control system, which can better perform in respect to set-point tracking and disturbance rejection [35].

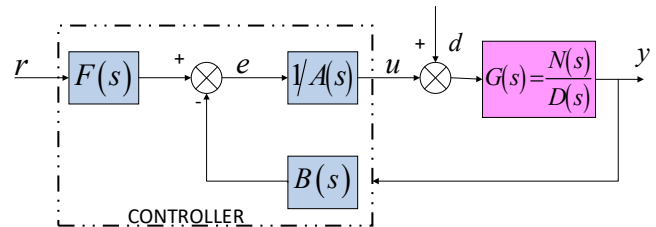


Fig.3. CDM based control system

In the CDM based control system in Fig. 3, $N(s)$ and $D(s)$ are numerator and denominator polynomials of the plant transfer function $G(s)$, respectively. $A(s)$ is denominator of the controller, $B(s)$ is feedback numerator of the controller, and $F(s)$ is the reference numerator of the controller [19].

The output of this control system can be written as follows [25].

$$y = \frac{N(s)F(s)}{P(s)}r + \frac{A(s)N(s)}{P(s)}d \quad (8)$$

Characteristic polynomial of the closed loop system is

$$P(s) = D(s)A(s) + N(s)B(s) = \sum_{i=0}^n a_i s^i \quad (9)$$

where the controller polynomials $A(s)$ and $B(s)$ are defined as

$$A(s) = \sum_{i=0}^p l_i s^i \quad \text{and} \quad B(s) = \sum_{i=0}^q k_i s^i ; p \geq q \quad (10)$$

Reference numerator of the controller $F(s)$ is

$$F(s) = \left(\frac{P(s)}{N(s)} \right)_{s=0} \quad (11)$$

There is an important point to be considered in the selection of $A(s)$ and $B(s)$ polynomials. In order to achieve the desired time response performance, if the system transfer function does not include an integrator, $l_0 = 0$ should be selected to completely eliminate the effects of the disturbance [28].

CDM design parameters are expressed with equivalent time constant τ , stability index γ_i and stability limit index γ_i^* [19]. They can be written in terms of the coefficients of the characteristic polynomial as follows:

$$\tau = \frac{a_1}{a_0}, \quad \gamma_i = \frac{a_i^2}{(a_{i-1}a_{i+1})}; i = 1, \dots, n-1 \quad (12)$$

$$\gamma_i^* = \frac{1}{\gamma_{i+1}} + \frac{1}{\gamma_{i-1}}; \gamma_0 = \gamma_n = \infty$$

Manabe proposed that the stability indexes are selected as $\gamma_1 = 2.5$, $\gamma_i = 2$ and $\gamma_0 = \gamma_n = \infty$ for $i = 2 \sim (n-1)$ so that CDM based controller design provides the desired performance [19]. These proposed values can be changed to achieve the desired control performance according to $\gamma_i > 1.5$ for all $i = 1 \sim (n-1)$ [28]. Using these parameters, the coefficients of the characteristic polynomial are determined as follows:

$$a_i = \frac{a_0 \tau^i}{\gamma_{i-1} \gamma_{i-2} \dots \gamma_1^{i-1}} \quad (13)$$

According to the standard Manabe form [19], the settling time is calculated with $\tau = t_s / (2.5 \sim 3)$. After determining the desired settling time and stability indexes of the system, the target characteristic polynomial is defined with the equation given below [28].

$$P_{\text{target}}(s) = a_0 \left[\sum_{i=2}^n \left(\prod_{j=1}^{i-1} \frac{1}{\gamma_{i-j}^j} \right) (\tau s)^i \right] + \tau s + 1 \quad (14)$$

By matching the target characteristic polynomial to the characteristic polynomial of the system, the controller parameters k_i and l_i can be calculated as follows.

$$A(s)D(s) + B(s)N(s) = P_{\text{target}}(s) \quad (15)$$

When the CDM based control system is adapted to the 2DOF control system as given in Fig. 4, it exhibits better performance for set point tracking and disturbance rejection [22]

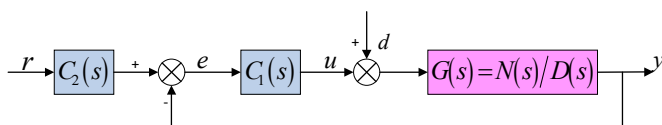


Fig. 4. Two-degree-of-freedom (2DOF) control system

The controllers $C_1(s)$ and $C_2(s)$ are defined by using the CDM polynomials as follows:

$$C_1(s) = \frac{B(s)}{A(s)}, \quad C_2(s) = \frac{F(s)}{B(s)} \quad (16)$$

This modification, which is made according to the 2DOF control system, facilitates the use of CDM in FOTITO systems with decoupling. The simplified and inverted decoupling block diagrams can be redesigned by using the 2DOF structure [31] as shown in Fig. 5, respectively. In Equation (16), $C_{11}(s)$ and $C_{12}(s)$ correspond to $C_1(s)$ and $C_2(s)$ for r_1 - y_1 relation and $C_{21}(s)$ and $C_{22}(s)$ correspond to $C_1(s)$ and $C_2(s)$ for r_2 - y_2 relation, respectively.

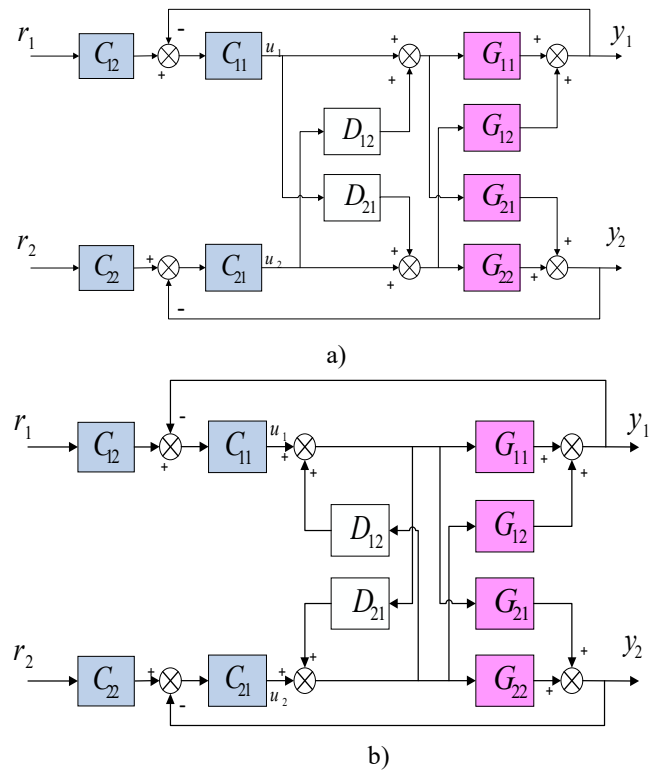


Fig. 5. Modified FOTITO systems based on 2DOF control: a) simplified decoupling b) inverted decoupling.

IV. APPROXIMATE MODELS FOR IMPLEMENTING FRACTIONAL ORDER TRANSFER FUNCTIONS

The realization of fractional order derivative and integrator operators is inherently difficult due to their long memory effects [36]. Therefore, approximate integer order models of fractional order operators are used in the realization task. Several approximation methods such as Oustaloup's method [37], Matsuda's method [38], SBL (Stability Boundary Locus) fitting method [39] and its modified version M-SBL (Modified Stability Boundary Locus) fitting method [20] were presented in the literature to obtain these approximate models. In this study,

authors prefer to use the M-SBL fitting approximation method [20], which is based on fitting the stability boundary curves of the fractional order operator with its integer order approximation model. Thus, it can be ensured that the integer order approximate models can preserve stabilization of fractional order transfer functions in the realization process [40].

Using M-SBL fitting approximation method, one can obtain an n^{th} integer order approximate model for a fractional order derivative operator s^α in a desired frequency range $\omega \in [\omega_l, \omega_h]$ rad/sec as follows:

$$s^\alpha = \frac{a_0 s^n + a_1 s^{n-1} + \dots + a_{n-1} s + a_n s^0}{a_n s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0 s^0}, \quad 0 < \alpha < 1 \quad (17)$$

The M-SBL fitting Matlab function [41] is used to obtain the integer order approximate transfer functions of the fractional order derivative operators. This subject was studied in [20] comprehensively.

V. SUB-OPTIMAL MODEL REDUCTION

The sub-optimal model reduction is a reduction method based on optimization, and its algorithms are simple [21]. This method reduces the size of the coefficient matrix of the system while preserving the dominant eigenvalues or most important states of the original system. Target of the model reduction method is that it can easily reduce the model to the most realistic and desired degree [21].

ISE (Integral Square Error) criterion commonly used in determining the quality of the reduced model according to the error signal $e(t)$ is expressed as follows.

$$I_{ISE} = \int_0^{\infty} e^2(t) dt \quad (18)$$

Using the ISE criterion, the error rate is minimized, and the model reduction problem is converted into an optimization problem. The approximate error of the model can be written as $\hat{e}(t, \theta)$ for a particular original model and input signal. The sub-optimal model reduction objective function can be given as follows:

$$J = \min_{\theta} \left[\int_0^{\infty} \omega^2(t) \hat{e}^2(t, \theta) dt \right] \quad (19)$$

where $\omega(t)$ is the weighting function,

$\theta = [\alpha_1, \dots, \alpha_k, \beta_1, \dots, \beta_{r+1}, \tau]^T$ is the parameter vector. In this study, the sub-optimal Matlab function algorithm given in [21] is used to reduce the high order transfer functions to the models of desired orders.

VI. SIMULATION RESULTS

In this section, two examples are given to demonstrate the decentralized controller design by using the coefficient diagram method for FOTITO systems. At first, the integer order approximate models of the FOTITO systems are obtained by using the M-SBL fitting approximation method. Then, the FOTITO systems are decoupled into two subsystems by appropriate decoupling methods. The integer order approximate models are remarkably high order models and not suitable for CDM based decentralized controller design. For this reason, the approximate models are brought into suitable forms for the design by using the mentioned model reduction method. Finally, the CDM based decentralized controller design procedure is applied.

Example 1: Let us consider an FOTITO system [17] which has the transfer function matrix

$$G(s) = \begin{bmatrix} \frac{1}{2s^{1.4} + s^{0.7} + 1} & \frac{1}{s^{1.4} + 0.5s^{0.7} + 0.5} \\ \frac{0.5}{0.5s^{1.4} + 0.5s^{0.7} + 0.4} & \frac{1}{0.5s^{1.4} + s^{0.7} + 1.2} \end{bmatrix} \quad (20)$$

M-SBL fitting approximation method is applied to obtain 4th integer order approximate models of fractional order operators s^α in the frequency range $\omega \in [10^{-1}, 10^1]$ rad/sec, and by using this approximate models, integer order approximate transfer functions of FOTITO systems are calculated as high order transfer functions. Details are given in appendix.

High order transfer functions complicate the design of the CDM based decentralized controller. Therefore, using the sub-optimal model reduction method [21], the reduced transfer function matrix elements are obtained as follows.

$$G_{r11}(s) = \frac{0.2136s^2 + 0.5841s + 0.06934}{s^3 + 1.584s^2 + 0.7807s + 0.07191} \quad (21)$$

$$G_{r12}(s) = \frac{0.4273s^2 + 1.168s + 0.1387}{s^3 + 1.584s^2 + 0.7806s + 0.07191} \quad (22)$$

$$G_{r21}(s) = \frac{0.2059s^2 + 2.764s + 0.3635}{s^3 + 4.879s^2 + 3.162s + 0.3043} \quad (23)$$

$$G_{r22}(s) = \frac{0.1995s^2 + 15.54s + 2.569}{s^3 + 20.35s^2 + 24.9s + 3.178} \quad (24)$$

Fig. 6 shows that step responses of the fractional order transfer matrix elements in Equation (20), their integer order approximate transfer functions and their reduced transfer functions match successfully. Step responses of the fractional order transfer functions are obtained by using fof function in [21]. Considering the matching of the unit step responses, the 2nd degree numerator and the 3rd degree denominator polynomials were used.

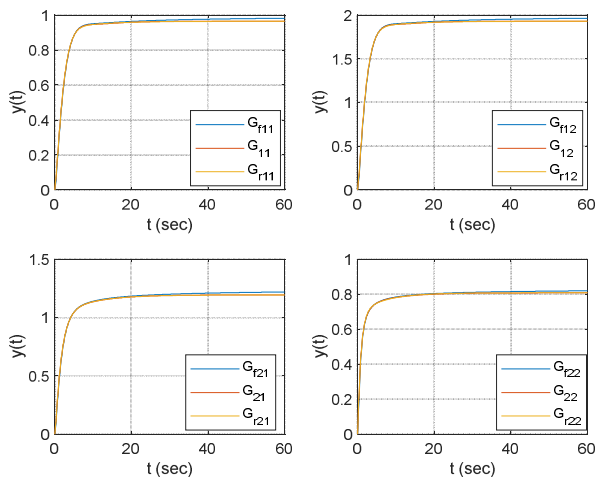


Fig. 6. The step responses of the fractional order transfer matrix elements $G_{ij}(s)$, their integer order approximate transfer functions $G_{ji}(s)$, and their reduced transfer functions $G_{rji}(s)$.

In this example, the system is decoupled with simplified decoupling to eliminate the interaction. Thus, the decoupler matrix elements of simplified decoupling, which are calculated with equation (4) by using the reduced transfer functions, are obtained as follows:

$$D_{r12}(s) = \frac{-G_{r12}(s)}{G_{r11}(s)} = \frac{-0.4273s^5 - 1.845s^4 - 2.323s^3 - 1.162s^2 - 0.1923s - 0.009972}{0.2136s^5 + 0.9225s^4 + 1.161s^3 + 0.5811s^2 + 0.09613s + 0.004986} \quad (25)$$

$$D_{r21}(s) = \frac{-G_{r21}(s)}{G_{r22}(s)} = \frac{-0.2059s^5 - 6.955s^4 - 61.74s^3 - 76.87s^2 - 17.83s - 1.155}{0.1995s^5 + 16.51s^4 + 79.01s^3 + 61.72s^2 + 12.85s + 0.7818} \quad (26)$$

The diagonal transfer matrix elements are calculated by using the approximate transfer functions of the FOTITO system in Equation (20). Since the obtained transfer functions are remarkably high order, the diagonal transfer matrix elements are redefined as follows using model reduction.

$$T_{r1}(s) = \frac{N_{r1}(s)}{D_{r1}(s)} = \frac{0.03348s^2 - 0.557s - 0.09187}{s^3 + 1.126s^2 + 0.4854s + 0.04872} \quad (27)$$

$$T_{r2}(s) = \frac{N_{r2}(s)}{D_{r2}(s)} = \frac{0.01349s^2 - 1.522s - 0.2131}{s^3 + 2.466s^2 + 1.405s + 0.1348} \quad (28)$$

After the suitable transfer functions are defined, the task of designing a CDM based decentralized controller can be performed. The numerator and denominator polynomials of the transfer functions $T_{r1}(s)$ and $T_{r2}(s)$ are given in the above equations.

According to Equation (10), polynomials $A(s)$ and $B(s)$ can be defined as follows.

$$A(s) = l_3s^3 + l_2s^2 + l_1s + l_0 \quad (29)$$

$$B(s) = k_3s^3 + k_2s^2 + k_1s + k_0 \quad (30)$$

where $l_0 = 0$ is used to provide the disturbance rejection on the system [28].

The characteristic polynomial of the CDM control system, which consists of decentralized controller parameters and the subsystem transfer function $T_{r1}(s)$, can be written as

$$P(s) = (l_3s^3 + l_2s^2 + l_1s^1)(s^3 + 1.126s^2 + 0.4854s + 0.04872) + (k_3s^3 + k_2s^2 + k_1s^1 + k_0)(0.03348s^2 - 0.557s - 0.09187) \quad (31)$$

The target characteristic polynomial, which is determined with CDM parameters $\tau = 10$, $\gamma_1 = 3.4$, $\gamma_2 = 2$, $\gamma_3 = 2$, $\gamma_4 = 2$, $\gamma_5 = 2$, is obtained as

$$P_{\text{target}}(s) = 2.149s^6 + 11.69s^5 + 31.8s^4 + 43.25s^3 + 29.41s^2 + 10s + 1 \quad (32)$$

By matching Equation (31) and Equation (32), the polynomials of the decentralized controller are obtained as follows:

$$A(s) = 2.1490s^3 + 11.6124s^2 - 18.0864s \quad (33)$$

$$B(s) = -69.9564s^3 - 95.5171s^2 - 52.4465s - 10.8849 \quad (34)$$

In addition, $F(s) = -10.8849$ can be calculated by using Equation (11). The decentralized controllers $C_{11}(s)$ and $C_{12}(s)$ given in Fig. 5a are obtained as follows:

$$C_{11}(s) = \frac{-69.9564s^3 - 95.5171s^2 - 52.4465s - 10.8849}{2.1490s^3 + 11.6124s^2 - 18.0864s} \quad (35)$$

$$C_{12}(s) = \frac{-10.8849}{-69.9564s^3 - 95.5171s^2 - 52.4465s - 10.8849} \quad (36)$$

The characteristic polynomial for the CDM control system using the subsystem $T_{r2}(s)$ is calculated as follows:

$$P(s) = (l_3s^3 + l_2s^2 + l_1s^1)(s^3 + 2.466s^2 + 1.405s + 0.1348) + (k_3s^3 + k_2s^2 + k_1s^1 + k_0)(0.01349s^2 - 1.522s - 0.2131) \quad (37)$$

Substituting the $\tau = 10$, $\gamma_1 = 4.5$, $\gamma_2 = 2.2$, $\gamma_3 = 2$, $\gamma_4 = 2$, $\gamma_5 = 2$ parameters in Equation (14), the target characteristic polynomial is obtained as follows:

$$P_{\text{target}}(s) = 0.3615s^6 + 2.863s^5 + 11.34s^4 + 22.45s^3 + 22.22s^2 + 10s + 1 \quad (38)$$

Polynomials of the decentralized controller $A(s) = 0.3615s^3 + 2.1044s^2 - 9.0065s$, $F(s) = -4.6926$ and

$B(s) = -9.8547s^3 - 26.1558s^2 - 19.1079s - 4.6926$ are computed for $T_{r2}(s)$ by using Equation (37) and Equation (38).

The decentralized controllers $C_{21}(s)$ and $C_{22}(s)$ given in Fig. 5a are obtained as follows:

$$C_{21}(s) = \frac{-9.8714s^3 - 26.1920s^2 - 19.1222s - 4.6926}{0.3615s^3 + 2.1047s^2 - 9.0291s} \quad (39)$$

$$C_{22}(s) = \frac{-4.6926}{-9.8714s^3 - 26.1920s^2 - 19.1222s - 4.6926} \quad (40)$$

A unit step function with an amplitude of 0.1 at $t = 50$ sec is given to the FOTITO system as a disturbance signal. When the FOTITO system is performed with CDM based decentralized controllers, the unit step responses are obtained as given in Fig. 7.

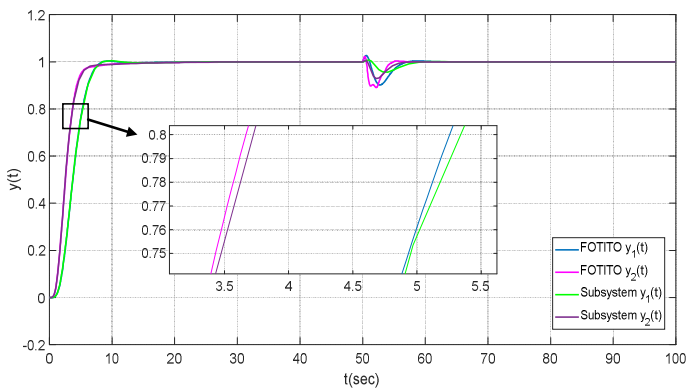


Fig. 7. The step responses of the FOTITO system with simplified decoupling

The unit step responses obtained for the subsystems $T_{r1}(s)$, $T_{r2}(s)$ and the FOTITO system are shown in Fig. 7. It can be seen from this figure that the CDM-based decentralized controllers obtained for the subsystems provide successful performance when placed in the FOTITO system using simplified decoupling.

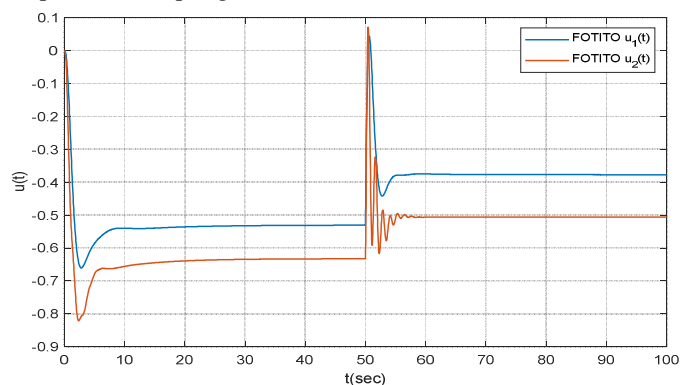


Fig 8. Control signals generated by the FOTITO system with simplified decoupling

The control signals generated by the decentralized controllers in the FOTITO system are given in Fig.8. Although the previous studies that controllers can be unstable in control systems with

CDM [22], the control signals of the FOTITO system are acceptable in terms of saturation limits.

The performance values of the step responses shown in Fig. 7 are listed in Table I. As seen in Table I and Fig. 7, the system controlled by the CDM based decentralized controllers exhibits a successful control performance in respect to the settling time and maximum overshoot. In addition, when a disturbance signal is applied to the system, the CDM based decentralized controllers provide satisfactory disturbance rejection.

TABLE I
PERFORMANCE VALUES OF THE STEP RESPONSES IN FIG. 7

PERFORMANCE VALUES	FOTITO y_1	FOTITO y_2	Subsystem y_1	Subsystem y_2
Settling Time	7.4843	7.0743	7.6272	6.3575
Maximum Overshoot (%)	0.3338	0	0.5021	0.0180

In this example, the stability index and equivalent time constant values, which are different from the values proposed by Manabe, are used to achieve better time response performance. Settling time and maximum overshoot are obtained as given in the Table I according to the selected values. Depending on the selected values, it can be seen that the ratio between t_s and τ suggested by Manabe [19] is relatively smaller. One can say that better results can be obtained depending on the selected stability index values and equivalent time constant.

Example 2: Consider the following FOTITO system with time delay as given in [15],

$$G(s) = \begin{bmatrix} \frac{1.2e^{-0.2s}}{2s^{0.5} + 1} & \frac{0.6e^{-0.3s}}{3s^{0.7} + 1} \\ \frac{0.5e^{-0.4s}}{s^{0.8} + 1} & \frac{1.5e^{-0.3s}}{3s^{0.6} + 1} \end{bmatrix} \quad (41)$$

Similar to the example 1, 4th integer order approximate models of fractional order operators s^α in the frequency range $\omega \in [10^{-1}, 10^1]$ rad/sec are obtained by using the M-SBL fitting approximation method. Then, the integer order approximate transfer functions of the FOTITO system are calculated by using the obtained 4th order approximate models (in appendix), and 1st order Pade approximation [22] is used for time delay elements. The following reduced transfer functions are obtained by applying the sub-optimal model reduction method [21] to integer order approximate transfer functions to facilitate controller design with CDM.

$$G_{r11}(s) = \frac{117.2s + 16.57}{s^4 + 4.879s^3 + 180.9s^2 + 209.1s + 16.89} \quad (42)$$

$$G_{r12}(s) = \frac{8.648s + 1.36}{s^4 + 4.155s^3 + 50.41s^2 + 31.25s + 2.519} \quad (43)$$

$$G_{r21}(s) = \frac{10.61s + 2.009}{s^4 + 5.885s^3 + 31.61s^2 + 29.91s + 4.092} \quad (44)$$

$$G_{r22}(s) = \frac{28.6s + 4.112}{s^4 + 3.859s^3 + 63.93s^2 + 45.47s + 3.289} \quad (45)$$

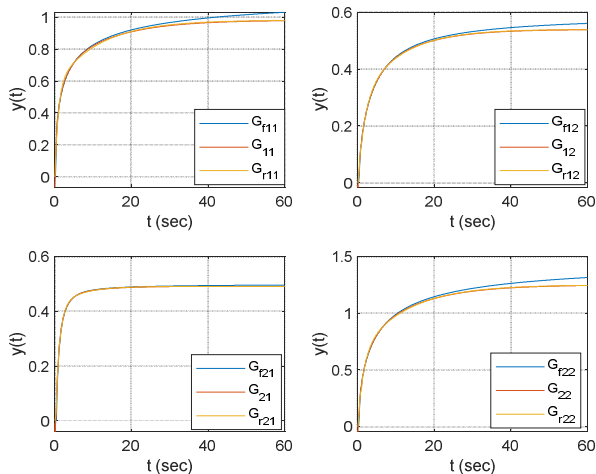


Fig. 9. The step responses of the fractional order transfer matrix elements $G_{ij}(s)$, their integer order approximate transfer functions $G_{ij}(s)$, and their reduced transfer functions $G_{rij}(s)$.

Fig. 9 shows step responses of the fractional order transfer matrix elements in Equation (41), their integer order approximate transfer functions and their reduced transfer functions. One can say that the step responses match satisfactorily.

In this example, the system is decoupled with inverted decoupling method. When reduced transfer functions are substituted in Equation (4), decoupler transfer function matrix elements are calculated as follows.

$$D_{r12}(s) = \frac{-8.648s^5 - 43.56s^4 - 1571s^3 - 2055s^2 - 430.4s - 22.97}{117.2s^5 + 503.6s^4 + 5977s^3 + 4497s^2 + 813s + 41.74} \quad (46)$$

$$D_{r21}(s) = \frac{-10.61s^5 - 42.96s^4 - 686.2s^3 - 611s^2 - 126.3s - 6.608}{28.6s^5 + 172.5s^4 + 928.3s^3 + 985.4s^2 + 240s + 16.83} \quad (47)$$

The subsystem transfer functions of the FOTITO system given in Equation (41) are obtained as follows for the inverted decoupling.

$$T_{r11}(s) = \frac{117.2s + 16.57}{s^4 + 4.879s^3 + 180.9s^2 + 209.1s + 16.89} \quad (48)$$

$$T_{r22}(s) = \frac{28.6s + 4.112}{s^4 + 3.859s^3 + 63.93s^2 + 45.47s + 3.289} \quad (49)$$

Since both $T_{r1}(s)$ and $T_{r2}(s)$ subsystems are 4th order, polynomials $A(s)$ and $B(s)$ are expressed as follows.

$$A(s) = l_4s^4 + l_3s^3 + l_2s^2 + l_1s \quad (50)$$

$$B(s) = k_4s^4 + k_3s^3 + k_2s^2 + k_1s + k_0 \quad (51)$$

CDM based decentralized controller design is applied by selecting parameters $\tau = 10$, $\gamma_1 = 6$, $\gamma_2 = 5$, $\gamma_3 = 2$, $\gamma_4 = 2$, $\gamma_5 = 2$, $\gamma_6 = 2$ and $\gamma_7 = 2$ for $T_{r1}(s)$ and $T_{r2}(s)$ transfer functions.

The parameters determined in the CDM based decentralized controller design are selected the same for both transfer functions, so their target characteristic polynomials are also the same:

$$P_{\text{target}}(s) = 6.977e - 07s^8 + 6.698e - 05s^7 + 0.003215s^6 + 0.07716s^5 + 0.9259s^4 + 5.556s^3 + 16.67s^2 + 10s + 1 \quad (52)$$

Characteristic polynomial for $T_{r1}(s)$ is obtained by substituting the determined parameters in Equation (9) as follows.

$$P_1(s) = (l_4^4 + l_3s^3 + l_2s^2 + l_1s^1) \left(\begin{matrix} s^4 + 4.879s^3 + 180.9s^2 \\ +209.1s + 16.89 \end{matrix} \right) + (k_4^4 + k_3s^3 + k_2s^2 + k_1s^1 + k_0)(117.2s + 16.57) \quad (53)$$

$A(s)$, $B(s)$ and $F(s)$ are obtained by matching the target characteristic polynomial to the characteristic polynomial of the transfer function $T_{r1}(s)$.

$$A(s) = 0.00001s^4 + 0.0001s^3 + 0.0028s^2 - 0.0744s \quad (54)$$

$$B(s) = 0.0011s^4 + 0.0064s^3 + 0.1564s^2 + 0.2525s + 0.0604 \quad (55)$$

$$F(s) = 0.0604 \quad (56)$$

The characteristic polynomial for CDM control system using $T_{r2}(s)$ is

$$P_2(s) = (l_4^4 + l_3s^3 + l_2s^2 + l_1s^1) \left(\begin{matrix} s^4 + 3.859s^3 + 63.93s^2 \\ +45.47s + 3.289 \end{matrix} \right) + (k_4^4 + k_3s^3 + k_2s^2 + k_1s^1 + k_0)(28.6s + 4.112) \quad (57)$$

By matching Equation (52) with Equation (57), $A(s)$, $B(s)$ and $F(s)$ for $T_{r2}(s)$ transfer function are obtained as follows:

$$A(s) = 0.00001s^4 + 0.0001s^3 + 0.0029s^2 - 0.3918s \quad (58)$$

$$B(s) = 0.0159s^4 + 0.0763s^3 + 1.0544s^2 + 1.0538s + 0.2432 \quad (59)$$

$$F(s) = 0.2432 \quad (60)$$

CDM based decentralized controllers shown in Fig. 5b are obtained according to the given parameters as follows:

$$C_{11}(s) = \frac{0.0011s^4 + 0.0064s^3 + 0.1564 + 0.2525s + 0.0604}{0.000001s^4 + 0.0001s^3 + 0.0028s^2 - 0.0744s} \quad (61)$$

$$C_{12}(s) = \frac{0.0604}{0.0011s^4 + 0.0064s^3 + 0.1564 + 0.2525s + 0.0604} \quad (62)$$

$$C_{21}(s) = \frac{0.0159s^4 + 0.0763s^3 + 1.0544s^2 + 1.0538s + 0.2432}{0.000001s^4 + 0.0001s^3 + 0.0029s^2 - 0.3918} \quad (63)$$

$$C_{22}(s) = \frac{0.2432}{0.0159s^4 + 0.0763s^3 + 1.0544s^2 + 1.0538s + 0.2432} \quad (64)$$

A unit step function with an amplitude of 0.1 at $t = 30$ sec is given to the system shown in Fig. 5b as a disturbance signal. When the system given in Fig. 5b is simulated for the obtained CDM based decentralized controllers, the unit step responses and the control signals given in Fig. 10 and Fig. 11 are obtained.

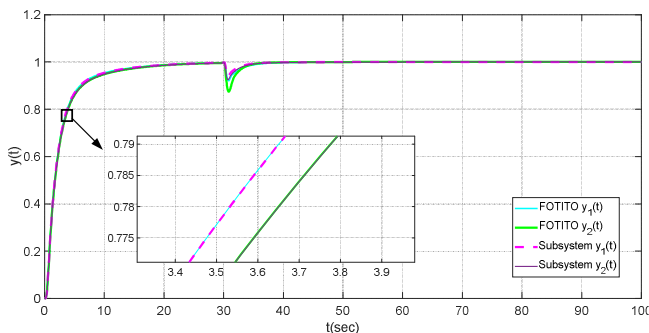


Fig. 10. The step responses of the FOTITO system with inverted decoupling.

Fig. 10 demonstrates that the CDM-based decentralized controllers obtained for subsystems exhibit good performance when placed in the FOTITO system using inverted decoupling. Similar to the example 1, one can say that the amplitudes of the control signals in Figure 11 do not affect the saturation negatively.

Performance values of the unit step responses shown in Fig. 10 are listed in Table II. The values in Table II validate that the maximum overshoot does not occur when the proposed the CDM-based decentralized controllers are used in the FOTITO system. As in the first example, the settling time is different because the equivalent time constant and stability index were selected different from those suggested by Manabe.

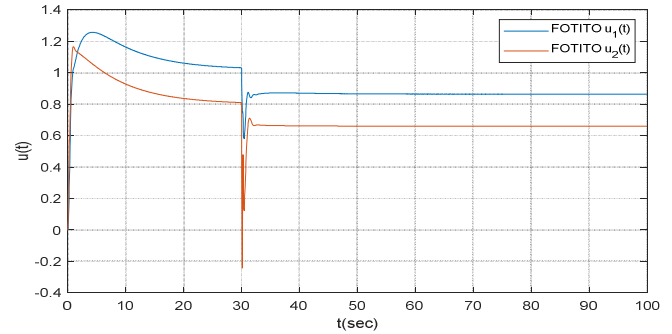


Fig 11. Control signals generated by the FOTITO system with inverted decoupling

TABLE II
PERFORMANCE VALUES OF THE STEP RESPONSES IN FIG. 10

PERFORMANCE VALUES	FOTITO y_1	FOTITO y_2	Subsystem y_1	Subsystem y_2
Settling Time	16.2669	17.2318	16.1515	17.0949
Maximum Overshoot (%)	0	0	0	0

VII. CONCLUSIONS

In this study, CDM based decentralized controllers are designed for FOTITO systems. Integer order approximate models of FOTITO systems are obtained by using the M-SBL fitting approximation method. These approximate models are decoupled into two subsystems using simplified and inverted decoupling methods. The transfer functions of the decoupled subsystems are obtained as high order. CDM-based decentralized controller design is difficult for the high-order subsystem transfer functions. Therefore, the orders of subsystem transfer functions are reduced by using the model reduction algorithm in order to facilitate the design of decentralized controller with CDM. The decentralized controllers are designed by using CDM for the transfer functions of reduced subsystems. Satisfactory results are obtained by placing the decentralized controllers, which are calculated for the subsystems, in the FOTITO systems configuration. In other words, it is concluded that CDM based decentralized controllers, which are designed for subsystems obtained by simplified and inverted decouplings, provide successful performance for FOTITO systems.

Appendix:

Integer order approximate transfer functions of the FOTITO system in Example 1 are

$$G_{11}(s) = \frac{s^8 + 99.68s^7 + 2632s^6 + 1.873e04s^5 + 5.04e04s^4 + 5.205e04s^3 + 2.174e04s^2 + 3205 + 149.9}{11.13s^9 + 909.1s^8 + 1.416e04s^7 + 7.667e04s^6 + 1.737e05s^5 + 1.898e05s^4 + 1.064e05s^3 + 3.046e04s^2 + 3700s + 155.5}$$

$$G_{12}(s) = \frac{s^8 + 99.68s^7 + 2632s^6 + 1.873e04s^5 + 5.04e04s^4 + 5.205e04s^3 + 2.174e04s^2 + 3205s + 149.9}{5.567s^9 + 454.6s^8 + 7081s^7 + 3.833e04s^6 + 8.687e04s^5 + 9.491e04s^4 + 5.32e04s^3 + 1.523e04s^2 + 1850s + 77.74} \quad [12]$$

$$G_{21}(s) = \frac{0.5s^8 + 49.84s^7 + 1316s^6 + 9363s^5 + 2.52e04s^4 + 2.602e04s^3 + 1.087e04s^2 + 1603s + 74.95}{2.784s^9 + 234.2s^8 + 3837s^7 + 2.245e04s^6 + 5.656e04s^5 + 6.923e04s^4 + 4.222e04s^3 + 1.249e04s^2 + 1516s + 62.74} \quad [13]$$

$$G_{22}(s) = \frac{s^8 + 99.68s^7 + 2632s^6 + 1.873e04s^5 + 5.04e04s^4 + 5.205e04s^3 + 2.174e04s^2 + 3205 + 149.9}{2.784s^9 + 248.4s^8 + 4482s^7 + 3.034e04s^6 + 9.218e04s^5 + 1.38e05s^4 + 9.948e04s^3 + 3.312e04s^2 + 4300s + 185.5} \quad [14]$$

and integer order approximate transfer functions of the FOTITO system in Example 2 are

$$G_{11}(s) = \frac{-1.2s^5 - 35.55s^4 + 289.3s^3 + 1744s^2 + 1179s + 107.5}{18.92s^5 + 427s^4 + 2844s^3 + 4836s^2 + 1795s + 109.6} \quad [15]$$

$$G_{12}(s) = \frac{-0.6s^5 - 36.18s^4 + 72.42s^3 + 1154s^2 + 976.3s + 107.7}{81.78s^5 + 1356s^4 + 6712s^3 + 9136s^2 + 3023s + 199.5} \quad [16]$$

$$G_{21}(s) = \frac{-0.5s^5 - 47.95s^4 - 23.64s^3 + 1150s^2 + 1122s + 136.3}{55.5s^5 + 837.9s^4 + 3905s^3 + 6078s^2 + 2858s + 277.5} \quad [17]$$

$$G_{22}(s) = \frac{-1.5s^5 - 64.89s^4 + 173.2s^3 + 1946s^2 + 1493s + 150.2}{46.07s^5 + 811.6s^4 + 4233s^3 + 6098s^2 + 2027s + 120.1} \quad [18]$$

REFERENCES

- [1] B. Halvarsson, "Interaction Analysis in Multivariable Control Systems," PhD Thesis, Uppsala University, 2010.
- [2] M. A. Üstüner, "Çok Değişkenli Sistemlerde Etkileşimin Yok Edilmesi : Proses Kontrol Sistemi Uygulanması," Master Thesis, Celal Bayar University, 2016.
- [3] E. Gagnon, A. Pomerleau, and A. Desbiens, "Simplified, ideal or inverted decoupling?," *ISA Trans.*, vol. 37, no. 4, pp. 265–276, 1998, doi: 10.1016/S0019-0578(98)00023-8.
- [4] F. Vázquez and F. Morilla, "Tuning decentralized pid controllers for MIMO systems with decouplers," *IFAC Proc. Vol.*, vol. 15, no. 1, pp. 349–354, 2002, doi: 10.3182/20020721-6-es-1901.00139.
- [5] T. Nguyen, L. Vu, and M. Lee, "Design of Extended Simplified Decoupling for Multivariable Processes with Multiple Time Delays," 2011, pp. 1822–1827.
- [6] K. Weischedel and T. J. McAvoy, "Feasibility of Decoupling in Conventionally Controlled Distillation Columns," *Ind. Eng. Chem. Fundam.*, vol. 19, no. 4, pp. 379–384, 1980, doi: 10.1021/i100005a026.
- [7] C. Rajapandayan and M. Chidambaram, "Controller design for MIMO processes based on simple decoupled equivalent transfer functions and simplified decoupler," *Ind. Eng. Chem. Res.*, vol. 51, no. 38, pp. 12398–12410, 2012, doi: 10.1021/ie301448c.
- [8] M. G. Bulut and F. N. Deniz, "Computation of Stabilizing Decentralized PI Controllers for TITO Systems with Simplified and Inverted Decoupling," *2020 7th Int. Conf. Electr. Electron. Eng. ICEEE 2020*, pp. 294–298, 2020, doi: 10.1109/ICEEE49618.2020.9102547.
- [9] M. G. Bulut and F. N. Deniz, "Computation of Stabilizing Decentralized PI Controllers for Fractional Order TITO (FOTITO) Systems," *IEEE 11th Annu. Comput. Commun. Work. Conf. CCWC 2021*, vol. 11, no. 1, pp. 1274–1280, 2021.
- [10] L. Liu, S. Tian, D. Xue, T. Zhang, Y. Q. Chen, and S. Zhang, "A Review of Industrial MIMO Decoupling Control," *Int. J. Control. Autom. Syst.*, vol. 17, no. 5, pp. 1246–1254, 2019, doi: 10.1007/s12555-018-0367-4.
- [11] A. Numsomran, T. Wongkhum, T. Suksri, P. Nilas, and J. Chaoraingern, "Design of Decoupled Controller for TITO System using Characteristic Ratio Assignment," 2007, vol. 12, no. 3, pp. 957–962.
- [12] S. Tavakoli, I. Griffin, and P. J. Fleming, "Tuning of decentralised PI (PID) controllers for TITO processes," pp. 1069–1080, 2006.
- [13] G. Kumar, "Control of TITO Process using Internal Model Control Technique," vol. 8, no. 16, pp. 87–92, 2020.
- [14] C. Hwang and Y. C. Cheng, "A numerical algorithm for stability testing of fractional delay systems," *Automatica*, vol. 42, no. 5, pp. 825–831, 2006, doi: 10.1016/j.automatica.2006.01.008.
- [15] D. Li, X. He, T. Song, and Q. Jin, "Fractional Order IMC Controller Design for Two-input-two-output Fractional Order System," *Int. J. Control. Autom. Syst.*, vol. 17, no. 4, pp. 936–947, 2019, doi: 10.1007/s12555-018-0129-3.
- [16] Z. Li and Y. Q. Chen, *Ideal, simplified and inverted decoupling of fractional order TITO processes*, vol. 19, no. 3. IFAC, 2014.
- [17] E. O. Mahdouani, M. Ben Hariz, and F. Bouani, "Design of fractional order controllers for TITO systems," *2019 Int. Conf. Signal, Control Commun. SCC 2019*, pp. 179–184, 2019, doi: 10.1109/SCC47175.2019.9116167.
- [18] A. San-Millan, D. Feliu-Talegón, V. Feliu-Batlle, and R. Rivas-Perez, "On the modelling and control of a laboratory prototype of a hydraulic canal based on a TITO fractional-order model," *Entropy*, vol. 19, no. 8, 2017, doi: 10.3390/e19080401.
- [19] S. Manabe, "Coefficient Diagram Method," *IFAC Proc. Vol.*, vol. 31, no. 21, pp. 211–222, 1998, doi: 10.1016/S1474-6670(17)41080-9.
- [20] F. N. Deniz, B. B. Alagoz, N. Tan, and M. Koseoglu, "Revisiting four approximation methods for fractional order transfer function implementations: Stability preservation, time and frequency response matching analyses," *Annu. Rev. Control*, vol. 49, pp. 239–257, 2020, doi: 10.1016/j.arcontrol.2020.03.003.
- [21] D. Xue, Y. Chen, and D. P. Atherton, *Linear Feedback Control - Analysis and Design with MATLAB*. Society for Industrial and Applied Mathematics, 2007.
- [22] H. Kayan, "Durum Zaman Gecikmeli Sistemler için Kontrol Sistem Tasarımı," Master Thesis, Inonu University, 2014.
- [23] M. Yardımcı, "Katsayı Diyagram Yönteminin (KDY) Ölü Zamanlı Sistemlere Uygulanması," Master Thesis, Istanbul Teknik University, 2005.
- [24] S. E. Hamamcı, "Zaman Gecikmeli Kararsız Sistemler için Katsayı Diyagram Metodu ile Kontrolör Tasarımı," vol. 6, no. 3, pp. 135–142, 2002.
- [25] S. E. Hamamcı, "İntegratörlü sistemler için Katsayı Diyagram Metodu ile kontrolör tasarımı," *İTÜ Derg.*, no. 422, pp. 3–12, 2004.
- [26] S. Manabe, "Application of Coefficient Diagram Method to Dual-Control-Surface Missile," *IFAC Proc. Vol.*, vol. 34, no. 15, pp. 499–504, 2001, doi: 10.1016/S1474-6670(17)40776-2.
- [27] S. Manabe, "Application of coefficient diagram method to MIMO design in aerospace," *IFAC Proc. Vol.*, vol. 15, no. 1, pp. 43–48, 2002, doi: 10.3182/20020721-6-es-1901.01233.
- [28] S. E. Hamamcı and M. Koksall, "Robust controller design for TITO processes with coefficient diagram method," *IEEE Conf. Control Appl. - Proc.*, vol. 2, no. November, pp. 1431–1436, 2003, doi: 10.1109/cca.2003.1223224.
- [29] R. Janani, I. Thirunavukkarasu, and V. S. Bhat, "Experimental implementation of cdm based two mode controller for an interacting 2 * 2 distillation process," *Int. J. Pure Appl. Math.*, vol. 118 (18), no. March, pp. 2241–2251, 2018.
- [30] C. Wuthithanyawat and S. Wangnippamto, "Decentralized PI controller with coefficient diagram method incorporating feedforward controller based on inverted decoupling for two input - Two output system," *Prz. Elektrotechniczny*, vol. 96, no. 9, pp. 159–166, 2020, doi: 10.15199/48.2020.09.33.
- [31] C. Wuthithanyawat and S. Wangnippamto, "Design of Decentralized PID Controller with Coefficient Diagram Method Based on Inverted Decoupling for TITO System," *iEECON 2018 - 6th Int. Electr. Eng. Congr.*, vol. 0, pp. 0–3, 2018, doi: 10.1109/IEECON.2018.8712222.
- [32] Q. G. Wang, B. Huang, and X. Guo, "Auto-tuning of TITO decoupling controllers from step tests," *ISA Trans.*, vol. 39, no. 4, pp. 407–418, 2000, doi: 10.1016/S0019-0578(00)00028-8.
- [33] K. V. T. Waller, "Decoupling in distillation," *AIChe Journal*, vol. 20, no. 3, pp. 592–594, 1974, doi: 10.1002/aic.690200321.
- [34] S. Fragoso, J. Garrido, F. Vázquez, and F. Morilla, "Comparative analysis of decoupling control methodologies and H ∞ multivariable robust control for variable-speed, variable-pitch wind turbines: Application to a lab-scale wind turbine," *Sustain.*, vol. 9, no. 5, 2017, doi: 10.3390/su9050713.
- [35] M. Araki and H. Taguchi, "Two-degree-of-freedom PID controllers,"

- Int. J. Control. Autom. Syst.*, vol. 1, no. 4, pp. 401–411, 2003, doi: 10.11509/isciesci.42.1_18.
- [36] N. Tan, A. Yüce, A. Özel, and F. N. Deniz, “Kesirli Dereceli Kontrol Sistemlerinde Tamsayı Dereceli Yaklaşım Metotlarının İncelenmesi,” in *Otomatik Kontrol Ulusal Toplantısı, TOK’14*, 2014, pp. 556–561.
- [37] A. Oustaloup, F. Levron, B. Mathieu, and F. M. Nanot, “Frequency-band complex noninteger differentiator: Characterization and synthesis,” *IEEE Trans. Circuits Syst. I Fundam. Theory Appl.*, vol. 47, no. 1, pp. 25–39, 2000, doi: 10.1109/81.817385.
- [38] K. Matsuda and H. Fujii, “ H_{∞} optimized wave-absorbing control: Analytical and experimental results,” *J. Guid. Control. Dyn.*, vol. 16, no. 6, pp. 1146–1153, 1993, doi: 10.2514/3.21139.
- [39] F. N. Deniz, B. B. Alagoz, N. Tan, and D. P. Atherton, “An integer order approximation method based on stability boundary locus for fractional order derivative/integrator operators,” *ISA Trans.*, vol. 62, pp. 154–163, 2016, doi: 10.1016/j.isatra.2016.01.020.
- [40] F. N. Deniz, “Kesir Dereceli Sistemlerde Modelleme ve Kontrol Uygulamaları,” PhD Thesis, Inonu University, 2017.
- [41] “Integer order approximation for fractional order derivative - File Exchange - MATLAB Central.” [Online]. Available: https://nl.mathworks.com/matlabcentral/fileexchange/87357-integer-order-approximation-for-fractional-order-derivative?s_tid=srchtitle.

BIOGRAPHIES



MİRAY GÜNAY BULUT received the B.S. degree in electrical and electronics engineering from Meliksah University, in 2016. She received the M.S in electrical and electronics engineering from Inonu University Malatya, Turkey in 2021. Her research interests are control systems applications.



FURKAN NUR DENİZ graduated from Inonu University department of Electrical and Electronics Engineering in 2008. She received her Ph.D. degree in Electrical and Electronics Engineering from Inonu University in 2017. Her research interests include fractional order control systems, modeling, and simulation.

The Use of Kalman Filter in Control The PanTilt Two-Axis Robot With Wearable System

Temel Sonmezocak


Abstract—Today, the use of micro electromechanical system (MEMS)-based control-based unmanned aerial vehicles is becoming widespread. The control of the vehicles during use can be done with joystick control connected to a fixed monitor in order to monitor the environment during flight. This situation brings with it some limitations in the activities of users to monitor the environment or direct the vehicle. However, it is important that the systems used for the vehicles control are used extremely sensitively and effectively. In this study, a 2-axis robotic pan-tilt system based on human-machine interaction and a wearable MEMS gyroscope-based headband is designed for remote control of unmanned aerial vehicles by creating an augmented reality perception on users. Therefore, it brings the perception of reality with the head movements of the users and the convenience of watching the environment during the flight. In addition, the signals produced by MEMS, the vibrations caused by electrical noise in the motors due to human interaction and environmental factors, are effectively eliminated with the Kalman filter. In this way, the images transmitted to the pilot become smoother. Therefore, it is cost-effective as it eliminates the need for additional hardware filtering structures.

Index Terms—Accelerometer, Kalman filter, micro-electromechanical, servo control, signal processing, wearable systems.

I. INTRODUCTION

MICROELECTROMECHANICAL SYSTEM (MEMS) gyroscope and accelerometers are used in today's systems as a highly developed structure in the measurement and control of motion acceleration, vibration and motor rotation angles [1-4]. In recent studies, filtering methods for the Kalman model can be used in the stabilization of MEMS accelerometer / gyroscope type systems, servo motor systems and in filtering the resulting vibrations. In the studies, balancing robots are designed to stabilize devices such as video recorders, sports cameras or smart phones during movement. In this way, the position of the camera is stabilized during movement [5,6]. For this type of systems, a

TEMEL SÖNMEZOCAK, is with Department of Electrical and Electronic Engineering, Faculty of Engineering, Dogus University, Istanbul, Turkey, (e-mail: tsonmezocak@dogus.edu.tr).

 <https://orcid.org/0000-0003-4831-9005>

Manuscript received Feb. 26, 2021; accepted April 22, 2022.
DOI: [10.17694/bajece.1079636](https://doi.org/10.17694/bajece.1079636)

new Orientation Axes Crossover Processing (OACP) algorithm based on vibration optimization by measuring vibrations with a MEMS accelerometer is developed to eliminate vibrations on the pan-tilt system. Kalman and low-pass filters are also used [7]. Similarly, for the most effective estimation of angular acceleration signals with 3-axis gyroscope, extended Kalman filter (EKF) and master-slave Kalman filter algorithm created with inverse ϕ -algorithm are used [8]. These studies, the roughness of the land, wind, etc. vibration-induced noises are prevented by using MEMS systems mounted on a pan-tilt platform. This is very important for the control of camera rotation angles in the pan-tilt system.

Unlike these studies, this study focuses on the effective use of wearable First Person View (FPV) / virtual reality (VR) type monitor glasses of a pan-tilt system on a unmanned aerial vehicle (UAV) based on human interaction. In addition, high-frequency noises caused by the sudden head position movement of the human body created by a MEMS-based 2-axis gyroscope sensor positioned on the human body, as well as electromagnetic-based noises caused by environmental factors, are effectively filtered using the Kalman filter without the need for an external hardware filtering. Therefore, the vibrations of the servo motors in the X and Y axis in the pan-tilt system are effectively eliminated. In this study, a headband wearable wireless headband is designed. With the pan-tilt camera system mounted on the UAVs, an augmented reality perception is created for the pilot by directing the wireless camera with the head movements of the pilot during the flight. This designed system does not require additional active filter hardware resources. Considering the wireless communication, the effectiveness of the Kalman filter in terms of speed performance due to human interaction only with the process noise parameter is observed in the experiments.

II. MATERIALS AND METHODS

A. Signal Acquisition

In this study, the X and Y axes of the MPU6050 system were used together as a MEMS accelerometer/gyroscope. MPU6050 can be used both as a 3-axis gyroscope for angular velocity measurement and as an accelerometer for linear acceleration measurement [9]. This chip has a 16bit A/D analog converter [10]. The angles of the MEMS system on the wearable platform are ρ for the X axis (Roll) and φ for the Y axis (Pitch), Eq. It can be determined in Eq. (1) and (2) [6].

$$\rho = \tan^{-1} \left(\frac{a_x}{\sqrt{a_y^2 + a_z^2}} \right) \quad (1)$$

$$\varphi = \tan^{-1} \left(\frac{a_y}{\sqrt{a_x^2 + a_z^2}} \right) \quad (2)$$

where, a_x , a_y and a_z , are the X, Y and Z axis components of the gravitational acceleration. In addition, the angular rotation speed of the axes (ω) is proportional to the capacitance changes of the gyro. Therefore, MEMS accelerometer/gyro systems work according to the capacitive effect [11,12]. The change of ω caused by this effect is expressed in Eq. (3).

$$\omega = \frac{dc}{dt} \quad (3)$$

where ω is proportional to velocity of slow change capacity according to time t . Therefore, when we rearrange all this information, the time-dependent changes of the angles in the X and Y axes are calculated as shown in Eq. (4) and Eq. (5).

$$p(t) = p(0) + \int_0^t \omega_x(\tau) d\tau \quad (4)$$

$$\varphi(t) = \varphi(0) + \int_0^t \omega_y(\tau) d\tau \quad (5)$$

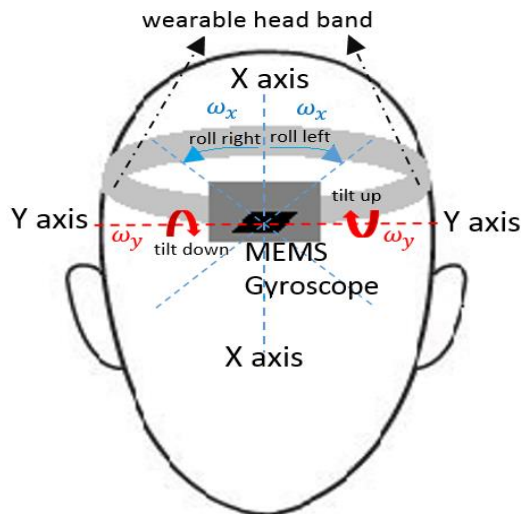


Fig.1. MEMS axes positioned on the wearable head

Fig. 1 shows the positioning of the MEMS accelerometer

axes placed on the wearable headband. The data obtained from both sensors creates involuntary noises due to their sudden changes. These noises cause sudden changes or vibrations in servo motors.

In this study, two 8-bit 20MHz Atmega328 microcontrollers are used in both the receiver and transmitter circuit structure for the processing of angular data obtained by MEMS. Data from the MEMS gyroscope sensor is sampled over the Inter-Integrated Circuit (I2C) serial port for 77 samples in 1 second (sample frequency 77Hz). NRF24L01 2.4 GHz transceiver module is used directly to transmit these sampled signals [13]. In this way, signals are transmitted to the pan-tilt system wirelessly. However, the data sent to the pan-tilt system consists of raw noisy data. These transmitted data are received by the wireless NRF24L01 2.4 GHz transceiver and converted into pulse with modulation (PWM) signals before being transmitted to the pan-tilt system in the receiving part. These translated signals are transmitted to the 8-bit 20MHz Atmega328 microcontroller built in the receiver part with the serial peripheral interface (SPI) protocol for processing by converting the signals into angle information in digital environment. Noise is also raw signals, as these received signals are carried by MEMS with the sudden change of body movements and over long distances. Therefore, the filtering of these signals is done on the receiving system. The Kalman model was used to filter these signals. In this filter, the process noise w_k parameter is used effectively in the filtering of the signals in this study before they are transmitted to the servo motors in the pan-tilt system in the receiving part. In addition, in terms of the performance of the system, the noisy signals are digitized raw and filtered with different w_k parameters and transferred to the PC environment for real-time analysis. MATLAB software was used for analysis. The experimental setups and block diagrams of these studies are shown in Fig. 2 and Fig. 3.

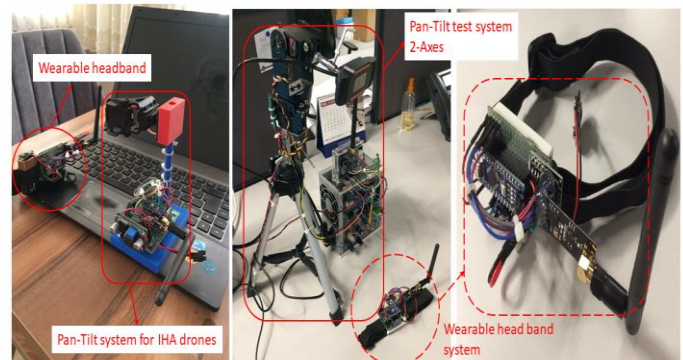


Fig.2. Experimental setups of the system

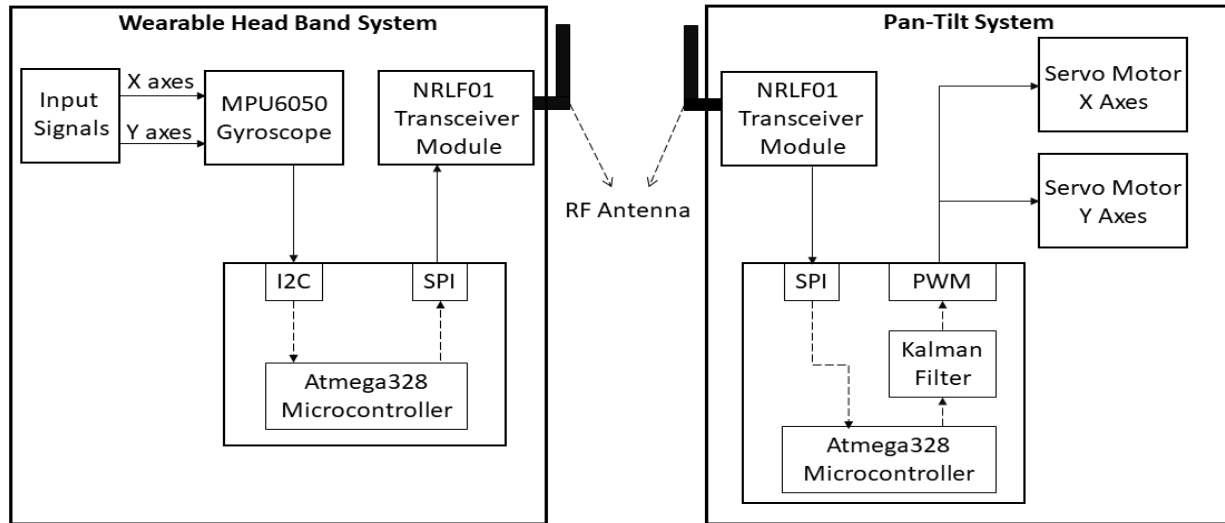


Fig.3. Block diagram of the designed system

B. Kalman Filter

Kalman filter (KF) was introduced by Rudolf Kalman. This dynamic model is being used in position tracking systems of space vehicles, robotic control systems, and biomedical signal processing systems [14,15]. In this method, the equation of a nonlinear signal discretized at each moment k and the measurement equation are modeled as in Eq. (6) and Eq. (7).

$$X_k = AX_{k-1} + w_k \quad (6)$$

$$Z_k = HX_k + v_k \quad (7)$$

In these equations, X_k is the value of signal X at moment index of k , Z_k is the measured value of signal X at moment index of k . w_k and v_k values are respectively process noise and measurement noise of the signal at moment index of k . These values are expressed as Q_k and R_k covariance matrices. It is easy to find parameter R_k from here; however, it is not easy to calculate parameter Q_k . Error covariance P_k is calculated from this parameter and is updated in a continuous loop according to the previous error covariance value \hat{P}_k^- and Kalman gain (K_k) at moment index of k in Eq. (9). Updating of the error covariance is expressed in Eq. (10) [6,16].

$$P_k^- = AP_{k-1}^- A^T + Q_k \quad (8)$$

$$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1} \quad (9)$$

$$P_k = (1 - K_k H) P_k^- \quad (10)$$

$$\hat{X}_k = \hat{X}_k^- + K_k (Z_k - HX_k^-) \quad (11)$$

The estimated output signal values are updated according to Eq. (11) by updating the covariance values for each moment index of k with the predicted error covariance expressed in Eq. (8) and K_k parameter. This updated signal represents the new filtered signal for each moment index of k .

III. EXPERIMENTAL RESULTS AND DISCUSSIONS

In this study, the time-dependent changes of the angular values of real-time signals sent to the pan-tilt system, both unfiltered and filtered, are examined comparatively. For this reason, 10 s data is used in the experiments with the user's head up/down movements. In order to test the filter performance in experimental studies, v_k parameters were kept constant in each filtering process, and experiments were made with different w_k parameters.

When the graphs in Fig. 4(a), (b) and (c) are examined separately, the existing noises can be eliminated with different w_k parameters by keeping the v_k values constant. Especially, as seen in the Fig. 4(a) and (b) graphs, filtering can be effective when the w_k parameter takes values between 0.1 and 0.01. In addition, as can be seen in Fig. 4(a), the 0.1 value of w_k , the temporal shift is 0.1 s, which is too small to be felt by the human eye. The angle values of the servo motors in the pan-tilt system remain almost the same, with a maximum deviation of 3 degrees depending on the angle value in the head movements of the user and the sudden change of direction of the head. In field experiments conducted by different people, this situation seems to be suitable for comfortable viewing performance from the camera, especially on external FPV/VR glasses screens of UAV users. On the other hand, with w_k value of 0.001, the noise in the signals is completely eliminated. The time shift in the filtered signals is in the order of 1 s compared to the original signal. Angular amplitude values in servo motors may shift about 40 degrees. This situation prevents the desired control of the pan-tilt system with human head movements.

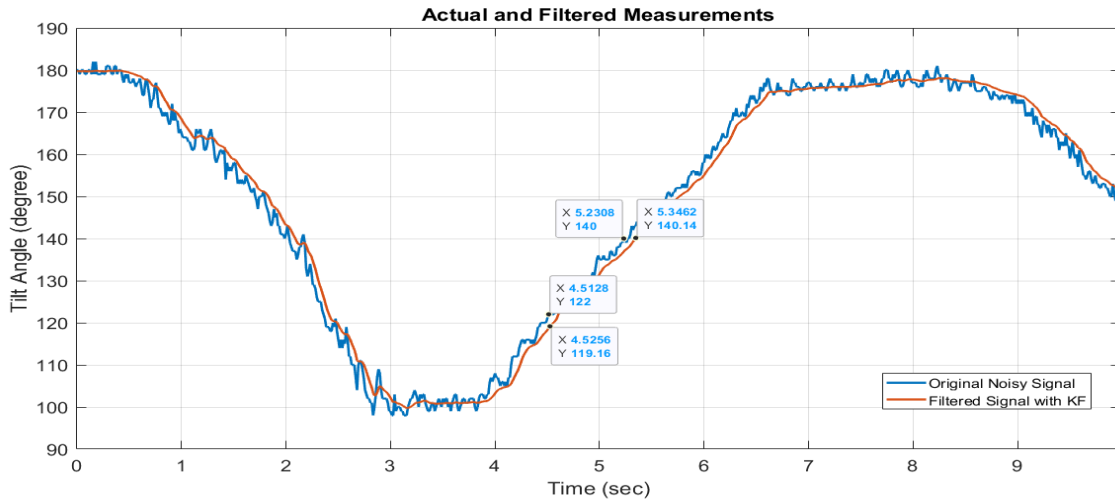


Fig.4(a). Time dependent variation of angular changes of signals filtered with process noise ($w_k: 0.1$)

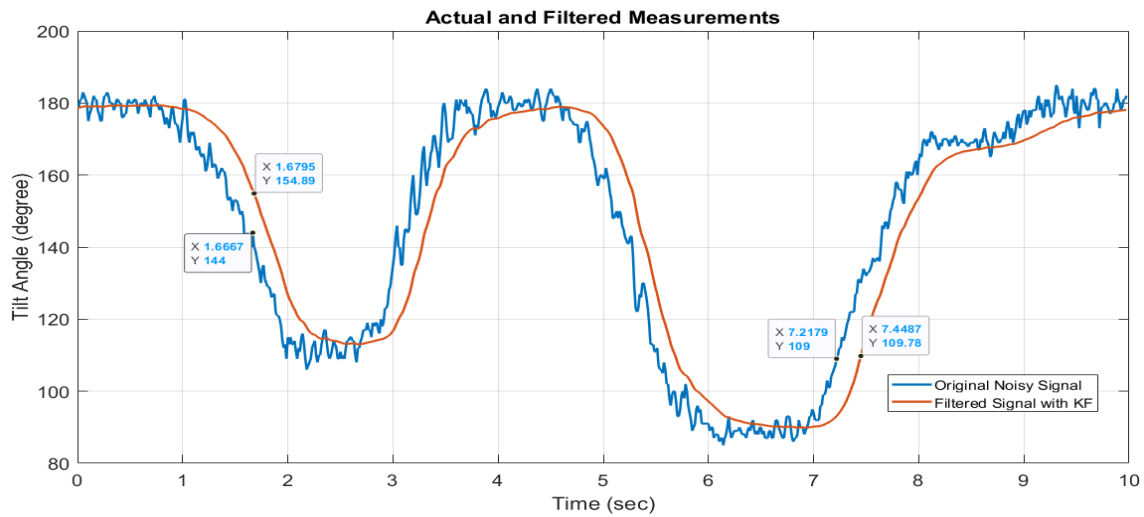


Fig.4(b). Time dependent variation of angular changes of signals filtered with process noise ($w_k: 0.01$)

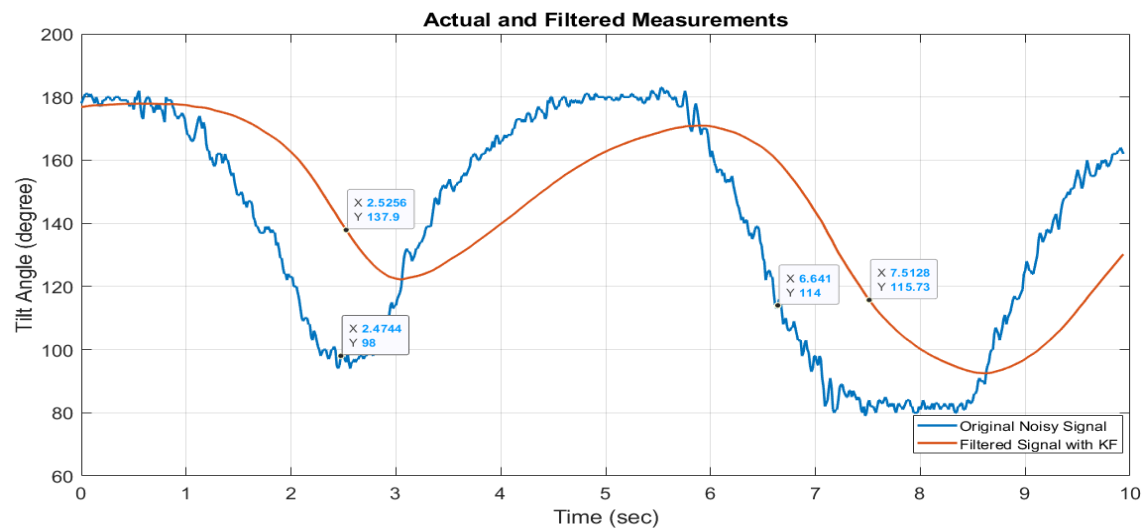


Fig.4(c). Time dependent variation of angular changes of signals filtered with process noise ($w_k: 0.001$)

As a result, time delays in the filtered signals and angular deviations in pan-tilt servos occur by decreasing the w_k parameter in software while filtering. On the other hand, by increasing the same parameter, this time the shifts in time are minimal and the angular amplitudes are the minimum deviations. When using this method, process noise parameters are used in the range of 0 to 1. The parameter values used in optimum filtering ranges between 0.1 to 0.01. In this way, there is no need for external hardware filtering structures, but it is a low-cost system, especially for human-interactive remote-control systems connected to FPV/VR wearable glasses systems.

IV. CONCLUSION

In this study, a headband working with human body interactive control is designed for the effective use of wearable FPV/VR type monitor glasses of a pan-tilt system on the UAV. Noisy signals from MEMS gyroscope detection and environmental factors are effectively filtered based on Kalman filter. In this way, the vibrations created by the electrical signals transmitted to the motors are eliminated. In addition, the effect of the process noise parameter used in filtering on the filtering performance was investigated. By using only this parameter, remote control is provided quickly and without delay, and it is cost-effective since there is no need for an external hardware filter structure.

REFERENCES

- [1] A. Harindranath, M. Arora. "MEMS IMU sensor orientation algorithms-comparison in a simulation environment", Int. Conf. Netw. Embedded Wireless Syst. IEEE Access, 2018.
- [2] P. Schopp, H. Graf, W. Burgard, Y. Manoli. "Self-calibration of accelerometer arrays", IEEE Trans. Instrum. Meas., vol. 65, 2016.
- [3] S.O. Shin, D. Kim, Y.H. Seo. "Controlling mobile robot using IMU and EMG sensor-based gesture recognition", 9th Int. Conf. Broadband Wireless Comput. Commun. & Appl., IEEE Access, China, 2014, pp.554-557.
- [4] G. Qinglei, L. Huawei, M. Shifu, H. Jian. "Design of a plane inclinometer based on MEMS accelerometer", Int. Conf. Inf. Acqui., IEEE Access, 2007, pp.320-323.
- [5] G. Li., Y. He, Y. Wei, S. Zhu, Y. Cao. "The MEMS gyro stabilized platform design based on Kalman filter", International Conference on Optoelectronics and Microelectronics (ICOM). IEEE Access, Harbin, 2013.
- [6] A. Nawrocka, M. Nawrocki, A. Kot. "The use of Kalman filter in control the balancing robot", 21st Inter. Carpathian Control Conf. (ICCC). IEEE Access, Slovakia, 2020.
- [7] M.L. Hoang, A. Petrosanto. "A new technique on vibration optimization of industrial inclinometer for MEMS accelerometer without sensor fusion", IEEE Access. vol. 9, 2021.
- [8] S. Evren, F. Yavuz, M. Unel. "High precision stabilization of Pan-Tilt systems using reliable angular acceleration feedback from a master-slave Kalman filter", J. Intell Robot syst. vol. 88, pp.97-127. 2017.
- [9] MPU6050 +-16 g Accel&Gyro, datasheet. In: Analog Devices [online]. 2022. Available at: www.analog.com.
- [10] H. Li, J. Liu, Y. Sun. "Bionic robot based on internet of things", Inter. Conf. on Information Tech. Big Data and Artificial Intelligence. IEEE Access, China, 2020.
- [11] S. Kardos, P. Balog, S. Slosarcik. "Gait dynamics sensing using IMU sensor array system", Journal of Advances in Elec. & Electronic Eng. (AEEE), vol.15, pp.71-76, 2017.

- [12] S. Kardos, S. Slosarcik, P. Balog. "Sensor array for evaluation of gait cycle. Journal of Advances in Elec. & Electronic Eng. (AEEE), vol.17, pp.459-465, 2019.
- [13] NRF24L01 2.4 GHz Transceiver, datasheet. In: Nordic Semiconductor [online]. 2022. Available at: www.alldatasheet.com.
- [14] P. Brandstetter, M. Dobrovsky. "Speed estimation motor using model reference adaptive system with Kalman filter", Journal of Advances in Elec. & Electronic Eng. (AEEE), 2013, vol.11, pp.22-28, 2013.
- [15] M. Ghanai, A. Medjghou, K. Chafaa. "Extended Kalman filter states estimation of unmanned quadrotors for altitude-attitude tracking control", Journal of Advances in Elec. & Electronic Eng. (AEEE), vol.16, pp.446-458, 2018.
- [16] Z. Peter, V. Wieser, M. Ghanai, A. Medjghou, K. Chafaa. "Radio channel state prediction by Kalman filter", Journal of Advances in Elec. & Electronic Eng. (AEEE), vol.4, pp.237-240, 2005.

BIOGRAPHIES



Temel SONMEZOCAK was born in Edirne in 1978. He graduated his B.S. degree from Electrical Education Department of Technical Education Faculty of University of Marmara, Istanbul, in 2001. In 2015, he completed his M.Sc. degree at Istanbul Aydin University, Department of Electrical and Electronics Engineering. In 2021, He

completed his PhD education in Yıldız Technical University Electronics Engineering program. He is currently working as Assistant Prof. Dr. in the Department of Electrical & Electronics Engineering at Doğuş University. His research interests include circuits, systems, biomedical, sensor networks, lighting systems, mechatronics and instrumentation.

Lidar-based Robot Detection and Positioning using Machine Learning Methods

Zahir Yilmaz and Levent Bayindir


Abstract— This paper presents a machine learning-based kin detection method for multi-robotic and swarm systems. Detecting surrounding objects and distinguishing robots from these objects (kin detection) are essential in most multi-robotic applications. While infrared, ultrasonic, and vision systems were mainly used for applying the robot detection and relative positioning task in the literature, studies using the Lidar-based approach are limited. The proposed method uses the Lidar sensor to discover the work area and determine the distance and the angle of all kin members relative to the observer robot. The main steps of the proposed method can be summarized as follows: 1) the Lidar distance points are read and stored as a vector with some preprocessing, 2) the acquired distance points representing different objects in the environment are separated from each other using a segmentation method, 3) to classify the segmented objects, the segment classification process starts with extracting five features for each object, then these features are fed to various machine learning classification algorithms to distinguish the kin robots, 4) the segments classified as a kin robot in the previous step are handled, and the relative position is found for each of them. A new mobile robot prototype has been modeled and equipped with a Lidar sensor using ROS platform. Lidar has been used to collect data, and four different classification methods have been tested to verify the efficiency of the method using Gazebo simulation platform.

Index Terms— kin detection, robot detection, relative positioning, machine learning, ROS, Gazebo


I. INTRODUCTION

IN THE PAST, the interest of scientists has been focused on finding ways to help humans in their daily life and facilitate some complex tasks they do. However, this attention has changed in the last decades to develop machines that can replace humans and replicate their actions. On the other side, while human takes days or even hours to perform a task with a possible error rate, robot, in turn, performs that task with much less time, more efficiently, and with a zero-error rate.

ZAHİR YILMAZ, is with Department of Computer Engineering, Ataturk University, Erzurum, Turkey, (e-mail: yilmaz92.tr@gmail.com).

 <https://orcid.org/0000-0002-5009-6763>

LEVENT BAYINDIR, is with Department of Computer Engineering, Ataturk University, Erzurum, Turkey, (e-mail: levent.bayindir@atauni.edu.tr).

 <https://orcid.org/0000-0001-7318-5884>

Manuscript received August 21, 2021; accepted April 28, 2022.
DOI: [10.17694/bajece.984744](https://doi.org/10.17694/bajece.984744)

Therefore, it is quite natural to witness a big revolution in robotics that is considered a branch of science that aims to design machines that can perform tasks on human commands or by themselves to make human work easier or more productive. Robotics is one of the most scientific fields that has seen significant progress in the last few decades. This progress ranges from personal assistant robots to multi-robotic and intelligent swarm robotics systems.

Robots are integrated machines usually made up of sensors, actuators, control systems, and some dedicated software and worked either depending on human commands or autonomously to perform various tasks. According to the degree of mobility of robots, there are two main types of robots: the fixed robots, which cannot move in their environment, and the mobile robots, which can travel in the environment by using various means of locomotion. However, mobile robots are more desirable nowadays due to their navigation and sensing capabilities useful for many different tasks [1, 2]. Mobile robots can be seen today in all aspects of life; autonomous vehicles, robot vacuums, assistant robot devices, search and rescue robots are some examples. Moreover, instead of working individually, mobile robots can now work in cooperation mode to accomplish critical tasks beyond a single robot's capabilities, as in the case of multi-robot and swarm robotic systems.

Swarm robotics is an innovative approach that allows large numbers of robots to collaborate and coordinate with each other to perform critical tasks that are beyond the capabilities of an individual robot. The robots used in the swarm are relatively simple and have limited abilities compared to standalone robot systems. The main idea of swarm robotics is inspired by studying the collective behavior of social animals that can cooperate and coordinate among themselves for solving everyday problems such as foraging and flocking [3]. The term "swarm" or "swarm intelligence" was first launched by Beni [4], who was interested in cellular automata systems at that time. Beni and others [5] used the term "cellular robots" to refer to a group of robots with unique characteristics that closely match the insect's swarms. After that, Beni chose the word "swarms" as a better term for cellular robots. Later, the term "swarm" evolved, and many new concepts and terms emerged from it, such as swarming, swarm optimization, "swarm engineering, and swarm robotics [6].

In the last few decades, and due to the tremendous technological progress, considerable development was achieved in swarm intelligence and multi-robot cooperation.

Swarm robotics studies have been developed primarily to address the tasks that require working according to cooperative behavior. Şahin [3] chose to classify tasks according to the nature of the task as follows:

- Tasks that cover a region such as surveillance and environmental monitoring
- Tasks that are critical and dangerous such as rescue and mine detection
- Tasks that scale up or scale down in time
- Tasks that require redundancy, such as establishing a dynamic communication network on the battlefield

Brambilla et al. [7] categorized tasks according to a slightly different perspective. They classify tasks according to behaviors where they proposed four categories of collective behaviors:

- Spatially-organizing behaviors including aggregation, pattern formation, chain formation, self-assembly and morphogenesis, and object clustering and assembling
- Navigation behaviors such as collective exploration, coordinated motion, and collective transport
- Collective decision-making such as agreement (consensus achievement) and specialization (task allocation)
- Other collective behaviors may contain some other works such as collective fault detection, group size regulation, and human–swarm interaction

However, regardless of the type of task required, in most cases, the robots in the swarm have to navigate in an environment populated by various objects. Provided that the robots are equipped with some sensors, it is crucial to distinguish the team members from the other objects and estimate their relative positions based on the sensors' data.

Detecting surrounding objects in the environment and distinguishing swarm members from these objects (often called "kin detection task") is essential for collective behaviors discussed above. It is also worthy to note that this process does not confine to detecting the robots, but it also means localization since the approximate coordinates of the robots can be recovered from the sensors data. Several studies have proposed solutions to tackle the kin detection and relative positioning task. Different sensors have been used to implement this task, such as infrared (IR), ultrasound, vision, Lidar sensors. The infrared-based approach is the most common approach for solving the robot detection and relative positioning task. A major advantage of using Lidars for robot detection and positioning is outdoor efficiency. Unlike IR-based systems, Lidars can be designed to work in outdoor environments more efficiently.

One of the IR-based robot detection and relative positioning systems was constructed by Kelly and Martinoli [15]. The proposed IR system can reveal the distance and orientation of each robot in the multi-robot system. Each robot in the team

was equipped with twelve IR Light Emitting Diodes (LEDs) and four photodiodes controlled by a PIC microcontroller. LEDs have been distributed around the robot's perimeter in a manner that ensures as much coverage as possible of the surroundings. The photodiode receivers have also been fixed at 90 degrees from each other to be able to sense signals in all directions. The distance from each robot to another and the bearing can be determined by comparing all four receivers' received signal strengths. The flocking task has been implemented using eight robots to test the method. Since the IR-based approach is the most method that has been used in literature, there are a lot of robot platforms that use this approach to tackle kin detection, relative positioning, and other related tasks [16-23].

As an example of using non-IR-based systems, Bolla et al. [8] developed a visual kin recognition and localization method that can detect and identify kin robots in a swarm robotic system. The main idea of the suggested method depends on using the Fast Fourier Transform (FFT) to extract the peak in the FFT spectrum related to the zebra pattern used to distinguish robots in the swarm. Moreover, the proposed method is not only able to detect the kin robots but can also estimate their distances and to make identification among them too. Other studies that have used the vision-based approach to achieve the detection and positioning problem for multi-robot systems are [9-11].

The relative positioning system, which uses ultrasonic sensors, developed by Rivard et al. [12], is another example of a non-IR-based system. The proposed system consists of one ultrasonic transmitter, three receivers, and one RF communication link. The detection method depends on evaluating the time-of-flight of ultrasonic pulses from the transmitter to the receiver that has revealed the signal first. The role of the RF link is to determine the time-of-flight of ultrasound pulses as it spreads much faster than ultrasound pulses. Other studies that use ultrasonic sensors are [13] and [14].

Although Lidars have been heavily used in robotics [24-26] and autonomous vehicles [27-29], Lidar-based kin detection and relative positioning methods are very limited in the literature, and most of these studies do not solve the entire problem. Premebida et al. [30] have presented some algorithms to perform segmentation and feature extraction based on data acquired from the Lidar sensor. For data segmentation, which aims to group segments into sets to distinguish the different objects in the environment, two different methods have been suggested: Point-Distance-Based Segmentation Methods (PDBS) and Kalman Filter-Based Segmentation Methods (KFBS). While in the feature extraction process, three different geometric primitives (lines, circles, and ellipses) have been used to formulate the extraction and fitting problem since these geometric shapes can be described as cases of conics. Teixidó et al. [31] have presented a proposal for circular marker detection and external mobile robot tracking. The proposed work uses an external fixed 2D Lidar to detect cylindrical objects attached to a mobile robot. Circle fitting based on the least-squares method with an algorithm for outlier avoidance

has been used in the cylindrical target detection. Several experiments have been accomplished to evaluate the positioning error obtained in the center estimation process of the cylindrical targets, where different distances, orientations, and target diameters have been used. In [32], two laser range finder sensors have been used to obtain the full field of view for a mobile robot. The proposed method lets a robot estimate a distance (range) and an angle (bearing) to another robot in the environment using measurements extracted from the raw data provided by the Lidar. The method uses circle fitting to find circular objects and filter out outliers in the set of detected circular objects. Kalman filter has also been used to improve the estimate of the relative robot position. Four distinct scenarios have been tested using MBot robots to determine the detection and tracking method's performance. In [33], Zhou et al. have chosen to use machine learning methods to implement the object detection task (similar to the kin detection task). Like others, they have used the Lidar to distinguish circular and straight objects from other targets. The proposed method depends on three main stages. The first stage is the segmentation process, where authors have proposed an improvement to the segmentation approach in [34] and called it the Improved Dietmayer method (IDIET). In the second stage, five different features have been defined for each. Finally, the Support Vector Machine (SVM) machine learning method has been used to detect the target circular objects. Another Lidar-based kin detection study was performed by us recently [35]. This study developed a geometric approach for kin detection and used the same simulation model and the Lidar for the proposed method.

This paper is organized as follows. Section 2 presents an overview of ROS, Gazebo, and URDF files used to simulate and model our robot. Section 3 introduces the proposed robot detection and positioning. Finally, we end this paper with the experiments and the experimental results.

II. ROBOT MODELING AND SIMULATION

In this study, a mobile robot model has been designed and simulated in the Gazebo simulation environment. The simulated robot has been equipped with a Lidar to detect different objects in the environment. The sensor data has been obtained using the ROS topics communication method. Then this raw data has been processed in ROS and utilized in the kin detection and relative positioning system. ROS, Gazebo, and the robot model are briefly explained in the following sections subsequently.

A. ROS

The ROS stands for Robot Operating System, an open-source robot software platform that provides services expected from any operating system like hardware abstraction, device control, communication between processes, and file management. It also offers various tools and libraries which give the ability to build, write, debug, and run code across different workstations [36]. ROS is not a real operating system in the conventional sense such as Windows, Linux, and Mac, but it is a meta-

operating system that runs over the installed operating system, and it can perform processes such as scheduling, data transmission/reception, loading, monitoring, and error handling by utilizing virtualization layer between applications and distributed computing resources. ROS has been built in small software modules called "nodes." Each node can be described as one executable program that can run independently and communicate with other nodes to send and receive data by establishing peer-to-peer links. Nodes exchange data using "messages" that could be either primitive or composed. ROS provides three mechanisms for communication between these nodes provided by passing messages: topics, services, and actions. Topics is an asynchronous unidirectional message transmission/reception method used to exchange data continuously. Services is a bidirectional synchronous communication method that depends on request/reply messages and is often used to command a robot to perform a specific action. Action is very similar to the service method, but it contains feedback messages that periodically report task states to the client.

B. Gazebo

Gazebo is one of the most popular simulators in the field of robotics. It has an integrated development environment that provides robots, sensors, environment models for 3D simulation and offers realistic simulation with its physics engine [37]. It has been selected as the official simulator of the DARPA Robotics Challenge in the US due to its high performance and having various plugins for different robot and sensor types. Using these plugins, we can read sensors data from Gazebo or send commands to motors by using APIs. Gazebo is developed and distributed by Open Robotics, which is in charge of ROS and its community, so it is compatible with ROS.

C. Robot description

In the world of robot simulation, it is so common to have the ability to model your robot, handle various sensor data received from this robot, control the robot by actuators, and test or evaluate algorithms. In ROS, robot models could be described in an XML format called Unified Robot Description Format (URDF). URDF files describe the physical configuration of the robot, such as how many wheels it has, where they are placed, and which directions they turn in. This format is designed to represent a wide variety of robots, from a two-wheeled toy to a walking humanoid. Regardless of the complexity of the robot, there are two essential elements to model any robot: links and joints. Links are the rigid parts of the robot, such as a chassis or a wheel, while joints work to connect these links, defining how they can move with respect to each other. In a URDF file, links are represented through `<link>` and `</link>` tags which represent one specified part of the robot's model (one link), while `<joint>` and `</joint>` tags are used to describe the type of joint between two specified links (parent and child links) [30]. Additionally, each link or joint can have some subtags that can define the characteristics of this link or joint. For instance, the link mass and inertia over different axes could be determined

inside the link tags. Fig. 1 shows one link and one joint of the modeled robot as an example of these tags.

```

<link name="base_link">
  <visual>
    <geometry>
      <cyylinder length="0.046" radius="0.1" />
    </geometry>
    <material name="silver" />
  </visual>
  <collision>
    <geometry>
      <cyylinder length="0.046" radius="0.1" />
    </geometry>
    <material name="silver" />
  </collision>
  <inertial>
    <mass value="1.0"/>
    <inertia ixx="0.015" iyy="0.0375" izz="0.0375"
            ixy="0" ixz="0" iyz="0"/>
  </inertial>
</link>

<joint name="middle1_link_joint" type="fixed">
  <parent link="base_link"/>
  <child link="middle1_link"/>
  <origin xyz="0 0 0.024"/>
</joint>

```

Fig.1. A portion of the URDF file containing one link (base_link) and one joint (middle1_link_joint) of the modeled robot.

The essential components (links and joints) for our robot model and the visual state of these components as a result of interpreting the URDF file using the RViz tool (a 3D visualization tool of ROS) are shown in Fig. 2.

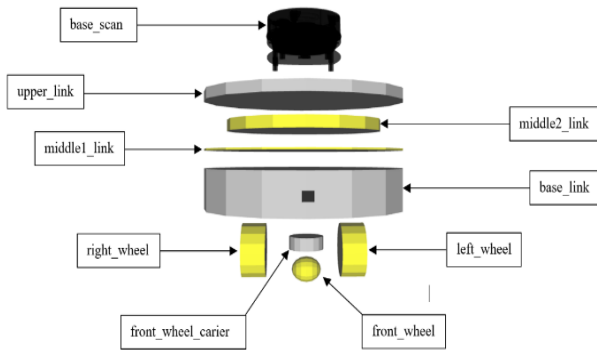


Fig.2. The URDF links that have been used to model our robot in ROS

After building the kinematics model of the robot in the previous stage, the model has to have some sensors to interact with the environment and solve our robot detection and relative positioning problem. The robot model has equipped with a differential driver and a Lidar sensor. The Lidar used to acquire surrounding data is the RPLidar A1, a 2D laser that can scan area ranges from 0.15cm up to 12m in all directions (360° of angular range) [39]. The RPLidar A1 runs clockwise to perform a 360° scan within a 12-meter range. The system measures distance data in more than 8000 times per second and with high-resolution distance output ($<1\%$ of the distance) and (≤ 1 of the angle). In this study, we have used a gazebo plugin related to

the laser sensor, where we have adjusted the plugin parameters to meet the specifications of the RPLidar sensor.

III. KIN DETECTION AND POSITIONING METHOD

The proposed method uses the Lidar sensor to collect the dataset and handle the kin detection and relative positioning task. As mentioned above, the Lidar gives 360 distance points (d_i , $0 \leq i \leq 359$) starting from the front of the sensor that corresponds to orientation 0° and goes clockwise to cover a 360° field of view.

Lidar-based studies commonly use circular robots or any circular object attached to the robot (as in our case where we use a circular-shaped Lidar) and take advantage of the circularity of these robots or objects attached to them for implementing robot detection. The main idea of these studies is to find segments from Lidar measurements corresponding to objects in the environment and apply geometric or machine learning methods to find member robots. As Fig. 3 shows, our method implements kin detection task by applying the following steps: (1) acquisition of laser data and preprocessing (2) segmentation of data using the point-distance-based segmentation method, (3) extracting five features for each segment and applying the machine learning methods to classify segment and distinguish the kin robots from them, and (4) finding the relative position of the segments classified as a kin robot in the previous step.

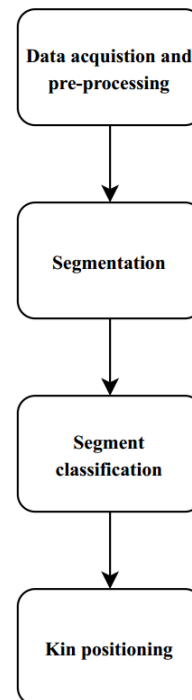


Fig.3. Flowchart of the kin detection method proposed in this paper

There are two assumptions for the proposed method.

- **Known radius:** The radius of the circular object/robot is known. In our case, the radius of Lidar RPLidar A1 is 0.035m.

- **No occlusion:** Each robot in the swarm is completely visible to the observer robot.

A. Data acquisition and preprocessing

As mentioned in the previous section, the Lidar being used measures 360 distance points. Each of the raw laser points is represented in the polar coordinate system as $\{(d_i, \theta_i); 0 \leq i \leq 359\}$, where d_i is the distance measured from the center of the observer robot to the object and θ_i the relative angle of the measurement (see Fig. 4). First, the acquired Lidar data are stored as vectors (d_i, θ_i) .

There is a difference between the simulated Lidar model and the real Lidar. In the real world, the Lidar returns the maximum range value for objects outside its operating range (which means that there is no obstacle against the laser beams). However, for the laser sensor plugin that we used, the simulated Lidar gives an infinity value for such a case. So, to tackle that, we convert the infinity values obtained from the laser sensor plugin to the Lidar max range value (d_{max}). In the same way, any object located at (d_{max}) from the observer robot will not be visible to the Lidar.

It is also possible to apply some filtering to remove noise from the Lidar data in this stage. However, we did not apply any filtering in this study because we did not add noise to the simulated Lidar data. Fig. 4 represents the scanning system of the Lidar used, where one object has been placed at a distance (d_i) and an angle (θ_i).

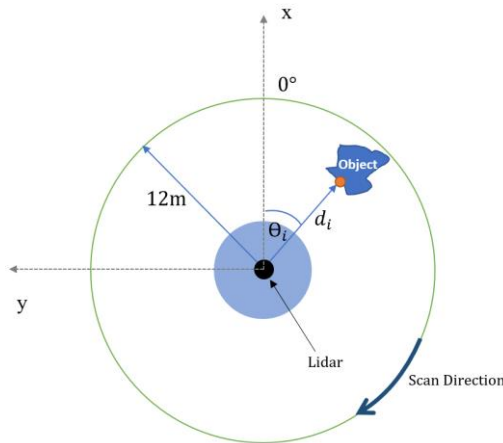


Fig.4. RPLidar A1 scanning system, where (d_i, θ_i) are the distance and the angle of the object relative to Lidar (Adapted from [30])

B. Segmentation

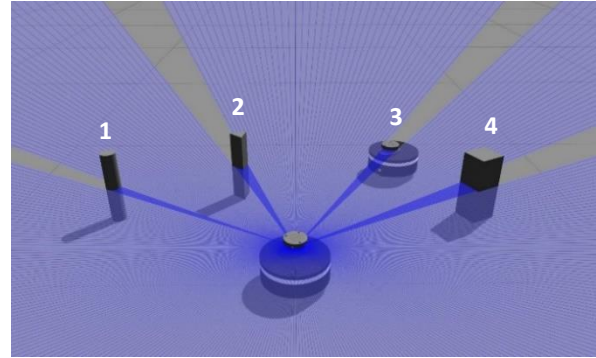
Segmentation is the process of transforming the raw laser points into groups of segments (useful data), which could be robots, humans, or other things. Segments can be defined as a set of range measurements (points) close to each other and probably belonging to one object. In this work, a PDBS based method is used for segmentation [30].

Segments are detected by using the derivative (∂_i) which can be calculated by finding the difference between each distance point (d_i) and the one before it (d_{i-1}). Then, small derivative values are filtered using a threshold value (d_{th}) which allows ignoring the small changes in scan data:

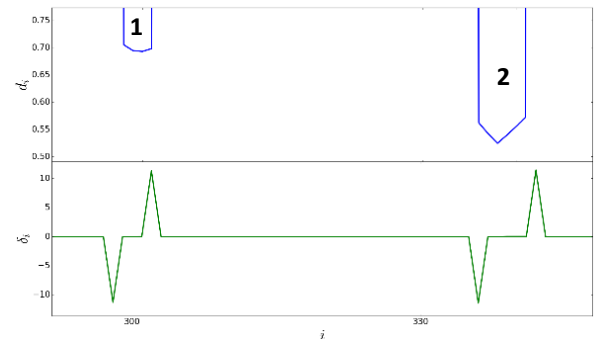
$$\partial_i = \{d_i - d_{i-1} \text{ if } |d_i - d_{i-1}| > d_{th}, 0 \text{ otherwise}\} \quad (1)$$

$$1 \leq i \leq 359$$

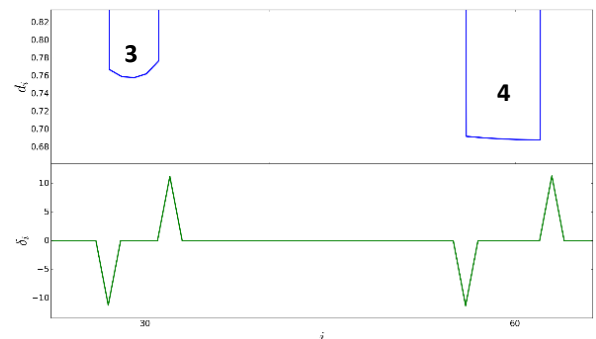
As a result of this operation, for each object in the environment, a falling slope ($\partial_i < -d_{th} < 0$) and a rising slope ($0 < d_{th} < \partial_i$) is detected (See Fig. 5 B and 5 C). By associating and combining each falling and rising edge, we obtain probable segments representing an object in the environment. Fig. 5 illustrates this step for four different objects (cube, kin, triangular prism, and cylinder) in the Gazebo simulator.



A. The observer robot and four different objects in the Gazebo simulator. The objects (from left to right) are: cylinder (1), triangular prism (2), kin robot (3), and a cube (4)



B. Representation of the Lidar scan data (blue lines) and the differences between each scan point and the previous one according to the ray number (green lines) for the two objects at the left of Fig. 5. A (object 1 and 2)



C. Representation of the Lidar scan data (blue lines) and the differences between each scan point and the previous one according to the ray number (green lines) for the two objects at the right of Fig. 5. A (object 3 and 4)

Fig.5. A representation of the segmentation process using four different objects in Gazebo simulator

C. Segment classification

Classification is the act or process of dividing things into groups according to their type (kin robot or not). Segment classification can be implemented using geometric [30] or machine learning [33] methods. In this study, the segment classification process has been performed using machine learning methods.

In machine learning, classification is a supervised learning approach. The computer program develops a model from the sample data (data set) and then uses this model to classify new observations. This data set may be bi-class (like our case where we attempt to identify whether the object is a robot or not) or multi-class.

This step aims to classify segments to distinguish the kin robots from the other objects. The classification process starts with collecting the segments' raw data and then converting these data to features. Five features have been extracted for each segment in the classification stage. Then various machine learning classification algorithms have been applied over these features, and results have been compared according to several performance measures (as described in the next section).

The features used are fitness, symmetry, radius error, circularity, and straightness, denoted by F_{fit} , F_{sym} , F_{rer} , F_{cir} and F_{str} respectively. These features have been suggested by [33], who have used these features to detect circular objects in polar coordinates. They have chosen SVM for classification, and just two object types (plastic cast and rectangle paper box) have been used to test the algorithm. However, this paper suggests using these features to distinguish kin robots from other possible obstacles and find their relative positions. We have tested several situations and scenarios for various objects using four different classification algorithms. The features used have been defined as follows:

1) Fitness

The fitness function evaluates how close a given solution is to the optimum solution of the desired problem. It determines how much a solution is fit. In our case, we can benefit from the fitness function to measure how the distance points related to a segment (object) fit the circular object representing the kin member that we are looking for by using the following equation.

$$F_{fit} = \frac{1}{\max(\theta_q) - \min(\theta_q)} \sum_{i=1}^N |\rho_q(i) - \rho(i)| \quad (2)$$

Where $(\rho(i), \theta(i))$ are the corresponding polar coordinates for scan distance points, and N is the number of these distance points. This feature could be used to extract circles and lines. If the target object is a circle or line, the corresponding F_{fit} is small. Otherwise, F_{fit} is a large number.

2) Circularity

The circularity feature describes how close an object is to a true circle. So, we can use this feature to show how an object is close to a circle with a known radius by using the following equations:

$$\alpha_{ct} = \sin^{-1} \frac{R_r}{\rho_{min} + R_r} \quad (3)$$

$$F_{cir} = |\alpha_{ct} - \alpha| \quad (4)$$

R_r is the radius of the Lidar used and ρ_{min} represents the minimal distance between Lidar and the object. While α represents the measured value of the angle between the maximal and the minimal distance between LRF and target object, α_{ct} denotes the theoretical calculation value of that angle. So, the difference of α_{ct} and α defines the circularity feature.

3) Radius error

For circle, we can calculate the difference between the theoretical calculation value of the circle radius (R_{est}) and the original value (R_r) which is in our case 0.035cm to define the radius error estimation feature as follows:

$$R_{est} = \frac{\rho_{min} \sin \alpha}{1 - \sin \alpha} \quad (5)$$

$$F_{rer} = |R_r - R_{est}| \quad (6)$$

In theory, for a circle, F_{cir} and F_{rer} are equal to zero. In practice, there are measurement errors, so F_{cir} and F_{rer} are not exactly equal to zero but are still close to zero. If the detected object is not a circle, F_{cir} and F_{rer} are large values.

4) Straightness

It is possible to calculate the straightness degree of the target object using the following equations:

$$\alpha_{lt} = \cos^{-1} \frac{\rho_{min}}{\rho_{max}} \quad (7)$$

$$F_{str} = |\alpha_{lt} - \alpha| \quad (8)$$

Where ρ_{min} and ρ_{max} represent the minimal and the maximal distance between Lidar and the target object. While (α) represents the measured value of the angle between the maximal and the minimal distance between LRF and the target object, (α_{lt}) denotes the theoretical calculation value of that angle. The difference of α_{lt} and α defines the straightness feature.

For both lines and circles, F_{str} should have a small value close to zero too. For objects with other shapes, F_{str} becomes a larger value.

5) Symmetry

Mathematically, symmetry means that one shape becomes exactly like another when you move it in some way: turn, flip, or slide. For two objects to be symmetrical, they must be the same size and shape but have a different orientation from each other. Since the Lidar used to distinguish robots is a cylindrical

object, the symmetry value we will obtain from the formula described below will measure the symmetry of a segment with a cylindrical object in Lidar size.

Generally, symmetry level is measured by the Hausdorff distance. However, in this paper, F_{sym} is simply computed by:

$$F_{sym} = \left| \frac{\max(\theta_q) - \min(\theta_q)}{2} - \theta_{fit} \right| \quad (9)$$

where θ_{fit} is the axis of symmetry, and it equals to $\theta_{fit} = -\frac{p_1}{2p_2}$.

According to the equations described above, it could be noticed that features will be affected by two factors:

- Whether an object is circular, line or not
- Whether the radius of the detected circular object is equal to the kin robot's lidar radius or not.

D. Kin positioning

After classifying the segments related to kin robots, it is necessary to identify the relative position of these kins. The kin positioning process involves estimating the distance and relative angle between the observer robot and other kins. As we have the distance points of the kin robot, we can obtain the relative position of the kin robot by getting the smallest value among the distance points. Finally, the index value corresponding to the smallest value can be easily used to obtain the relative angle of the kin robot.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

Experiments are performed in three stages. In the first stage, raw data collection is performed. In order to verify the validity of the proposed method, various object types (including the kins that we are looking for) have been used in the data collection stage (see Fig. 6). Different distances, positions, and sizes have been used for each object type. The data collection process covered the distance 0.2 m (the minimum distance that the Lidar can detect) to 1 m. The collected raw data contains the index of the first and the last rays that hit the object (or segment), the distance values in that range, and a class label ("1" for kins and "0" for non-kin objects). Some samples of the raw data set are shown in Table 1.

In the second stage, the raw data set is transformed into a feature data set. The feature data set contains five feature values for each segment and is stored in a CSV file. Some samples of the feature data set are shown in Table 2.

In the last stage, the feature data set is fed to different machine learning classifiers, where we used four different classification algorithms for this study which are: random forest (RF), SVM, k-nearest neighbor (kNN), and decision tree (DT). The collected datasets contain 2230 segments (850 frames for kins and 1380 for other shapes). The feature data set was divided into training and testing sets using the 20-fold cross-validation method. Different hyperparameters have been tried for each classifier to get the best result. Table 3 shows the best parameters obtained for each classification algorithm using the

grid search method. Table 4 shows the performance measures of the classification algorithms used where the best algorithm was the random forest with an 86% accuracy score.

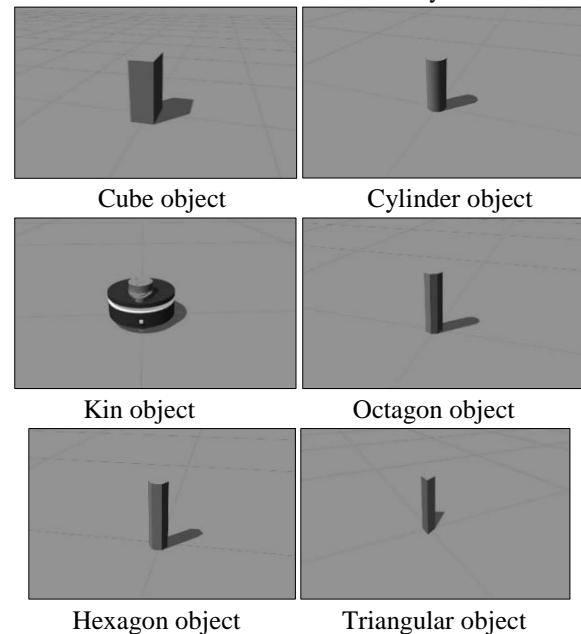


Fig.6. The objects used to collect data and extract the features in the segment classification process

TABLE I
SAMPLES OF THE GENERATED RAW DATA SET OBTAINED BY LIDAR FOR KINS AND OTHER OBJECTS. THE KIN ROBOTS WERE LABELLED WITH (1), AND NON-KINS WERE LABELLED WITH (0)

Seg No	Start index	End index	Distance Points			Kin or not	
1	173	184	0.21	0.21	0.21	...	0
2	264	270	0.43	0.42	0.41	...	0
3	239	242	0.77	0.76	0.76	...	0
4	174	183	0.38	0.37	0.37	...	0
5	201	204	1.09	1.08	1.07	...	0
6	267	272	0.70	0.69	0.68	...	0
7	155	171	0.23	0.22	0.22	...	1
8	165	171	0.59	0.58	0.57	...	1
9	121	125	0.75	0.74	0.74	...	1
10	204	208	0.81	0.79	0.79	...	0
11	216	225	0.35	0.35	0.34	...	0
12	263	267	0.89	0.87	0.87	...	0
13	174	184	0.34	0.33	0.33	...	1
14	177	181	0.88	0.86	0.86	...	1
15	262	265	0.98	0.96	0.96	...	1
16	174	183	0.29	0.29	0.29	...	0
17	270	274	0.73	0.73	0.73	...	0
18	272	286	0.22	0.22	0.22	...	0
19	143	151	0.47	0.46	0.45	...	1
20	100	104	0.75	0.75	0.74	...	1
21	340	346	0.53	0.52	0.52	...	1
22	172	182	0.32	0.32	0.32	...	0
23	175	180	0.62	0.62	0.61	...	0

TABLE II

THE FIVE FEATURES EXTRACTED FOR THE SEGMENTS SHOWN IN TABLE 1. THE FEATURES ARE FITNESS, SYMMETRY, ESTIMATE RADIUS ERROR, CIRCULARITY, AND STRAIGHTNESS DENOTED BY F_{fit} , F_{sym} , F_{rer} , F_{cir} AND F_{str}

Seg No	F_{fit}	F_{cir}	F_{rer}	F_{str}	F_{sym}	Kin or not
1	0.03	0.04	0.01	0.25	3.11	0
2	0.04	0.03	0.01	0.23	4.66	0
3	0.03	0.01	0.01	0.17	4.19	0
4	0.12	0.00	0.00	0.30	3.11	0
5	0.33	0.01	0.02	0.24	3.53	0
6	0.12	0.00	0.00	0.24	4.70	0
7	0.08	0.01	0.00	0.34	2.84	1
8	0.06	0.01	0.00	0.24	2.93	1
9	0.04	0.01	0.01	0.19	2.14	1
10	0.04	0.01	0.01	0.16	3.59	0
11	0.07	0.01	0.00	0.26	3.84	0
12	0.02	0.00	0.00	0.21	4.62	0
13	0.05	0.01	0.00	0.27	3.12	1
14	0.05	0.00	0.00	0.20	3.12	1
15	0.02	0.00	0.00	0.14	4.60	1
16	0.00	0.05	0.02	0.26	3.11	0
17	0.00	0.02	0.02	0.02	4.74	0
18	0.00	0.10	0.03	0.05	4.87	0
19	0.09	0.00	0.00	0.27	2.56	1
20	0.03	0.01	0.01	0.18	1.77	1
21	0.02	0.01	0.01	0.19	5.98	1
22	0.12	0.03	0.01	0.35	3.08	0
23	0.15	0.02	0.01	0.24	3.09	0

TABLE III

THE BEST VALUES OF HYPERPARAMETERS OBTAINED USING THE GRID SEARCH METHOD

Algorithm	Parameter	Value
Decision tree	criterion	entropy
	max depth	6
	min samples leaf	8
Random forest	criterion	entropy
	number of trees	100
KNN	number of neighbours	38
SVM	kernel function	rbf
	regularization parameter	9

This paper proposed a kin detection and positioning algorithm. Due to the lack of studies that use the Lidar sensor to implement such a task, we have chosen to use the lidar-based approach to tackle this task. Our method uses a robot model equipped with a Lidar sensor to detect the different objects in an environment, distinguish the kin robot member and find the relative position of each kin using the machine learning

classification methods. The features used in the classification process have been suggested by Zhou et al. [33] to detect circular objects using SVM. However, in this study, we have recommended using these features to distinguish kin robots equipped with Lidar sensors from other possible obstacles and then find the relative positions of each detected kin.

TABLE IV

PERFORMANCE MEASURES OBTAINED ACCORDING TO THE CLASSIFICATION METHODS USED FOR THE KIN DETECTION TASK WHERE THE RANDOM FOREST ALGORITHM GIVES THE BEST ACCURACY SCORE (86%).

	Accuracy	Precision	Recall	F1_score
DT	85%	87%	85%	84%
RF	86%	86%	86%	86%
kNN	84%	87%	84%	83%
SVM	82%	86%	82%	81%

Several situations and scenarios for various objects are tested using four different classification algorithms. The algorithms used are decision tree, random forest, k-nearest neighbors, and SVM. The random forest algorithm gave the best result for the segment classification with an 86% accuracy score.

The robot model used in the experiments is a model of an early version of the swarm robotic platform called Layka, developed by the second author. The Layka swarm robotic platform being developed is designed to be used with or without ROS. There are some advantages of using ROS for developing swarm robotic systems. First, tools and mechanisms supported by ROS can simplify debugging swarm robotic systems considerably. Second, ROS mechanisms and abstractions allow developing modular and platform-independent components. Third, ROS has many high-level libraries (such as navigation and planning) that can be valuable for swarm robotic systems. In future works, we plan to finalize the development of the Layka swarm robotic platform, test proposed methods on the real robot platform, and analyze the effect of sensor noise which is one of the limitations of this study and the previous one [35].

REFERENCES

- [1] R. S. Ortigoza, M. Marcelino-Aranda, G. S. Ortigoza, V. M. H. Guzman, M. A. Molina-Vilchis, G. Saldana-Gonzalez, J. C. Herrera-Lozada, M. Olguin-Carbajal. "Wheeled mobile robots: a review." *IEEE Latin America Transactions*, 10(6), 2012, pp2209-2217.
- [2] F. Rubio, F. Valero, C. Llopis-Albert. "A review of mobile robots: Concepts, methods, theoretical framework, and applications." *International Journal of Advanced Robotic Systems*, 16(2), 2019, 1729881419839596.
- [3] E. Şahin, "Swarm Robotics: From sources of inspiration to domains of application," in *Swarm Robotics*, E. Şahin and W. M. Spears, Eds. 2005, pp. 1–9.

- [4] G. Beni, "The concept of cellular robotic system," in *IEEE International Symposium on Intelligent Control*, 1988, pp. 57–62, DOI: 10.1109/isic.1988.65405.
- [5] T. Fukuda and S. Nakagawa, "Approach to the dynamically reconfigurable robotic system," *Journal of Intelligent and Robotic Systems*, vol. 1, no. 1, pp. 55–72, 1988, DOI: 10.1007/bf00437320.
- [6] G. Beni, "From swarm intelligence to swarm robotics," in *International Workshop on Swarm Robotics*, 2004: Springer, pp. 1-9.
- [7] M. Brambilla, E. Ferrante, M. Birattari, and M. Dorigo, "Swarm robotics: a review from the swarm engineering perspective," *Swarm Intelligence*, vol. 7, no. 1, pp. 1–41, 2013, DOI: 10.1007/s11721-012-0075-2.
- [8] K. Bolla, T. Kovacs, and G. Fazekas, "Compact Image Processing Based Kin Recognition, Distance Measurement and Identification Method in a Robot Swarm," in *International Joint Conference on Computational Cybernetics and Technical Informatics*, 2010, pp. 419–424, DOI: 10.1109/iccocyb.2010.5491237.
- [9] I. Rekleitis, G. Dudek, and E. Milios, "Multi-robot collaboration for robust exploration," *Annals of Mathematics and Artificial Intelligence*, vol. 31, no. 1–4, pp. 7–40, 2001, DOI: 10.1023/a:1016636024246.
- [10] Y. Han and H. Hahn, "Visual tracking of a moving target using active contour based SSD algorithm," *Robotics and Autonomous Systems*, vol. 53, no. 3–4, pp. 265–281, 2005, DOI: 10.1016/j.robot.2005.09.005.
- [11] K. Bolla, Z. Istenes, T. Kovacs, and G. Fazekas, "A Fast Image Processing Based Robot Identification Method for Surveyor SRV-1 Robots," in *IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*, 2011, pp. 1003–1009, DOI: 10.1109/aim.2011.6027147.
- [12] F. Rivard, J. Bisson, F. Michaud, and D. Létourneau, "Ultrasonic Relative Positioning for Multi-Robot Systems," in *IEEE International Conference on Robotics and Automation*, 2008, pp. 323–328, doi: 10.1109/robot.2008.4543228.
- [13] L. E. Navarro-Serment, C. J. Paredis, and P. K. Khosla, "A beacon system for the localization of distributed robotic teams," in *International Conference on Field and Service Robotics*, 1999, vol. 6, pp. 1–6.
- [14] C.-J. Wu and C.-C. Tsai, "Localization of an Autonomous Mobile Robot Based on Ultrasonic Sensory Information," *Journal of Intelligent and Robotic Systems*, vol. 30, no. 3, pp. 267–277, 2001, DOI: 10.1023/a:1008154910876.
- [15] I. Kelly and A. Martinoli, "A scalable, on-board localisation and communication system for indoor multi-robot experiments," *Sensor Review*, vol. Volume 24, no. Issue 2, pp. 167–180, 2004, DOI: 10.1108/02602280410525968.
- [16] G. Caprari and R. Siegwart, "Design and control of the mobile micro robot alice," in *2nd International Symposium on Autonomous Minirobots for Research and Edutainment*, 2003, pp. 23--32.
- [17] F. Arvin, K. Samsudin, and A. R. Ramli, "A Short-Range Infrared Communication for Swarm Mobile Robots," in *International conference on signal processing systems*, 2009, pp. 454–458, DOI: 10.1109/icsps.2009.88.
- [18] F. Mondada *et al.*, "The e-puck, a robot designed for education in engineering," in *9th conference on autonomous robot systems and competitions*, 2009, vol. 1, pp. 59--65.
- [19] S. Kornienko, "IR-based Communication and Perception in Microrobotic Swarms," in *7th Workshop on Collective & Swarm Robotics*, 2010.
- [20] A. E. Turgut, F. Gokce, H. Celikkanat, L. Bayindir, and E. Sahin, "Kobot: A mobile robot designed specifically for swarm robotics research," *METU-CENG-TR Tech. Rep*, vol. 5, no. 2007, Middle East Technical University, Ankara, Turkey, 2007.
- [21] F. Mondada, A. Guignard, M. Bonani, D. Bär, M. Lauria, and D. Floreano, "SWARM-BOT: From Concept to Implementation," 2003, vol. 2, pp. 1626–1631, doi: 10.1109/iro.2003.1248877.
- [22] F. Arvin, J. Murray, C. Zhang, and S. Yue, "Colias: An Autonomous Micro Robot for Swarm Robotic Applications," *International Journal of Advanced Robotic Systems*, vol. 11, no. 7, p. 113, 2014, DOI: 10.5772/58730.
- [23] J. McLurkin, A. McMullen, N. Robbins, G. Habibi, A. Becker, A. Chou, H. Li, M. John, N. Okeke, J. Rykowski, S. Kim, W. Xie, T. Vaughn, Y. Zhou, J. Shen, N. Chen, Q. Kaseman, L. Langford, J. Hunt, A. Boone, K. Koch. A robot system design for low-cost multi-robot manipulation. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2014)*, 2014.
- [24] F. B. Malavazi, R. Guyonneau, J. B. Fasquel, S. Lagrange, F. Mercier. "LiDAR-only based navigation algorithm for an autonomous agricultural robot.",

- Computers and electronics in agriculture*, 154, 2018, 71-79.
- [25] I. Belkin, A. Abramenko, D. Yudin. "Real-time lidar-based localization of mobile ground robot." *Procedia Computer Science*, 186, 2021, 440-448.
- [26] D. Zhang, J. Cao, G. Dobie, C. MacLeod. "A framework of using customized LIDAR to localize robot for nuclear reactor inspections." *IEEE Sensors Journal*, 2021.
- [27] C. Benedek, A. Majdik, B. Nagy, Z. Rozsa, T. Sziranyi. "Positioning and perception in LIDAR point clouds." *Digital Signal Processing*, 119, 2021, 103193.
- [28] S. Royo, M. Ballesta-Garcia. "An overview of lidar imaging systems for autonomous vehicles." *Applied sciences*, 9(19), 4093.
- [29] Y. Li, L. Ma, Z. Zhong, F. Liu, M. A. Chapman, D. Cao, J., Li. "Deep learning for LiDAR point clouds in autonomous driving: a review." *IEEE Transactions on Neural Networks and Learning Systems*, 32(8), 2020.
- [30] C. Premebida and U. Nunes, "Segmentation and geometric primitives extraction from 2d laser range data for mobile robot applications," *Robotica*, vol. 2005, pp. 17--25, 2005.
- [31] M. Teixidó, T. Pallejà, D. Font, M. Tresanchez, J. Moreno, and J. Palacín, "Two-Dimensional Radial Laser Scanning for Circular Marker Detection and External Mobile Robot Tracking," *Sensors*, vol. 12, no. 12, pp. 16482–16497, 2012, doi: 10.3390/s121216482.
- [32] A. Wąsik, R. Ventura, J. N. Pereira, P. U. Lima, and A. Martinoli, "Lidar-based relative position estimation and tracking for multi-robot systems," in *Robot 2015: Second Iberian Robotics Conference*, 2016, pp. 3–16, DOI: 10.1007/978-3-319-27146-0_1.
- [33] X. Zhou, Y. Wang, Q. Zhu, and Z. Miao, "Circular object detection in polar coordinates for 2D LIDAR data," in *Chinese Conference on Pattern Recognition*, 2016, pp. 65–78, DOI: 10.1007/978-981-10-3002-4_6.
- [34] K. Dietmayer, "Model-Based Object Classification and Object Tracking in Traffic Scenes from Range-Images," in *IV2001*, 2001, pp. 25--30.
- [35] Z. Yılmaz and L. Bayındır, "Simulation of Lidar-Based Robot Detection Task using ROS and Gazebo," *European Journal of Science and Technology*, pp. 513–529, 2019, doi: 10.31590/ejosat.642840.
- [36] M. Quigley *et al.*, "ROS: an open-source Robot Operating System," in *ICRA workshop on open source software*, 2009, vol. 3, p. 5.
- [37] N. Koenig and A. Howard, "Design and use Paradigms for Gazebo, An Open-Source Multi-Robot Simulator," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (IEEE Cat. No.04CH37566)*, 2004, vol. 3, pp. 2149–2154, DOI: 10.1109/iros.2004.1389727.
- [38] L. Kunze, T. Roehm, and M. Beetz, "Towards Semantic Robot Description Languages," 2011, vol. 1, pp. 5589–5595, DOI: 10.1109/icra.2011.5980170.
- [39] "Laser range scanner RPLIDAR A1 Datasheet." [Online]. Available: <https://download.slamtec.com/api/download/rplidar-a1m8datasheet/2.1?lang=en>. [Accessed: 14-Mar-2020]

BIOGRAPHIES



ZAHİR YILMAZ graduated from Damascus University in Syria Department of Computer Engineering and Automation in 2015. He received his MSc in Computer Engineering from Ataturk University in Turkey in 2020. His primary research interests are swarm robotics and machine learning.



LEVENT BAYINDIR received his PhD in Computer Engineering from Middle East Technical University, Turkey. He is currently an assistant professor of Computer Engineering at Atatürk University in Erzurum, Turkey. His research interests include swarm robotics and pervasive computing.

A 0.6-V 11- μ W PVT tolerant DTMOS inverter based OTA

Mesut Atasoyu

Abstract—This paper presents the design of a process-voltage-temperature (PVT) variation tolerant inverter-based operational transconductance amplifier (OTA) employing both a dynamic threshold MOS (DTMOS) technique and a constant voltage biasing (CVB) scheme. The proposed inverter-based OTA offers a higher bandwidth due to implemented DTMOS technique, which realizes higher input transconductance value than a conventional inverter-based OTA design. Simulation results show that the proposed OTA achieves superior slow-slow (SS) corner performance under PVT variations than the conventional inverter-based OTA while consuming only 11- μ W and providing a figure of merit (FoM) of 7.0-MHz-pF/ μ A. As a result, DC gain and unity-gain bandwidth (UGBW) of the proposed OTA improve by 27% and 32% at SS corner under the PVT variations, respectively.

Index Terms—DTMOS, inverter based OTA, PVT tolerant.

I. INTRODUCTION

Mobile devices demand low-power analog to digital converter (ADC) techniques, and continuous-time (CT) circuit techniques are exploited due to low-power constraint rather than discrete-time [1, 2]. Continuous-time analog-to-digital converters (CT-ADCs) should have high linearity on a low supply voltage [1]. Integrators are the main block in CT-ADCs, and an operational transconductance amplifier (OTA) in integrators is a power-hungry unit. Inverter-based OTAs offer higher power efficiency than conventional OTA amplifiers due to their bias current reuse mechanism. Recently, low power capable inverter-based OTAs have been reported [1–4]. Inverter-based OTA functioning in the sub-threshold region (digital-like amplifier) [4] is to be used in the second and third integrators of CT-ADCs [1]. However, inverter-based OTAs are vulnerable to process, voltage, and temperature (PVT) variations, which degrade CT-ADC performance [3, 5]. On the other hand, sufficient DC gain of inverter-based OTA at lower supply voltage is remaining a problem [4, 6].

It is important to say that proper biasing of inverter-based OTAs is a difficult task under PVT variations. Previous PVT compensation techniques in [3], have mitigated PVT variations with auto RC tuning and constant g_m current source circuitry. In [6] using variable current sources at the output stage by controlling common-mode voltage. In [7], a low-dropout voltage regulator controls the supply voltage. In [8], the main inverter amplifier PVT compensation was realized by controlling the bulk voltage of the inverter gate amplifier.

This work proposes the PVT tolerant inverter-based OTA using the constant biasing voltage (CVB) mechanism, composed of a current source with a diode-connected MOS device. On the other hand, in this bias method, the drain and source voltage (V_{DS}) of inverter gate transistors are clamped to a constant voltage. Furthermore, the unity-gain bandwidth (UGBW) of the proposed OTA is improved with the dynamic threshold MOS technique (DTMOS) [9], thus sub-threshold DTMOS provides additional g_{mb} improves the current drivability of the MOS device [10].

This paper is organized in the following manner. Section II explains the principle and circuit implementation of the PVT invariant inverter-based OTA. Section III reports the performance results of the inverter-based OTAs presented in detail. Section IV concludes final discussions on the proposed OTA design.

II. CIRCUIT IMPLEMENTATION

The OTAs in this work are biased in the weak inversion region. The DTMOS technique is implemented in a triple-well process. An NMOS bulk is connected to a p-well by adding a deep n-well layer in a triple-well process. In the conventional CMOS process, an NMOS transistor gate and bulk can not connect to each other due to short circuit concerns. However, the designing of a PMOS transistor is the same for these two processes. The DTMOS technique on the design of an inverter-based OTA increases the input transconductance of the OTA due to bulk transconductance value g_{mb} . Moreover, the DC-gain performance of inverter-based OTA is enhanced due to increased drain current drivability of the input transistors of the inverter-based OTA as the DTMOS technique modulates threshold voltages (V_{th}) of these transistors.

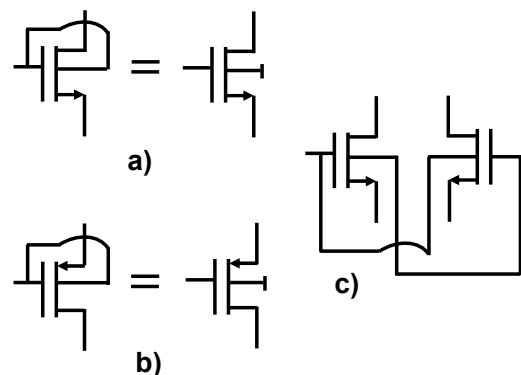


Fig. 1: The symbols of DTMOS transistors.

MESUT ATASOYU is with Department of Electrical and Electronics Engineering at Ataturk University.
E-mail: mesutatasoyu@atauni.edu.tr

<https://orcid.org/0000-0002-0029-7436>
Manuscript received October 7, 2021; accepted Apr 14, 2022.
DOI: 10.17694/bajece.1005797

Fig. 1 shows the symbols of DTMOS transistors. Fig. 1a and Fig. 1b is the symbol of DTMOS NMOS, and DTMOS PMOS transistors, respectively. Fig. 1c is the cross-coupled DTMOS (CDTMOS) symbol for NMOS and PMOS transistors in the same way, but an NMOS CDTMOS symbol is shown in this figure. The DTMOS operated in subthreshold region can provide 16% more input transconductance [11] as given below equation:

$$g_{mb} = (n - 1)^2 * g_m \quad (1)$$

where n is the subthreshold slope factor and taken as a value of ≈ 1.4 [11]. In this case, the input transconductance improves 32 % due to the input stage of the inverter-based OTA composed of an NMOS and PMOS transistor. In this work, the current starved inverter-based OTAs were simulated in 65nm CMOS technology, and their schemes were shown in Fig. 2 and Fig. 3. Fig. 2a shows the schematic of the conventional inverter-based OTA, two inverter gates combined of M_{2-5} transistors consist of the differential input and output stage, the biasing circuitry is consists of $M_{0,1}$ and $M_{6,7}$ transistors, and the common-mode feedback (CMFB) circuitry is consists of M_{8-10} . Fig. 3a shows the schematic of the proposed OTA, two inverter gates combined of M_{2-5} transistors consist of the differential input and output stage, the biasing circuitry is consists of $M_{0,1}$ and $M_{6,7}$ transistors, and the common-mode feedback (CMFB) circuitry is consists of $M_{8,9,10}$. The CVB biasing scheme is consists of M_{C1-C4} .

Fig. 2b and in Fig. 3b illustrate the small-signal equivalent circuits of the inverter-based OTAs used for the differential gain calculations of the conventional and proposed inverter-based OTAs. The differential AC gain formulas was given in the below equations, assuming $g_m * r_o \gg 1$:

$$G_{M[\text{conv}]} = g_{m2} + g_{m4} \quad (2)$$

$$R_{O[\text{conv}]} = \frac{(g_{m2} * r_{o2} * 2r_{o1})}{//(g_{m4} * r_{o4})} \quad (3)$$

$$G_{M[\text{proposed}]} = g_{m2} + g_{mb2} + g_{m4} + g_{mb4} \quad (4)$$

$$R_{O[\text{proposed}]} = \frac{(g_{m2} * r_{o2} * 2r_{o1})}{//(g_{c1} * g_{m4} * r_{o1} * r_{o4})} \quad (5)$$

$$A_{V,\text{differential}[\text{conv,proposed}]} = -G_M * R_O \quad (5)$$

The following equations for UGBW calculations of the conventional inverter-based OTA and proposed OTA are:

$$UGBW_{[\text{conv}]} = (g_{m2} + g_{m4})/C_L \quad (6)$$

$$UGBW_{[\text{proposed}]} = (g_{m2} + g_{mb2} + g_{m4} + g_{mb4})/C_L \quad (7)$$

In the voltage transfer characteristic of an inverter gate, the switching threshold voltage (V_{sw}) is the value of a point where input and output voltage are equal. In the ideal case, the desired value of V_{sw} is $(V_{DD}=0.6V)/2$. At this value, the noise margin is maximized [12]. On the other hand, the value of V_{sw} at $V_{DD}/2$ may be degraded due to process variations. To ensure V_{sw} voltage stability, common-mode feedback (CMFB) structure adopted from [13] is necessary to control the output common-mode voltage value at $(V_{DD}=0.6V)/2$. Transistors of CMFB circuitry are $M_{9,10}$ transistors operated in the triode

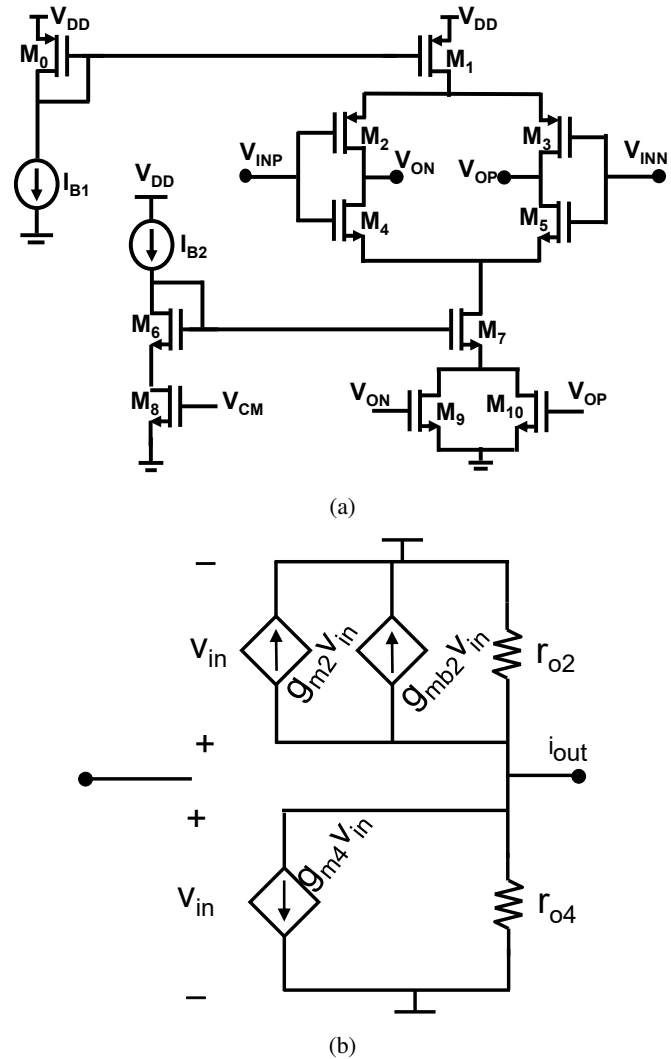


Fig. 2: (a)The schematic of the conventional inverter based OTA and (b)corresponding differential-mode small signal equivalent model of (a).

region, and M_8 transistor, driven by V_{CM} voltage source has a value of $V_{DD}/2$.

All transistor channel lengths are $1.33\mu\text{m}$ to provide sufficient intrinsic gain, except for M_{C3} device, which has a value of $L=10\mu\text{m}$. All gate widths of transistors were sized using g_m/I_D methodology with drain current normalization coefficients (I_D^*) of 65nm PMOS and NMOS transistors where were taken from [14]. I_D^* was extracted experimentally in the 65-nm CMOS process for NMOS the value of $0.49\text{-}\mu\text{A}$ and PMOS the value of $0.15\text{-}\mu\text{A}$ [14]. The inversion coefficient (IC) is a transistor sizing parameter, experimental values of IC for NMOS and PMOS transistors in this work obtained from [14]:

$$IC = \frac{I_D}{I_D^* \frac{W}{L}} \quad (8)$$

where I_D is drain current, W is the gate width of a MOS transistor, and L is the channel length of a MOS transistor. IC levels for different operating regions of MOS transistors: IC values < 0.1 correspond to the sub-threshold region, values

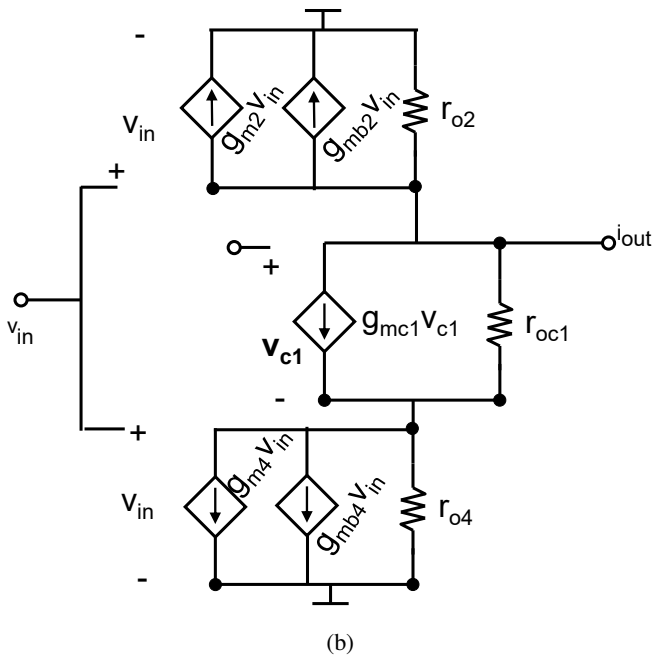
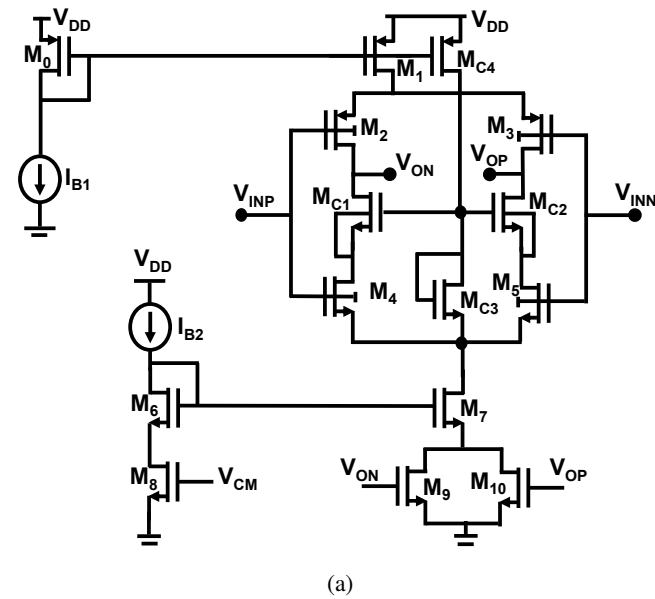


Fig. 3: (a)The schematic of the proposed inverter-based OTA and (b)corresponding differential-mode small-signal equivalent model of (a).

of $0.1 < IC < 10$ correspond to the linear region, IC values > 10 correspond to the saturation region. A transistor can operate in weak inversion, where IC values are less than 0.1. The IC value is 0.05 for the NMOS transistor sizing in this work. The corresponding value of an NMOS transistor current density (I_D/W) is 0.033 determined from Fig. 4. W value of an NMOS transistor calculates from the following equation:

$$W = \frac{I_{bias}}{I_D/W} \quad (9)$$

where I_{bias} is an applied biasing current. The CVB biasing scheme or the clamped V_{DS} biasing scheme is dedicated to the PVT compensation of the proposed OTA. The CVB scheme

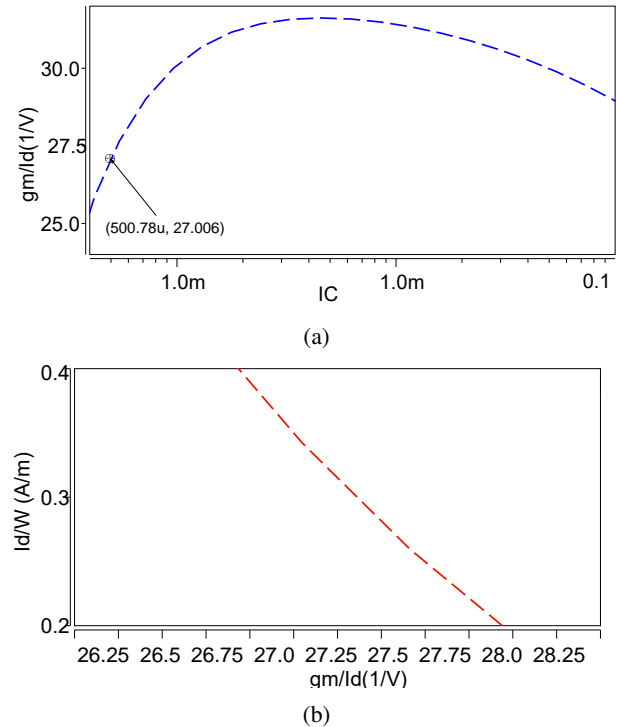


Fig. 4: The schematics of the inverter based OTAs.

is composed of M_{C3} diode-connected MOS device sized with $W=130\text{nm}$ and $L=10\mu\text{m}$, the current source transistor M_{C4} was sized with $W=1\mu\text{m}$, and $M_{C1,C2}$ voltage-clamp transistors are sized with $W=27\mu\text{m}$, as given in Table. I.

TABLE I: Transistor Sizes

Devices	($W(\mu\text{m})$)
$M_{0,1}$	162
$M_{2,3}$	81
$M_{4,5}$	27
M_{6-10}	54
$M_{C1,C2}$	27
M_{C3}	0.13
M_{C4}	1

III. SIMULATION RESULTS

The inverter-based OTAs were implemented in 65-nm CMOS technology. The simulated AC gain and the UGBW of conventional and proposed amplifiers are 35.9-dB and 5.82-MHz, 50.6-dB and 7-MHz, respectively. Regarding the phase margin (PM), the conventional and proposed OTAs are 90.0° and 85.7° , respectively, as given in Fig.5.

Figure 6 shows the transient analysis of the proposed OTA. The proposed OTA was designed with a capacitive feedback network, and the input signal was a square signal, which has a step of 600mV_{pp} and period of $1\mu\text{s}$, for the transient simulations. As for the slew rate, the proposed OTA has better performance of the slew rate over the conventional OTA.

To verify the effectiveness of the proposed inverter-based OTA, the corner analysis in HSPICE was performed under

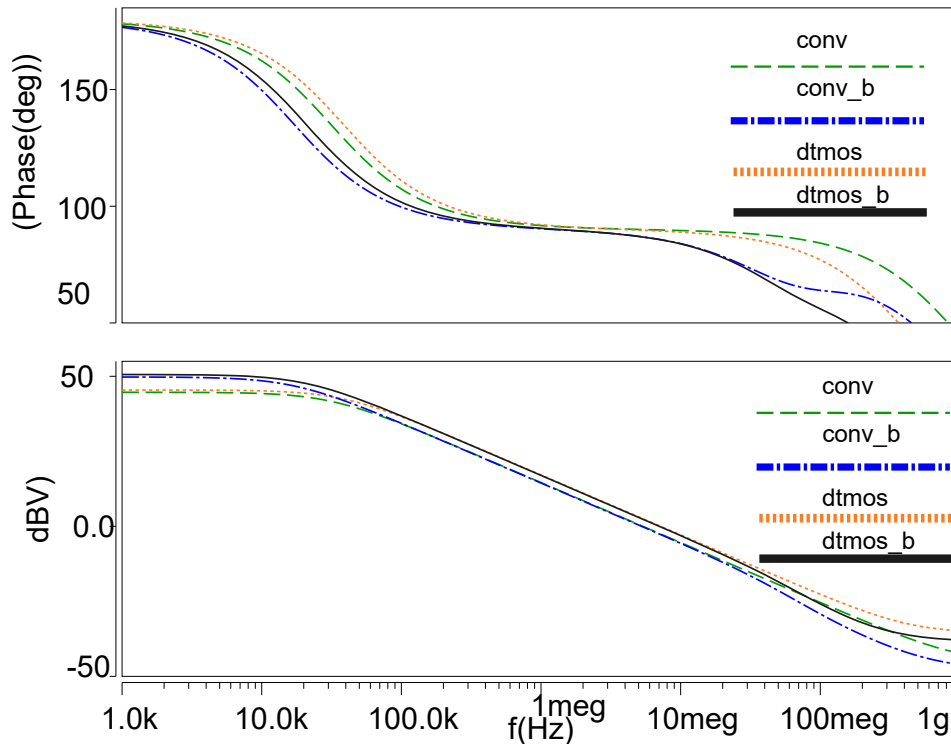


Fig. 5: AC response of the proposed inverter-based OTA. Both the inverter-based OTAs including CMFB circuits are biased with $10\text{-}\mu\text{A}$ current for 10-pF capacitive load under 0.6-V supply.

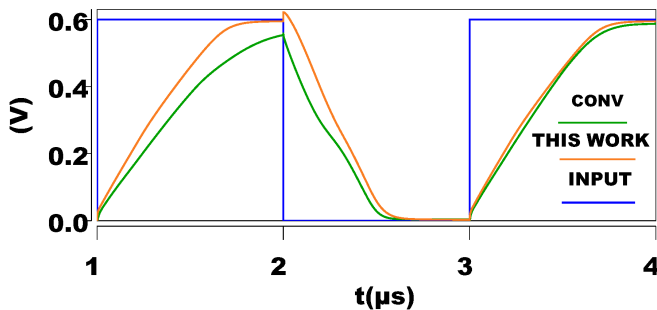


Fig. 6: Transient analysis of the proposed inverter-based OTA.

PVT variations setting up slow-slow (SS) corner for -40° and 0.54-V , typical-typical (TT) corner for 27° and 0.60-V , and fast-fast (FF) corner for 80° and 0.66-V [6]. On the other hand, especially in the SS process corner, an inverter-based OTA performance highly deteriorates. As a comparison of SS corner performance for the proposed OTA, the conventional OTA, the conventional OTA with CVB (Conv + CVB), the only DTMOS technique (DTMOS), the cross-coupled DTMOS (CDTMOS) with CVB (CDTMOS + CVB), and the Class-AB OTA [15] were performed. The proposed OTA has better AC gain and better UGBW performance in the SS corner, as shown in Table. II. The potential UGBW value which is highly dependent on PVT variation in an inverter based OTA design, [3]. As a result of the corner analysis under PVT variations, the gain degradation (from TT corner to SS corner) of the tested designs as follows: 34.6% for the Conv, 30.78% for the

Conv+CVB, 6.6% for the DTMOS, 12.6% for the proposed (DTMOS+CVB), and 13.9% for the CDTMOS+CVB. The UGBW degradation of the tested designs as follows: 87.7% for the Conv, 87.16% for the Conv+CVB, 70.25% for the DTMOS, 70% for the proposed (DTMOS+CVB), and 70.36% for the CDTMOS+CVB. However, the proposed OTA has the worse performance at the FF corner, as obtained result in [3].

Table. III summarizes the simulation results and provides a comparison with state-of-the-art inverter-based OTAs. The widely used one of the figures of merit (FoM) has been calculated for the performance evaluation of OTAs in Table. III:

$$FoM = UGBW * C_L / I_{supply} \quad (10)$$

where I_{supply} is the total current consumption. The proposed OTA design achieves better FoM performance and as the FoM value of $7.0 \text{ MHz}\cdot\text{pF}/\mu\text{A}$.

IV. CONCLUSION

In this paper, an improved PVT inverter-based OTA design was suggested. Compared to the conventional inverter-based OTA, the proposed circuit shows better AC gain, UGBW, and PVT tolerant design performances, thus having better FoM performance while consuming almost the same power budget. The proposed OTA operates in the sub-threshold region to enable low supply voltage circuit operation at the lowest 0.6V . The gain value of 50.6-dB , UGBW value of 7.0-MHz , and providing the FoM of $7.0 \text{ MHz}\cdot\text{pF}/\mu\text{A}$. The proposed inverter-based OTA design is one of the promising solutions for low

TABLE II: The OTA performances across different process corners under PVT variations.

	Gain(dB),UGBW(MHz)		
	SS (0.54-V, -40°)	TT (0.6-V, 27°)	FF (0.66-V, 120°)
Conv	(34.6, 0.68)	(44.6, 5.3)	(35.9, 5.82)
Conv+CVB	(34.4, 0.67)	(49.7, 5.22)	(34.2, 5.52)
DTMOS	(42.4, 2.13)	(45.4, 7.16)	(31.2, 6.59)
The proposed (DT-MOS+CVB)	(44.2, 2.09)	(50.6, 7.0)	(28.9, 6.10)
CDTMOS+CVB	(41.4, 1.55)	(48.1, 5.23)	(26.5, 4.65)
Class-AB OTA [15]	(37.96, 0.064)	(43.96, 0.112)	(43.55, 0.197)

TABLE III: Simulation performance results

Parameter	Conventional	([5])	Proposed	[15]	[16]	[17]	[18]
Power supply(V)	0.6	0.6	0.6	0.6	1.0	1.0	0.6
Technology(μm)	0.65	0.18	0.65	0.13	0.18	0.13	0.18
Power consumption(μW)	10	48.65	11	1.8	50	15	0.18
Load capacitance (pF)	10	8	10	6	140	50	20
DC gain (dB)	38.9	41.93	50.6	43.86	78.6	58.7	75.4
Unity-gain bandwidth (MHz)	2.7	15.55	7.0	0.112	2.02	10.43	0.074
Phase Margin($^\circ$)	90.0	88.06	85.7	88.8	85.4	78.8	78.86
Bias current(μA)	10	52	10	3.05	50	15	0.3
FoM ([17])	5.3	2.39	7.0	0.226	5.6	34.77	4.95

voltage and low power continuous-time ADC design in deep sub-micron CMOS technology.

ACKNOWLEDGEMENT

REFERENCES

- [1] B. Gönen, S. Karmakar, R. van Veldhoven, and K. A. Makinwa, "A continuous-time zoom adc for low-power audio applications," *IEEE Journal of Solid-State Circuits*, vol. 55, no. 4, pp. 1023–1031, 2019.
- [2] P. C. C. de Aguirre, E. Bonizzoni, F. Maloberti, and A. A. Susin, "A 170.7-dB fom-dr 0.45/0.6-V inverter-based continuous-time σ - δ modulator," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 67, no. 8, pp. 1384–1388, 2019.
- [3] Y. Guo, J. Jin, X. Liu, and J. Zhou, "An inverter-based continuous time sigma-delta adc with latency-free dac calibration," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 67, no. 11, pp. 3630–3642, 2020.
- [4] L. Lv, X. Zhou, Z. Qiao, and Q. Li, "Inverter-based subthreshold amplifier techniques and their application in 0.3-V delta-sigma modulators," *IEEE Journal of Solid-State Circuits*, vol. 54, no. 5, pp. 1436–1445, 2019.
- [5] P. de Aguirre and A. Susin, "Pvt compensated inverter-based ota for low-voltage ct sigma-delta modulators," *Electronics Letters*, vol. 54, no. 22, pp. 1264–1266, 2018.
- [6] A. Ismail and I. Mostafa, "A process-tolerant, low-voltage, inverter-based ota for continuous-time sigma-

delta adc,” *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 24, no. 9, pp. 2911–2917, 2016.

- [7] L. Breems, M. Bolatkale, H. Brekelmans, S. Bajoria, J. Niehof, R. Rutten, B. Oude-Essink, F. Fritschij, J. Singh, and G. Lassche, “A 2.2 ghz continuous-time δ - σ adc with- 102 dbc thd and 25 mhz bandwidth,” *IEEE Journal of Solid-State Circuits*, vol. 51, no. 12, pp. 2906–2916, 2016.
- [8] P. C. C. de Aguirre and A. A. Susin, “A 0.6-v, 74.2-db dr continuous-time sigma–delta modulator with inverter-based amplifiers,” *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 65, no. 10, pp. 1310–1314, 2018.
- [9] F. Assaderaghi, D. Sinitsky, S. A. Parke, J. Bokor, P. K. Ko, and C. Hu, “Dynamic threshold-voltage mosfet (dtmos) for ultra-low voltage vlsi,” *IEEE Transactions on Electron Devices*, vol. 44, no. 3, pp. 414–422, 1997.
- [10] M.-E. Hwang and K. Roy, “A 135mv 0.13 μ w process tolerant 6t subthreshold dtmos sram in 90nm technology,” in *2008 IEEE Custom Integrated Circuits Conference*. IEEE, 2008, pp. 419–422.
- [11] P.-H. P. Wang, H. Jiang, L. Gao, P. Sen, Y.-H. Kim, G. M. Rebeiz, P. P. Mercier, and D. A. Hall, “A near-zero-power wake-up receiver achieving- 69-dbm sensitivity,” *IEEE Journal of Solid-State Circuits*, vol. 53, no. 6, pp. 1640–1652, 2018.
- [12] M.-E. Hwang, “Supply-voltage scaling close to the fundamental limit under process variations in nanometer technologies,” *IEEE transactions on electron devices*, vol. 58, no. 8, pp. 2808–2813, 2011.
- [13] R. Wei, W. Wang, X. Xiao, and Q. Chen, “A low-power σ - δ capacitance-to-digital converter for capacitive sensors,” *IEEE Access*, vol. 7, pp. 78 281–78 288, 2019.
- [14] M. Manghisoni, L. Gaioni, L. Ratti, V. Re, and G. Traversi, “Assessment of a low-power 65 nm cmos technology for analog front-end design,” *IEEE Transactions on Nuclear Science*, vol. 61, no. 1, pp. 553–560, 2014.
- [15] F. Centurelli, A. Fava, M. Olivieri, P. Tommasino, and A. Trifiletti, “A low-voltage class-ab ota exploiting adaptive biasing,” *AEU-International Journal of Electronics and Communications*, vol. 122, p. 153282, 2020.
- [16] X. Zhao, Y. Wang, D. Jia, and L. Dong, “Ultra-high current efficiency single-stage class-ab ota with completely symmetric slew rate,” *AEU-International Journal of Electronics and Communications*, vol. 87, pp. 65–69, 2018.
- [17] J. M. Algueta-Miguel, A. Lopez-Martin, M. P. Garde, A. Carlos, and J. Ramirez-Angulo, “ ± 0.5 v 15 μ w recycling folded cascode amplifier with 34767 mhz· pf/ma fom,” *IEEE Solid-State Circuits Letters*, vol. 1, no. 7, pp. 170–173, 2018.
- [18] Q. Zhang, Y. Wang, X. Zhao, and L. Dong, “Single-stage multipath class-ab bulk-driven ota with enhanced power efficiency,” *AEU-International Journal of Electronics and Communications*, vol. 107, pp. 39–48, 2019.



Mesut Atasoyu Mesut Atasoyu received the Bachelor of Science (B.Sc.) in Electrical and Electronics Engineering, from Ataturk University, Turkey, in 2009. He completed his Master of Science (MSc) in 2014 and a Ph.D. degree in Electronics Engineering from Istanbul Technical University. In 2022, He joined the Department of Electrical and Electronics Engineering, Ataturk University, as an assistant professor. His research interests lie in analog and digital integrated circuit design with nano-device modeling.

Publication Ethics

The journal publishes original papers in the extensive field of Electrical-electronics and Computer engineering. To that end, it is essential that all who participate in producing the journal conduct themselves as authors, reviewers, editors, and publishers in accord with the highest level of professional ethics and standards. Plagiarism or self-plagiarism constitutes unethical scientific behavior and is never acceptable.

By submitting a manuscript to this journal, each author explicitly confirms that the manuscript meets the highest ethical standards for authors and coauthors

The undersigned hereby assign(s) to *Balkan Journal of Electrical & Computer Engineering* (BAJECE) copyright ownership in the above Paper, effective if and when the Paper is accepted for publication by BAJECE and to the extent transferable under applicable national law. This assignment gives BAJECE the right to register copyright to the Paper in its name as claimant and to publish the Paper in any print or electronic medium.

Authors, or their employers in the case of works made for hire, retain the following rights:

1. All proprietary rights other than copyright, including patent rights.
2. The right to make and distribute copies of the Paper for internal purposes.
3. The right to use the material for lecture or classroom purposes.
4. The right to prepare derivative publications based on the Paper, including books or book chapters, journal papers, and magazine articles, provided that publication of a derivative work occurs subsequent to the official date of publication by BAJECE.
5. The right to post an author-prepared version or an official version (preferred version) of the published paper on an internal or external server controlled exclusively by the author/employer, provided that (a) such posting is noncommercial in nature and the paper is made available to users without charge; (b) a copyright notice and full citation appear with the paper, and (c) a link to BAJECE's official online version of the abstract is provided using the DOI (Document Object Identifier) link.



ISSN: 2147- 284X
Year: April 2022
Volume: 10
Issue: 2

CONTENTS

- S. Buyrukoğlu, A. Akbaş;** Machine Learning based Early Prediction of Type 2 Diabetes: A New Hybrid Feature Selection Approach using Correlation Matrix with Heatmap and SFS,.....110 - 117
- S. İkizoğlu;** Distributed Formation Control Algorithm for Improved String Stability in Heterogenic Vehicle Platoons,.....118 - 124
- H. Çelebi;** A Dual-Radio Hybrid Mesh Topology for Multi-Hop Industrial IoT Networks in Harsh Environments,.....125 - 131
- İ. E. Saçu;** Fractional Integration Based Feature Extractor for EMG Signals,.....132 - 138
- S. Yılmaz;** Stabilization of Chaos in a Cancer Model: The Effect of Oncotripsy,.....139 - 149
- Ö. F. Ertuğrul, Y. Sönmez, N. Sezgin, E. Akil;** Assessment of Epileptic Seizures and Non-Epileptic Seizures via Wearable Sensors and Priori Detection of Epileptic Seizures,.....150 - 155
- L. Kouhalvandi;** A Prospective Look on Optimization Methods For RFID Systems: Requirements, Challenges and Implementation Aspects,.....156 - 169
- D. Çelik;** Performance Analysis of Three Levels Three Switches Vienna-Type Rectifier Based on Direct Power Control,.....170 - 177
- D. Ozcelik, Ö. Taştan;** A Weakly Supervised Clustering Method for Cancer Subgroup Identification,.....178 - 186
- Z. O. Davarci, M. Şahin, O. Akar;** Estimation by ANN of Luminous Efficacy of Lamps Used for Lighting,....187 - 197
- F. N. Deniz, M. Günay;** Coefficient Diagram Method Based Decentralized Controller for Fractional Order TITO Systems,.....198 - 208
- T. Sönmezocak;** The Use of Kalman Filter in Control The PanTilt Two-Axis Robot With Wearable System,....209 - 213
- Z. Yılmaz, L. Bayindir;** Lidar-based Robot Detection and Positioning using Machine Learning Methods,.....214 - 223
- M. Atasoyu;** A 0.6-V 11-uW PVT tolerant DT MOS inverter based OTA,.....224 - 229

BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING

(An International Peer Reviewed, Indexed and Open Access Journal)

Contact

Batman University
Department of Electrical-Electronics Engineering
Bati Raman Campus Batman-Turkey

Web: <http://dergipark.gov.tr/bajece>
<http://www.bajece.com>
e-mail: bajece@hotmail.com

