



Balkan Journal of Electrical & Computer Engineering

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

www.bajece.com

Vol : 8
No : 2
Year : 2020
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, iijFactor. 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)
Ozgur 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: 8
No : 2
Year: July 2020

CONTENTS

- N.H. Kaplan, I.Karabey Aksakalli, U. Kilic, I. Erer;** A Novel Approach for Copy-move Forgery Detection using Bilateral Filtering.....**114-120**
- E. Belge, H. K. Kaba, A. Parlak, A. Altan, R. Hacıoğlu ;** Estimation of Small Unmanned Aerial Vehicle Lateral Dynamic Model with System Identification Approaches,**121-126**
- M. Yılmaz, M. F. Çorapsız, M. R. Çorapsız;** Voltage Control of Cuk Converter with PI and Fuzzy Logic Controller in Continuous Current Mode.....**127-134**
- P. Arikan, S. Balci, F. Battal;** Determination of the roll-off value in the air-gapped inductor of a DC-DC boost converter circuit with FEA parametric simulations,**135-141**
- O. Akgun, A. Akan and O. Yılmaz;** The Spectral and Statistical Analysis of Gait Dynamics in ALS Disease,**142-146**
- Y. Sarica and M. Cetin;** Artificial Intelligence Based Game Levelling,**147-153**
- S. Bilgili, A. K. Demir;** Determination of the Ideal Protocol Stack for the Transmission of Health Data over 6LoWPAN IoT Networks,.....**154-163**
- A. Dönük;** Determination of Leakage Inductance Percentage for Gapped Iron-Core Shunt-Reactors with M4 Steel as Core Material,.....**164-169**
- A. Kaplan and E. Akagunduz;** A Hybrid Framework for Matching Printing Design Files to Product Photos,.....**170-180**
- AA. Abro, E. Taşci, A. Uğur;** A Stacking-based Ensemble Learning Method for Outlier Detection,.....**181-185**
- M.A. Çelik, D. Gökçen, F.E. Aydos;** Nested Miller Compensation Based Op-Amp Design for Piezoelectric Actuators,.....**186-192**

**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

A Novel Approach for Copy-move Forgery Detection using Bilateral Filtering

N.H. KAPLAN, I.KARABEY AKSAKALLI, U. KILIC, I. ERER


Abstract— Digital image processing methods have a wide area of usage and their complexity is increasing, as well as the tampering methods. A widely used tampering method is copy-move forgery. In this study, a hybrid method combining the Discrete Cosine Transform (DCT) and Bilateral filtering is developed. In this method, first overlapping blocks are obtained from the input image. Then, bilateral filtering and DCT of these blocks are multiplied to obtain the refined block features. The block features are scanned by a zig-zag process followed by a lexicographic sorting. Finally, a similarity detection by a predetermined threshold parameter is applied to detect the forgery. Both visual and quantitative results demonstrated that the proposed method can determine the copy-move forgery regions.

Index Terms— copy-move forgery, bilateral filtering, zigzag scanning, DCT (Discrete Cosine Transform)


I. INTRODUCTION

NOWADAYS, digital images are used in important areas such as medical, law and public. Digital images can be manipulated and regulated easily by malicious people using various image regulation software tools. With the emergence of this software, the reliability of the images and their authentication have become an important problem. Therefore, image fraud detection has become an important research focus. Image fraud detection methods are generally divided into two categories as active and passive approaches. Active


NUR HUSEYIN KAPLAN, is with Department of Electrical Engineering University of Erzurum Technical University, Erzurum, Turkey, (e-mail: huseyin.kaplan@erzurum.edu.tr).

 <https://orcid.org/0000-0002-4740-3259>


ISIL KARABEY AKSAKALLI, is with Department of Computer Engineering University of Erzurum Technical University, Erzurum, Turkey, (e-mail: isil.karabey@erzurum.edu.tr).

 <https://orcid.org/0000-0002-4156-9098>

UGUR KILIC, is with Computer Engineering University of Erzurum Technical University, Erzurum, Turkey, (e-mail: ugur.kilic@erzurum.edu.tr).

 <https://orcid.org/0000-0003-4092-3785>

ISIN ERER, is with Department of Electronical and Communication Engineering University of Istanbul Technical University, Istanbul, Turkey, (e-mail: ierer@itu.edu.tr).

 <https://orcid.org/0000-0002-2225-6379>

Manuscript received November 27, 2019; accepted Feb 6, 2020.
DOI: [10.17694/bajece.651435](https://doi.org/10.17694/bajece.651435)

approaches are based on additional information embedded into digital images such as digital watermarks or digital signatures. By using these additional information, the originality of the image can be detected. Unfortunately, active approaches require additional information to be embedded in the image by authorized personnel in the process of capturing the false image or at a later stage. If there is no information about the original image, applying an active approach is not useful [1].

On the other hand, passive approaches are used to determine the manipulated image without any additional information. Passive approaches are divided into two groups named tampering detection and source device identification. This approach detects copied image by extracting real features in the image. Tampering detection is also divided into dependent and independent classes. The dependent class of copy-move forgery is the commonly used method in fraud image. The image content is manipulated by copying an object that exists in the image and paste this object to another location within the same image. The transition between copied object and original image is masked using a variety of retouching tools. Since the features such as noise, color and contrast in the source and target regions have a statistical match, the detection of copied zones is a challenge.

Fridrich et al. [1], pioneers of copy-move forgery detection algorithm (CMFD), has handled the various requirements of the detection algorithm. The first requirement is that the detection algorithm should allow the approximate matching of the small image segments. Secondly, while the detection algorithm determines the mismatched fields (false positive), it must have an acceptable execution time. Furthermore, the authors mentioned that a fake segment will likely have a dependent component rather than very small patches or individual pixels.

In this study, a novel method consisting of a combination of Discrete Cosine Transform (DCT) and zigzag scanning is proposed by applying bilateral filtering. DCT is one of the most used watermarking algorithms among many data hiding methods to protect digital multimedia files. It states a limited sequence of data points in terms of a sum of cosine functions in different frequencies. Bilateral filtering has been used in many image processing algorithms [2-4]. Bilateral filters take into consideration both spatial and spectral properties of the image. By this way, the edge information is kept during the filtering process. The method is compared with traditional and state of art methods and the proposed method gives 99.7% accuracy rate in a standard dataset called “CoMoFoD” used for benchmarking the detection of tampering or copied images.

II. RELATED WORK

It is known that there is a correlation between the original image and the pasted object in copy-move forgery [1]. This correlation is used to detect forgery successfully. The methods for approximate matching of copy-moved and real segments using retouching tools or other image processing tools are not fully sufficient to detect forgery, so various approaches for detecting the forgeries are increasing day by day.

In most of the methods proposed for CMFD, the basic procedure is divided into pre-processing, feature extraction, matching, filtering and post-processing stages, respectively [5-8]. In the pre-processing stage, image data is improved to enhance image features or to reduce undesirable distortions within the image. One of the most commonly used methods at this stage is the conversion of RGB color channels of the input image into a single grayscale image [1], [5-9]. Besides, if the image is stored in a compressed format, the files are decompressed in the preprocessing phase [3]. The feature extraction phase is then applied to the selection of related information representing the properties of the image. This phase is carried out in two ways: dividing into blocks and key-point detection. In block-based approach, the image is divided into blocks in a square or round shape. These blocks can be divided by overlapping or non-overlapping division [8]. Then, properties are extracted from these blocks by using various features (frequency transform, texture and intensity, moments invariant, log-polar transform, dimension reduction etc.) and similarity comparison is performed between the blocks in the image. The third stage, matching, is a process in which the similarities between different segments of the image are detected. This process is carried out for each feature that is extracted and measured to define the manipulated area. In the literature, block-based matching can be performed by several methods such as sorting, hashing, correlation and Euclidean distance [8]. In the key point-based method, the image is subdivided into subfields and the feature points are removed only in certain regions by using methods such as SIFT, SURF, etc. [10]. Then, similar to the block-based method, each of the feature points is paired with each other using methods such as clustering, Euclidean distance, etc. If the focus is on computational complexity, the use of keypoint-based algorithms should be preferred. If the aim is to achieve a higher accuracy rate, block-based methods provide a higher accuracy rate [7]. In the scope of this study, it is aimed to obtain a higher accuracy rate than complexity.

In the literature, many block-based methods have been proposed for copy-move forgery detection and among these methods, DCT (Discrete Cosine Transform) based methods generally presents higher performance. Since DCT is used in the scope of this study, literature studies involving this method are given priority. Parveen et al. [7] have proposed a pixel-based method for copy-move forgery detection by converting the color image into a gray-scale image. Then the image is split into 8x8 block sizes and applied feature extraction using DCT (Discrete Cosine Transform) on different datasets. After the feature extraction, k-means algorithm is used for block clustering and the radix sort algorithm is used for feature matching. Wandji et al. [11] convert each block into DCT coefficients and apply feature extraction obtained from DCT

coefficients with Red, Green and Blue (RGB) channels. Afterward, similar block pairs are searched by ordering the properties alphabetically. In the last stage, duplicated regions, if any, are detected by using Euclidean distance as the similarity criterion. With this method, it has been proved that there is no decrease in the performance of the image with a slight rotation, JPEG compression, scrolling, scaling, blur and noise addition. A similar study is performed by Huang et. al [12] and DCT coefficients are used as a feature. It is observed that PCA (Principle Component Analysis) analysis method [13] is better for feature extraction in copy-move forgery detection when compared to detection forgery based on lexicographic sorting of DCT block coefficients. Wandji et al. [11] converted each block into DCT coefficients, then apply feature extraction obtained from these coefficients with red, green and blue (RGB) channels. Afterwards, similar block pairs are searched by ordering the properties alphabetically. In the last stage, duplicated regions, if any, are detected by using Euclidean distance as the similarity criterion.

Using the block-based method, not only the detection of the copied object, but also the image tampering is detected. Manu and Mehre [14] proposed two methods for detecting forgery by identifying the boundaries and location of the tampered image without any prior knowledge. The first method is the classification of tamper patterns with standard deviations of block discrete cosine transformations (BDCT) of textures. In the second method, the entropy of the histograms and the image quality artifacts due to image forgery are combined with the texture patterns of the first method. In these methods, a SVM based classifier is used to determine whether the image is tampered or not. It is emphasized that the method is implemented on various datasets and it is stated that the highest accuracy belongs to the second method with 98.89%. In another study [15], the RGB image is converted to grayscale and the resulting image is divided into blocks of $m \times m$ pixels. After the 2D DCT coefficients of the blocks have been calculated, zigzag scanning is performed on each block to reposition the coefficients to a feature vector. Extracted feature vectors are sorted by lexicographic sorting. Lastly, duplicated blocks are determined with Euclidean distance. The proposed method is used for extracting false positive and false negative values in the CoMoFoD dataset using different overlapping block sizes. As a result of this method, it is stated that the accuracy performance of different overlapping block sizes is affected by the size of the forged area, and the 4x4 overlapping block size causes a high false-positive compared to 8x8 overlapping block size, decreasing the accuracy of tampering detection in terms of precision.

III. PROPOSED METHOD

On the basis of all copy-move forgery detection methods in the literature, it is seen that there should be determined at least one copy of the same object. The areas within the image, which are exactly the same, are compared in size and shape. For this purpose, the image is divided into overlapping blocks by using block-based detection method. Then, feature extraction is performed from each block. Thus, each block has some features and the possibility of replicating regions having similar features is considered.

A. DCT (Discrete Cosine Transform)

During feature extraction, Discrete Cosine Transform (DCT) is first applied to the blocks in the proposed method. The intensity of the image at pixel (x, y) is I (x, y) and the block size is "b". The overlapping blocks of the image are represented by corresponding DCT coefficients as D (u, v) as per equation (1).

$$D(u, v) = \frac{2}{b} C(u) C(v) \sum_{x=0}^{b-1} \sum_{y=0}^{b-1} I(x, y) \cos \frac{\pi u(2x+1)}{2b} \cos \frac{\pi v(2y+1)}{2b} \quad (1)$$

where

$$C(u) = \begin{cases} \frac{1}{\sqrt{2}} & \text{if } u = 0 \\ 1 & \text{otherwise} \end{cases} \quad (2)$$

B. Zigzag Scanning

Zigzag scanning is used to group Direct Current (DC) low frequency and Alternative Current (AC) high frequency coefficients obtained from the DCT process. This scanning method maps 8x8 blocks to 1x64 block. It increases the rate of image and video compression. A sample of zigzag scanning process is shown in Fig.1.

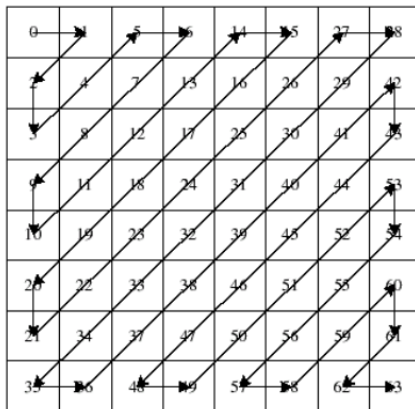


Fig.1. Sample of zigzag scanning process on 8x8 picture

C. Bilateral Filtering

Let Bilateral Filter0(I) (BF0(I)) be equal to the original image I at level 0, BF1(I) be the approximation image at level 1, p be the vector defining the location of the pixel to be filtered, S be the set of the neighbor pixels of p and q be the vector defining the location of a pixel in the set S. In the classical bilateral filtering, the lth approximation layer of a current pixel location p is estimated by:

$$BF(\mathbf{I})_p = \frac{1}{W_p} \sum_{q \in S} G_{\sigma_s}(\| \mathbf{p} - \mathbf{q} \|) G_{\sigma_r}(\| \mathbf{I}_p - \mathbf{I}_q \|) \mathbf{I}_q \quad (3)$$

with normalization parameter:

$$W_p = \sum_{q \in S} G_{\sigma_s}(\| \mathbf{p} - \mathbf{q} \|) G_{\sigma_r}(\| \mathbf{I}_p - \mathbf{I}_q \|) \quad (4)$$

and Gaussian kernel as:

$$G_{\sigma}(x) = \frac{1}{2\sigma^2} e^{-\frac{x^2}{2\sigma^2}} \quad (5)$$

The first detail plane is obtained by extracting the approximation image from the original image as

$$D^1(\mathbf{I}) = \mathbf{BF}^0(\mathbf{I}) - \mathbf{BF}^1(\mathbf{I}) \quad (6)$$

with $\mathbf{BF}^0(\mathbf{I}) = \mathbf{I}$ denoting the original image.

To obtain a multiscale decomposition, the spatial parameter σ_s is doubled and the range parameter σ_r is halved at each level.

The detail layer at level j can be obtained by the difference between the two consecutive approximation levels as

$$D^j(\mathbf{I}) = \mathbf{BF}^{j-1}(\mathbf{I}) - \mathbf{BF}^j(\mathbf{I}) \quad (7)$$

The original image can be reconstructed by simply adding the detail layers and the final level approximation layer as,

$$\mathbf{I} = \sum_{i=1}^L D^i(\mathbf{I}) + \mathbf{BF}^L(\mathbf{I}) \quad (8)$$

D. Algorithm Framework

This section describes the procedures performed for the proposed method, respectively. The framework of the method is shown in Fig.2.

Step-1: The input image has been converted into an mxn sized gray-scale image using Luminance values. A colored image is converted into grayscale using a standard formula of YUV conversion: $Y=0.299R+0.587G+0.114B$. In the formula; R, G, B represent three different color components of the RGB color model.

Step-2: The grayscale obtained in Step 1 is divided into 8x8 blocks. The purpose is to identify overlapping (m-b+ 1) (n-b + 1) blocks by shifting one pixel consisting of fixed-size bxb square windows from the top left to the bottom right corner of the image.

Step-3: After blocking, Bilateral filtering described in (3) is applied to each block to obtain the first level approximation layer for each block. In order to carry on decomposition Bilateral filtering is applied to the filtered blocks by doubling the spatial parameter and halving the spectral parameter to obtain the second level approximation layers for each block. The difference between these two approximation layers construct the second detail layer. Since each block is 8x8, the resulting detail layer will also be 8x8 and is named as Bilateral blocks.

Step-4: A normal image is a time-based image. Although the human eye can easily detect transitions at low frequency, it

cannot detect transitions in the high-frequency region. Therefore, copy-move forgery is usually performed in high-frequency regions. In this step, DCT is applied to 8x8 unfiltered blocks to convert from time domain to frequency domain. In DCT, the DC value of each block represents the brightness in that block.

Step-5: DCT applied blocks and Bilateral filtered blocks are multiplied and new 8x8 DCT-Bilateral blocks are created.

Step-6: Zigzag scanning has been applied to DCT-Bilateral filtered blocks to obtain a 1x64 sized vector and first 16 values of the vector are considered. $P \times B^2$ part of 1x64 sized vector is taken to reduce dimension in DCT-based image fraud. $P \in [0,1]$ is generally selected as 0.25 [12]. Thus, vector size drops to 1x16. In JPEG compression, fake data is searched in the low frequency region (first 16 values) due to the data loss occurs in the high frequency region.

Step-7: The quantization process is performed by dividing each vector composed with zigzag scanning into 16.

Step-8: Matrix A with $(N-7 * M-7, 16)$ elements is ordered lexicographically. In this step, vectors obtained from each block are placed in matrix A. Each vector is compared with all vectors placed after itself. Euclidean distance is used for similarity detection.

Euclidean distance: This method measures the difference between vectors and if the result is less than a certain threshold value, it is determined that these vectors are similar. The choice of the threshold value is very important, since it is used to determine how similar vectors are. In this step, lexicographically sorting has been performed to reduce comparison complexity. This reduction may result in disappearance of the information about which block the vector belongs. To avoid this, block values are stored by adding 2-bit location information to the end of the vector.

Step-9: After the lexicographic sorting, each element in the matrix A is compared with the element up to threshold value after itself using Euclidean distance. The initial coordinates of the vectors that are found to be similar are stored in a separate sequence. This sequence is called the shift-vector. If an area is selected as similar, the density of shift vectors is checked to ignore the false detected areas. The dense area is determined as the place of copying.

Step-10: The number of shift-vector is considered and if this number is up to threshold value, these areas are marked as a forgery.

Step-11: In the final step, the accuracy ratio and false negative values are calculated. The higher is the block size, the higher is the possibility of false negatives. The false-negative value increases when shift vector is used but this error is detected if there is multiple forgeries.

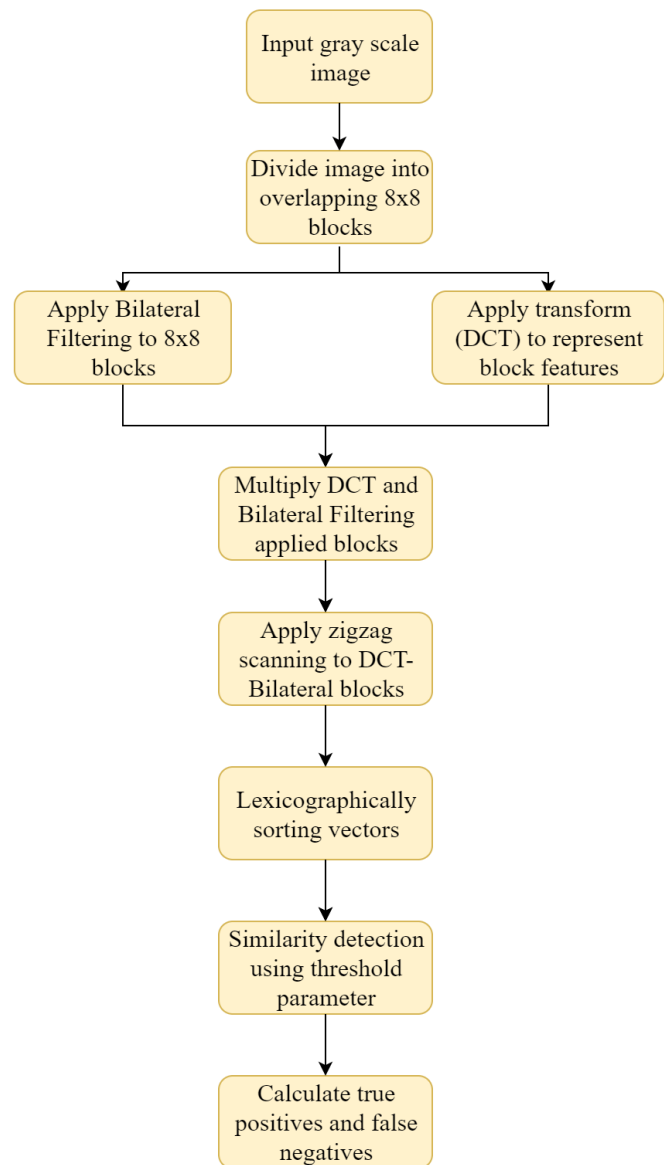


Fig.2. The framework of proposed method

IV. EXPERIMENTAL RESULTS

A. Dataset

According to our investigations, although many methods have been proposed to detect copy-move forgery, there was no common database created to test the proposed methods until 2013. In 2013, a common database called “CoMoFoD” is created by Video Communication Laboratory (VCL) [16] to detect copy-move forgery. The CoMoFoD dataset [17] contains two separate image sets containing 512 images of 512 x 512 and 60 images of 3000 x 2000. Each group has unreal images with five different manipulations such as translation, rotation, scaling, merging, and distortion with various composite operations. Also, each image set contains 40 different images containing the original image, color masking, dual masking, and fake images as shown in Fig. 3.

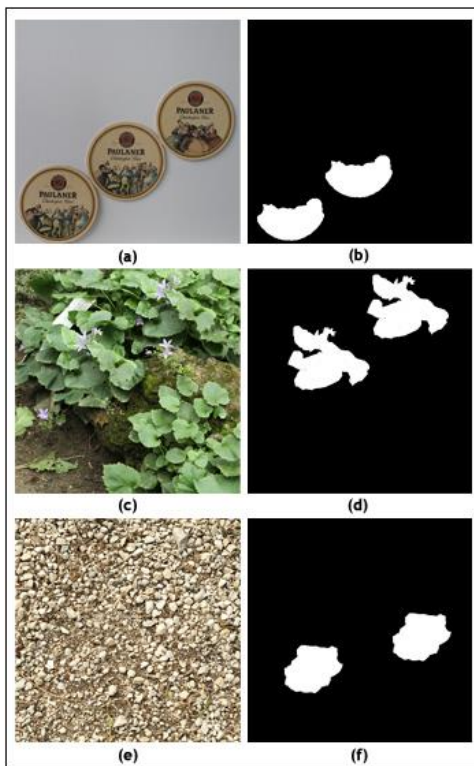


Fig.3. (a), (c) and (e) represent original images in CoMoFoD dataset, (b),(d) and (f) represent mask of these images

In addition, there are image forms in distorted original and fake using various last operation techniques for each image such as JPEG compression, blurring, noise addition, brightness change, contrast adjustment and color reduction [16-18].

B. Parameter Adjustment and Visual Comparisons

The performance of the bilateral filtering is highly correlated to three parameters, namely window size, spatial parameter and range parameter. In order to determine the optimum parameters, the proposed method is applied to several images and the accuracy rate of the resulting image is recorded. According to this comparisons, the optimum window size is determined as 3x3, spatial parameter as 1 and range parameter as 0.1.

After determining the optimum parameters, the proposed method has been tested on 40 different images on the CoMoFoD dataset. Among the 40 tested images, a copied region is pasted to more than one region in 12 images and more than one different regions are reproduced in 7 images with copy-move forgery. Fig. 4 shows the resulting images obtained by the proposed method we applied to the CoMoFoD dataset. According to the figure, the copy move forgeries are detected in a good accuracy, and a very little amount of false positives.

In order to make visual comparisons, the proposed method, DCT, Local Binary Pattern (LBP) and DCT-LBP methods have been applied to several images in CoMoFoD dataset and

sample results are shown in Fig. 5. Fig.5.a and Fig.5.b show the input image and the mask determining the copy move forgery area, respectively. Fig. 5.c, Fig. 5.e, Fig. 5.g and Fig. 5.i demonstrates the results of the DCT, LBP, DCT-LBP and the proposed methods, respectively. Fig. 5.d, Fig. 5.f, Fig. 5.h and Fig. 5.j show the copy move forgery areas obtained by the DCT, LBP, DCT-LBP and the proposed methods, respectively. According to Fig. 5., the closest result is achieved by the proposed method, with lower false positives. In order to make an objective comparison, quantitative assessments are made as well.

C. Quantitative Comparisons

In the copy-move forgery detection technique, two different assessments are made according to whether the copied area is regular or irregular.

In areas that are properly copied in square or rectangular form, the TP (True Positive) ratios of block-based copy-move forgery detection algorithms may be high and the FP (False Positive) ratios low. Considering this situation, the success of the proposed method on the experimental images is high.

TABLE I
COMPARISON OF PROPOSED METHODS IN LITERATURE

Method	Block Size			
	8x8		12x12	
	TP	FP	TP	FP
DCT [18]	0.75	0.12	0.65	0.40
LBP [18]	0.82	0.25	0.67	0.60
DCT-LBP [18]	0.89	0.08	0.85	0.12
DCT-Bilateral (proposed method)	0.96	0.042	0.95	0.051

The results of the proposed method in Table 1 are given in comparison with the other methods used previously using TP (True Positive) and FP (False Positive) criteria. The results are calculated based on the pixels that the algorithms correctly identified and marked as false positive. The calculation of TP and FP is given in equation (9) and (10), respectively. True positive (TP) represents the ratio of both result of the method and points labeled by the reference data.

$$TP = \frac{(C_1 \cap F_1) + (C_2 \cap F_2)}{(C_1 + C_2)} \quad (9)$$

The higher is the TP value, the better is the method. False-positive (FP) represents the proportion of points that are labeled by the method but not labeled in the reference image. C_1 and C_2 are copy-move areas, F_1 and F_2 are the fields determined by the method.

$$FP = \frac{(C_1 \cup F_1) + (C_2 \cup F_2)}{(C_1 + C_2)} - TP \quad (10)$$

The lower is the FP value, the better is the method.

In Table 1, average accuracy rate (TP) and false positive values (FP) are determined using some copy-move forgery detection methods in literature and the proposed DCT-Bilateral method by applying on the same images in CoMoFod dataset. When the results in Table 1 are compared, the best TP and FP scores are achieved by proposed DCT-Bilateral method. The worst TP values are obtained by DCT method, whereas the worst FP values are achieved by LBP method. Therefore, it is observed that the copy-move forgery detection in DCT-Bilateral applied images yields at least 7% more accurate results. It seems that DCT and LBP alone give less accuracy rates than combined methods.

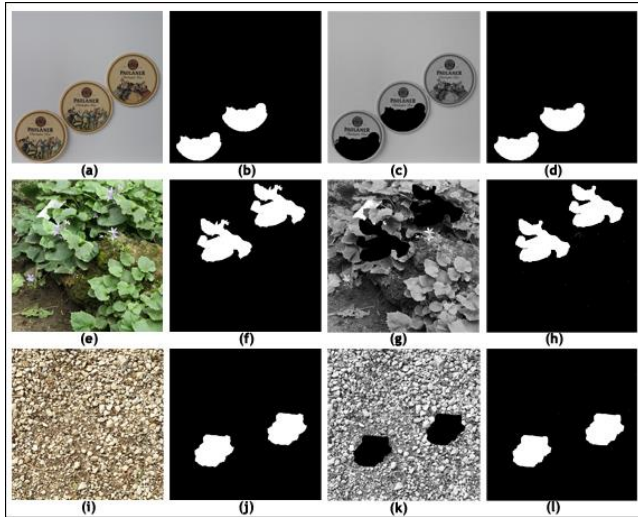


Fig.4. (a), (e) and (i) represent original images in ComoFod dataset; (b),(f) and (j) represent mask of these images; (c), (g), (k) DCT – Bilateral applied images; (d), (h), (l) represent mask of (c), (g), (k)

The images and their masks obtained from the proposed DCT-Bilateral method and the other methods in literature are listed in Fig.4. According to the obtained masks, only LBP applied method finds more false positive rates than the only DCT applied method. Similarly, when DCT-LBP and DCT-Bilateral images are compared, it is seen that DCT-Bilateral gives higher accuracy and lower false positive rates in both 8x8 and 12x12 images. These findings are also overlapped with given values in Table 1.

V. CONCLUSION

In this study, a block-based method is proposed for the first time using DCT and Bilateral filtering to detect copy-move forgery. Two different Gaussian kernels, namely spatial and range kernels of the bilateral filtering keep the edge information in the process of filtering. After dividing the image into overlapping bxb blocks, we applied bilateral filtering and DCT to each block separately. Then we multiply the obtained new blocks with each other to obtain the feature vectors. When we compare the proposed method with other copy-move forgery methods, it is seen that the true positive rate of the proposed method is higher. The proposed method is the first method in copy-move forgery detection technique using DCT and Bilateral combination. This method also works

on both gray-level images and color images. Although the proposed method in this study gives good results after blurring and noise addition on the image, it performs poorly in forgery by performing scaling or rotating operations on the copied image while performing copy-move forgery.

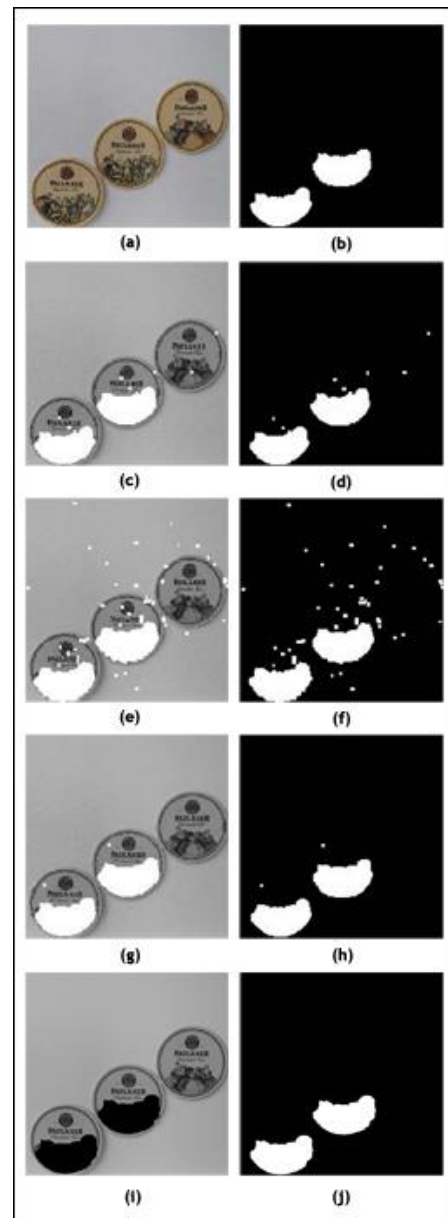


Fig. 5. (a) and (b) original images; (c) and (d) only applied DCT image; (e) and (f) only applied LBP image; (g) and (h) DCT-LBP image; (i) and (j) DCT-Bilateral image

REFERENCES

- [1] A. J. Fridrich, B. D. Soukal, and A. J. Lukáš, "Detection of copy-move forgery in digital images," in *in Proceedings of Digital Forensic Research Workshop*, 2003: Citeseer.
- [2] N. H. Kaplan and I. Erer, "Bilateral pyramid based pansharpening of multispectral satellite images," in

- 2012 *IEEE International Geoscience and Remote Sensing Symposium*, 2012: IEEE, pp. 2376-2379.
- [3] N. H. Kaplan and I. Erer, "Bilateral filtering-based enhanced pansharpening of multispectral satellite images," *IEEE geoscience and remote sensing letters*, vol. 11, no. 11, pp. 1941-1945, 2014.
- [4] S. Paris, P. Kornprobst, J. Tumblin, and F. Durand, "Bilateral filtering: Theory and applications," *Foundations and Trends® in Computer Graphics and Vision*, vol. 4, no. 1, pp. 1-73, 2009.
- [5] H. A. Alberry, A. A. Hegazy, and G. I. Salama, "A fast SIFT based method for copy move forgery detection," *Future Computing and Informatics Journal*, vol. 3, no. 2, pp. 159-165, 2018.
- [6] A. Novozámský and M. Šorel, "Detection of copy-move image modification using JPEG compression model," *Forensic science international*, vol. 283, pp. 47-57, 2018.
- [7] A. Parveen, Z. H. Khan, and S. N. Ahmad, "Block-based copy-move image forgery detection using DCT," *Iran Journal of Computer Science*, vol. 2, no. 2, pp. 89-99, 2019.
- [8] N. B. A. Warif *et al.*, "Copy-move forgery detection: survey, challenges and future directions," *Journal of Network and Computer Applications*, vol. 75, pp. 259-278, 2016.
- [9] R. Lionnie, R. B. Bahaweres, S. Attamimi, and M. Alaydrus, "A study on pre-processing methods for copy-move forgery detection based on SIFT," in *TENCON 2017-2017 IEEE Region 10 Conference*, 2017: IEEE, pp. 1142-1147.
- [10] D. Chauhan, D. Kasat, S. Jain, and V. Thakare, "Survey on keypoint based copy-move forgery detection methods on image," *Procedia Computer Science*, vol. 85, pp. 206-212, 2016.
- [11] N. D. Wandji, S. Xingming, and M. F. Kue, "Detection of copy-move forgery in digital images based on DCT," *arXiv preprint arXiv:1308.5661*, 2013.
- [12] Y. Huang, W. Lu, W. Sun, and D. Long, "Improved DCT-based detection of copy-move forgery in images," *Forensic science international*, vol. 206, no. 1-3, pp. 178-184, 2011.
- [13] A. C. Popescu and H. Farid, "Exposing digital forgeries by detecting duplicated image regions," *Dept. Comput. Sci., Dartmouth College, Tech. Rep. TR2004-515*, pp. 1-11, 2004.
- [14] V. Manu and B. Mehtre, "Tamper detection of social media images using quality artifacts and texture features," *Forensic science international*, vol. 295, pp. 100-112, 2019.
- [15] M. H. Alkawaz, G. Sulong, T. Saba, and A. Rehman, "Detection of copy-move image forgery based on discrete cosine transform," *Neural Computing and Applications*, vol. 30, no. 1, pp. 183-192, 2018.
- [16] VCL. "Comofod - image database for copy-move forgery detection,," <http://www.vcl.fer.hr/comofod/> (accessed September 27, 2019).
- [17] D. Tralic, I. Zupancic, S. Grgic, and M. Grgic, "CoMoFoD—New database for copy-move forgery detection," in *Proceedings ELMAR-2013*, 2013: IEEE, pp. 49-54.
- [18] A. Boz and H. Ş. Bilge, "Copy-move image forgery detection based on LBP and DCT," in *2016 24th Signal Processing and Communication Application Conference (SIU)*, 2016: IEEE, pp. 561-564.

BIOGRAPHIES



NUR HUSEYIN KAPLAN received his PhD in Electronics and Telecommunication Engineering from Istanbul Technical University, Turkey. He is currently an Associate Professor at Electrical and Electronics Engineering Department, Erzurum, Turkey. His primary research interests include digital signal and image processing.



ISIL KARABEY AKSAKALLI graduated from Gazi University in 2013 and started to work as a research assistant at Atatürk University in 2014. After receiving her MSc degree from Atatürk University in 2015, she was appointed as a research assistant to Erzurum Technical University in the Department of Computer Engineering. She started her PhD. in 2016 at Hacettepe University and still continues to this program. Her research topics include microservice architectures, optimization methods, distributed systems, machine learning and deep learning techniques.



UGUR KILIC graduated from Harran University in Turkey and he received his MSc in Computer Engineering from Ataturk University, Turkey. After the graduation from the Ataturk University, he started PhD in 2016 at Karadeniz Technical University. He is currently work as a research assistant at Erzurum Technical University. His primary research interests are IOT security, image processing and digital signal processing.



ISIN ERER is an Associate Professor in the Department of Electronics and Communication Engineering, Istanbul Technical University, Istanbul, Turkey. Her research interests include: Statistical signal processing, image processing for remote sensing, high-resolution radar imaging and ground penetrating radar.


Estimation of Small Unmanned Aerial Vehicle Lateral Dynamic Model with System Identification Approaches

E. BELGE, H. K. KABA, A. PARLAK, A. ALTAN, R. HACIOĞLU


Abstract— Modeling of unmanned aerial vehicle (UAV) with system identification is very important in terms of its model-based effective control. The modeling of UAV is required for aircraft crashes, analyzing autonomous aircrafts, preventing external disturbances, pre-flight analysis. However, since UAV has nonlinear inherent dynamics including inherent chaoticity and fractality, it becomes difficult to obtain a mathematical model under external disturbance. In this study, some of the inherent nonlinear dynamics of UAV are linearized and the model of UAV is obtained by system identification approaches under external disturbance. The linearized lateral dynamics of a fixed wing UAV is used in this study. Further, the flight motion equations applied to fixed wing UAV have been utilized for obtaining the coefficients of lateral model for straight and level flight. The roll angles are calculated using transfer functions for aileron, rudder and deflections inputs. The autoregressive exogenous (ARX), autoregressive moving average with exogenous (ARMAX) and output error (OE) parametric system identification approaches are performed to estimate UAV lateral dynamic system response as using empirical input-output data sets. The accuracy of parametric model estimation and model degrees are compared for different external disturbance effects.

Index Terms— System identification, UAV, lateral dynamic model, parameter estimation, external disturbance.

EGEMEN BELGE, is with Department of Electrical Electronics Engineering, Zonguldak Bülent Ecevit University, Zonguldak, Turkey, (e-mail: egemenbelge@beun.edu.tr).

 <https://orcid.org/0000-0001-5852-1085>

HIZIR KAN KABA, is with Department of Electrical Electronics Engineering, Zonguldak Bülent Ecevit University, Zonguldak, Turkey, (e-mail: hkaankaba@gmail.com).

 <https://orcid.org/0000-0001-5203-8092>

AHMET PARLAK, is with Department of Electrical Electronics Engineering, Zonguldak Bülent Ecevit University, Zonguldak, Turkey, (e-mail: mimikli6167@gmail.com).

 <https://orcid.org/0000-0003-0751-2263>

AYTAÇ ALTAN, is with Department of Electrical Electronics Engineering, Zonguldak Bülent Ecevit University, Zonguldak, Turkey, (e-mail: aytacaltan@beun.edu.tr).

 <https://orcid.org/0000-0001-7923-4528>

RIFAT HACIOĞLU, is with Department of Electrical Electronics Engineering, Zonguldak Bülent Ecevit University, Zonguldak, Turkey, (e-mail: hacirif@beun.edu.tr).

 <https://orcid.org/0000-0002-2480-0729>

Manuscript received Dec 03, 2019; accepted Feb 17, 2020.

DOI: [10.17694/bajece.654499](https://doi.org/10.17694/bajece.654499)

I. INTRODUCTION

UNMANNED AERIAL VEHICLE (UAV) modeling has become an important research subject in recent years. Modeling enables analysis of UAV's flight system under different conditions and determination of system reliability. System identification techniques are frequently used in modeling of dynamic systems such as UAV [1]. In this study, the model parameters of UAV is obtained with system identification approaches including Autoregressive Exogenous (ARX), Autoregressive Moving Average with Exogenous Variable (ARMAX) and Output Error (OE) under different external disturbances.

In the literature, there are studies conducted with various system identification methods in the modeling of UAV. In [2], an adaptive system identification technique has been used in real-time dynamic modeling of UAV. For a small UAV, an appropriate approach is proposed, which deals with the online prediction in the frequency domain. In [3], transfer function model with pitch and roll response of UAV are investigated using system identification techniques. The practical system identification procedure of the fixed-wing UAV is specified in [4]. It is emphasized that the system identification procedure used minimize the complexity of UAV movements. In [5], both online and offline models of nonlinear and complex UAV have been obtained using system identification procedure based on Artificial Neural Network (ANN). In [6], the flight control optimization has been handled together with the model of UAV. It is mentioned that hardware installation, flight test, flight data collection and processing, system identification, model verification, controller optimization and verification process steps have been performed. In [7], Hammerstein model is obtained for route tracking under zero averaged white Gaussian noise external disturbing effect based on experimentally obtained input and output data of three-axis camera system on the autonomously acting six-rotor UAV. In [8], the performance of the obtained model has been compared with Nonlinear Autoregressive and Moving Average (NARMA) model performance. The transfer function of the three-axis gimbal system has been obtained by linearly structured OE model using experimentally obtained data under different external disturbance effect. Model degree has been determined and data set based verification has been applied. Also, the performance has been compared by examining the effect of

external disturbance in the transfer function obtained. In [9], the model of Load Transporting System (LTS) originally designed on UAV has been obtained by linear ARX model structure and the Model Predictive Controller has been performed. In [10], the Multi Input Multi Output (MIMO) model of helicopter is obtained using ARX system identification model. The flight dynamics of the helicopter have been analyzed with various transfer functions. The system identification of a quadrotor-based aerial manipulator is presented in [11]. ARX and ARMAX models have been obtained from linear accelerations and yaw angular accelerations. The different control-oriented models of a quadrotor UAV have been obtained by applying different identification methods in [12]. Parametric techniques, neural networks, neuro-fuzzy inference systems and hybridization of some of them have been applied. The system identification techniques are utilized in the literature for UAV dynamics. The difference of this study from studies in literature is that the parametric system identification techniques are used for modelling roll angle with aileron and rudder input in fixed wing UAV under external disturbances. The proposed system identification model is more robust to external disturbance variances different from literature studies.

In this study, system identification method is used to define the roll angle for aileron and rudder inputs. The modelling of roll angle is substantial property to define lateral dynamics of small fixed-wing UAV. The OE model proposed to allow the UAV to track the specified route with minimum error is compared with other system identification approaches. This paper is organized as follows: small fixed-wing UAV is modelled, and ARX, ARMAX, OE system identification procedure is explained in Section II. The estimation results of UAV's lateral dynamic model are presented in Section III. Conclusions are finally given in Section IV.

II. MATERIALS AND METHOD

A. Fixed-Wing UAV Modelling

The lateral dynamic model of fixed wing UAV in Fig. 1 is discussed as in [2] where five states $x = (v, p, r, \phi, c)$, two inputs $u = (\delta_a, \delta_r)$, and five outputs (β, p, r, ϕ, c) are given.

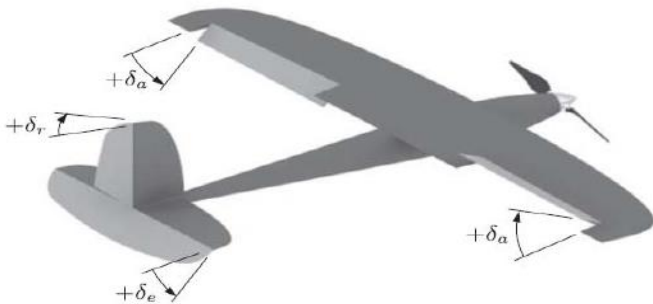


Fig.1. Control surface of fixed wing UAV.

The velocity, roll angular rate, yaw angular rate, roll, yaw, sideslip angle, aileron deflection, rudder deflection parameters are represented by $v, p, r, \phi, c, \beta, \delta_a$, and δ_r , respectively. The linearized lateral state space equations can be written as in [13]

$$\dot{x} = Ax + Bu \quad (1)$$

$$y = Cx \quad (2)$$

where state matrices

$$A = \begin{bmatrix} -0.8726 & 0.8789 & -16.82 & 9.791 & 0 \\ -2.823 & -16.09 & 3.367 & 0 & 0 \\ 0.702 & 0.514 & -2.775 & 0 & 0 \\ 0 & 1 & 0.05406 & -4.088 * 10^{-24} & 0 \\ 0 & 0 & 1.001 & -7.573 * 10^{-23} & 0 \end{bmatrix}$$

$$B = \begin{bmatrix} -0.8726 & 5.302 \\ -156.5 & -5.008 \\ 11.5 & -82.04 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$C = \begin{bmatrix} 0.05882 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Transfer function for linearized roll (ϕ) output due to aileron (δ_a) is given as

$$\frac{\phi(s)}{\delta_a(s)} = \frac{-155.8s^3 - 525.8s - 4283}{s^4 + 19.74s^3 + 90.49s^2 + 502.2s + 6.89} \quad (3)$$

Also the transfer function for linearized roll (ϕ) output due to rudder (δ_r) is given as

$$\frac{\phi(s)}{\delta_r(s)} = \frac{-9.443s^3 - 384.4s - 4370}{s^4 + 19.74s^3 + 90.49s^2 + 502.2s + 6.89} \quad (4)$$

The roll output data of dynamic lateral system are obtained with normal distribution aileron and rudder inputs. The obtained data is used in parametric estimation process. Both transfer function can be easily discretised by using appropriate sampling time.

B. System Identification Approaches

The system identification is a phenomenon to construct the mathematical modelling of dynamic systems using measured input output data [14]. There are different type of linear model parametric system identification techniques such as ARX, ARMAX and OE. In this study, the roll output, aileron and rudder input data used the model identification procedure and least-squares based method used to estimate the parameters. The block diagram of parameter estimation process of used system in this study is shown in Fig. 2.

General model structure of discrete-time input-output linear dynamic system with disturbance can be represented as

$$A(q)y(n) = \frac{B(q)}{F(q)}u(n) + C(q)e(n) \quad (5)$$

where $u(n)$, $y(n)$ and $e(n)$ are the input, output and disturbance respectively, Also $A(q)$, $B(q)$, $C(q)$ and $F(q)$ are polynomial with the delay operator q^{-1} .

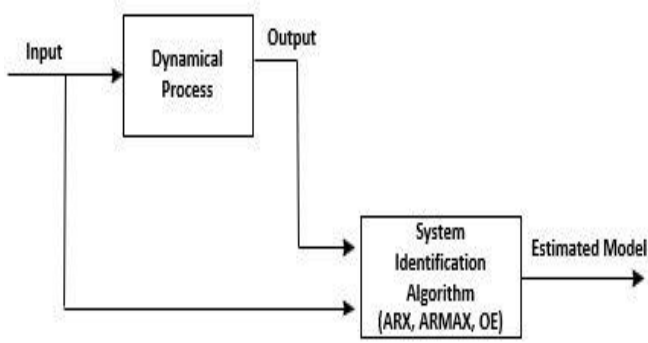


Fig.2. System identification block diagram.

The structure of ARX, ARMAX and OE models are presented in Figs. 3, 4 and 5, respectively.

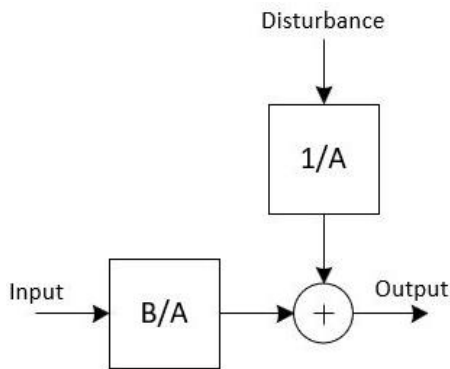


Fig.3. ARX model.

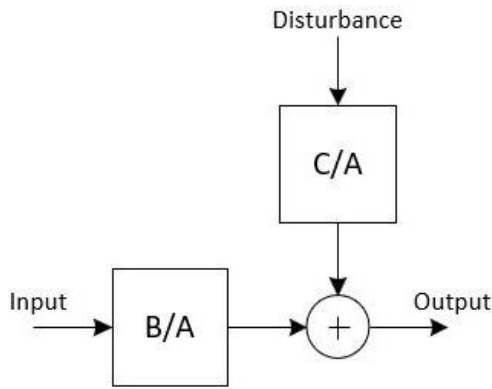


Fig.4. ARMAX model.

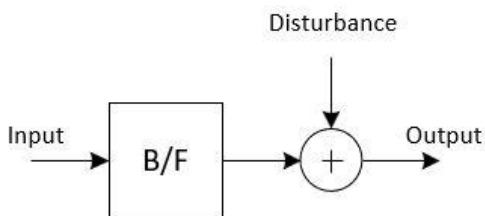


Fig.5. OE model.

ARX model parameters a_k, b_k are given as

$$A(q) = 1 + a_1q^{-1} + a_2 + \dots + a_{na}q^{-na} \quad (6)$$

$$B(q) = b_1q^{-1} + \dots + b_{nb}q^{-nb} \quad (7)$$

where model degree is na , q^{-1} is time delay operator. ARMAX model parameters a_k, b_k, c_k are presented as in Eqs. (6) and (7).

$$C(q) = 1 + c_1q^{-1} + \dots + c_{nc}q^{-nc} \quad (8)$$

ARMAX model includes $C(q)$ parameter unlike ARX model. Model degree nf OE model parameters are indicated in Eqs. (7) and (9).

$$F(q) = 1 + f_1q^{-1} + \dots + f_{nf}q^{-nf} \quad (9)$$

III. RESULTS AND DISCUSSION

The input-output data of fixed wing UAV with 0.2 disturbance variance is shown in Fig. 6 for 100 sec normally distribution aileron input where the sampling time 0.1 sec. Model order has been changed four to six to get the performance of linear ARX, ARMAX and OE identification approaches. The input output data is divided into test data with the first 700 samples of 1000 samples and validation data with the last 300 samples.

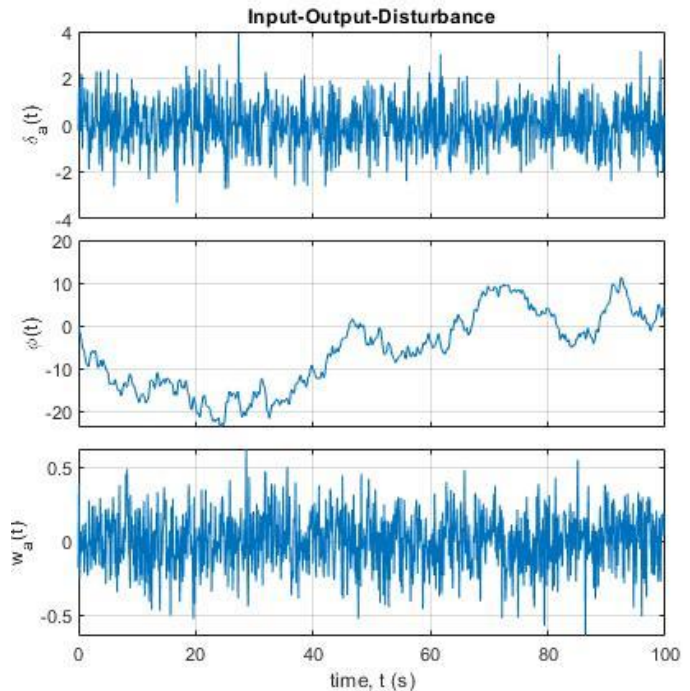


Fig.6. The input-output of fixed-wing UAV with disturbance for aileron input.

The estimated model outputs and squared errors for six model degrees and aileron input are shown in Fig. 7. These graphs express that the OE model accuracy is higher than the other models. The OE and ARMAX model are closer to the system output.

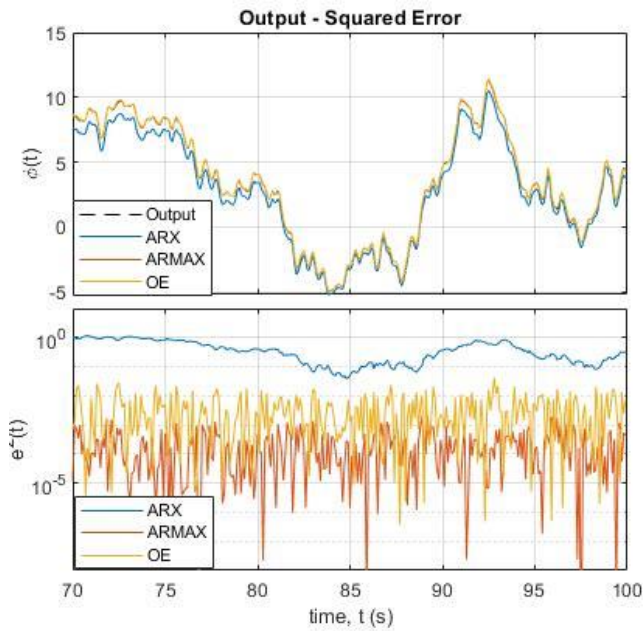


Fig.7. Model outputs and squared errors for aileron input.

The degree of models and estimation ratio system output-aileron input with/without (w/wo) disturbance variance are presented in Table 1.

TABLE 1
DEGREE OF MODELS AND ESTIMATION RATIO w/wo DISTURBANCE FOR AILERON INPUT

Model Type	Model Degree		Estimation Ratio w/wo Disturbance (%)	
	ARX (na, nb)	ARMAX (na, nb, nc)	OE (nb, nf)	
ARX	(4, 4)		84.78	91.33
ARMAX	(4, 4, 4)		86.06	97.10
OE	(4, 4)		86.22	97.15
ARX	(5, 5)		85.23	94.11
ARMAX	(5, 5, 5)		86.52	99.12
OE	(5, 5)		86.57	99.47
ARX	(6, 6)		86.34	96.33
ARMAX	(6, 6, 6)		86.65	99.58
OE	(6, 6)		87.81	99.66

The maximum parameter estimation accuracy of system without disturbance for aileron input obtained for (nf=6, nb=6) OE model. The accuracy value of OE model is 99.66%. Further, the maximum accuracy for system with disturbance is obtained for (na=6, nb=6, nc=6) ARMAX and (nf=6, nb=6) OE model. The accuracy values of ARMAX and OE model are 86.65% and 87.81%, respectively.

The rudder input, system output and disturbance graphs are shown in Fig. 8. The disturbance variance is defined as 0.2. The estimated model outputs and squared errors for 6 model degrees and rudder input are shown in Fig. 9. These model error and model output graphs state that the OE and ARMAX model accuracy is higher than ARX model.

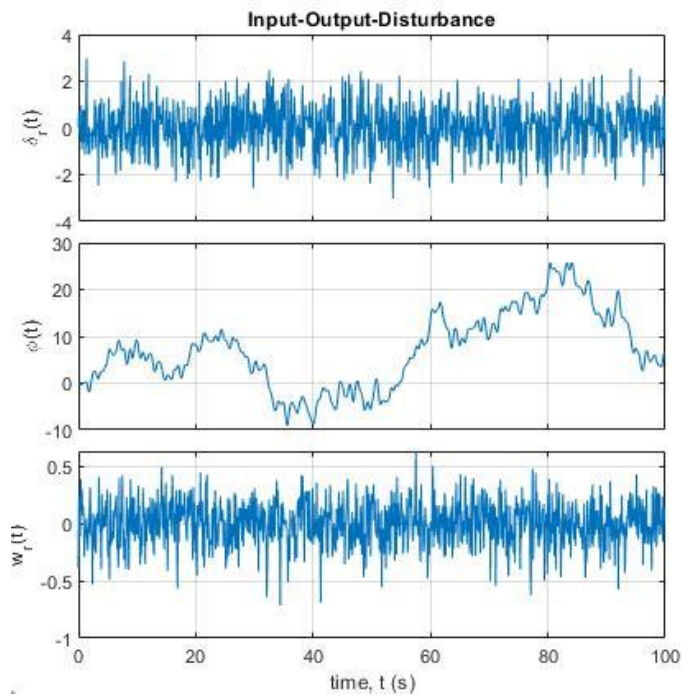


Fig.8. The input-output of fixed-wing UAV with disturbance for rudder input.

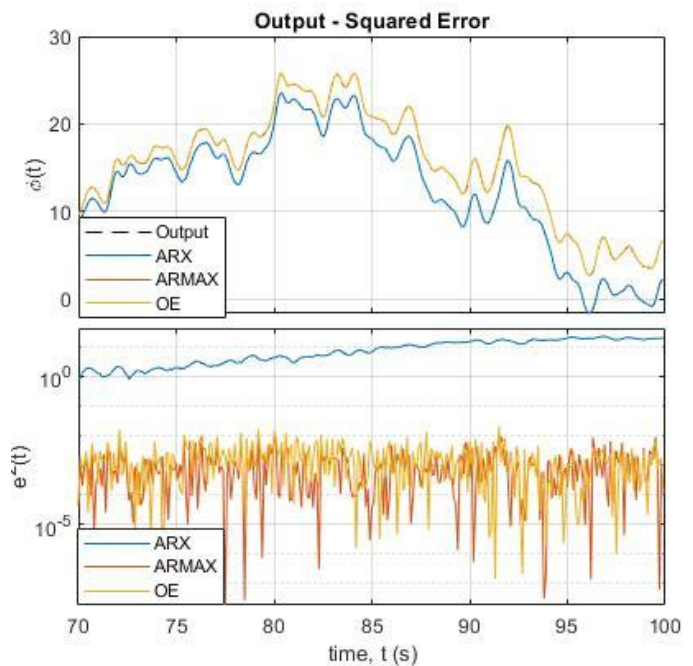


Fig.9. Model outputs and squared errors for rudder input.

The degree of models and estimation ratio system output-rudder input with/without disturbance are presented in Table 2. The maximum parameter estimation accuracy of system without disturbance for rudder input obtained for (na=6, nb=6 and nc=6) ARMAX and (nf=6 and nb=6) OE model. The accuracy of ARMAX and OE model is 99.97%. Also, the maximum accuracy with disturbance system obtained for (nf=6 and nb=6) OE model. The accuracy of OE model is 88.36%.

The lateral dynamics of fixed-wing UAV estimated through ARX, ARMAX and OE system identification models. According to results, lateral dynamics of a fixed wing UAV will

be modelled using parametric system identification approaches. The obtained results also will be basic for the other dynamics of fixed-wing UAVs.

TABLE 2
DEGREE OF MODELS AND ESTIMATION RATIO w/wo
DISTURBANCE FOR RUDDER INPUT

Model Type	Model Degree	Estimation Ratio	
	ARX (na, nb) ARMAX (na, nb, nc) OE (nb, nf)	w/wo Disturbance (%)	
ARX	(4, 4)	80.63	98.60
ARMAX	(4, 4, 4)	86.96	99.90
OE	(4, 4)	87.23	99.93
ARX	(5, 5)	81.59	99.62
ARMAX	(5, 5, 5)	87.12	99.94
OE	(5, 5)	87.57	99.95
ARX	(6, 6)	81.62	99.65
ARMAX	(6, 6, 6)	88.24	99.97
OE	(6, 6)	88.36	99.97

IV. CONCLUSION

The linearized lateral transfer functions of small fixed-wing UAV are used to form input (aileron and rudder) output (roll angle). The dynamic system model is formed as using ARX, ARMAX and OE system identification models and obtained data. The performance of dynamic systems is handled for different noise variances. The results of study show that estimated dynamic models are robust to noise and has minimum model complexity. The maximum estimation accuracy for different noise variances is obtained for OE model. With the help of the model obtained in this study, pre-flight analysis of fixed wing UAV can be performed. Thus, the risks of accidents of fixed wing UAV can be reduced. Thanks to the roll angle obtained with the OE model, the fixed wing UAV will be able to track the specified route with the minimum error. The proposed model will ensure that the payload is carried with high performance to the specified coordinates by tracking a certain route, even under external disturbance, and facilitates safe flight.

REFERENCES

- [1] A. Loya, M. Duraid, K. Maqsood, R. R. Khan. "The implementation and lateral control optimization of a UAV based on phase lead compensator and signal constraint controller." *Engineering*, vol. 10, 10, 2018.
- [2] P. D. Jameson, A. Cooke, "Developing Real-time System Identification for UAVs." 2012 UKACC International Conference on Control, 2012.
- [3] W. Saengphet, S. Tantrairatn, C. Thumtae, J. Srisertpol, "Implementation of System Identification and Flight Control System for UAV." 3rd International and Flight Control System for UAV, 2017.
- [4] A. Dorobantu, A. Murch, B. Mettler, G. Balas. "System identification for small, low-cost, fixed-wing unmanned aircraft." *Journal of Aircraft*, vol. 50, 4, 2013, pp.1117-1130.
- [5] V. Puttige, S. Anavatti. "Real-time system identification of unmanned aerial vehicles: A multi-network approach." *Journal of Computers*, vol. 3, 7, 2008.
- [6] W. Wei, M. B. Tischler, N. Schwartz, K. Cohen. "System identification and flight control of an unmanned quadrotor." *Advanced UAV Aerodynamics, Flight Stability and Control: Novel Concepts, Theory and Applications*, 2017, pp.695-727.

- [7] A. Altan, R. Hacıoğlu. "Model predictive control of three-axis gimbal system mounted on UAV for real-time target tracking under external disturbances." *Mechanical Systems and Signal Processing*, 2020.
- [8] A. Altan, R. Hacıoğlu, "Modeling of Three-axis Gimbal System on Unmanned Air Vehicle (UAV) under External Disturbances." 2017 25th Signal Processing and Communications Applications Conference (SIU), 2017.
- [9] A. Altan, Ö. Aslan, R. Hacıoğlu, "Model Predictive Control of Load Transporting System on Unmanned Aerial Vehicle (UAV)." Fifth International Conference on Advances in Mechanical and Robotics Engineering, 2017.
- [10] Y. H. Aleed, T. A. Tutunji, "RC Helicopter Modeling Using Re-engineering and System Identification." 14th International Multi-Conference on Systems, Signals & Devices (SSD), 2017.
- [11] C. Dube, J. O. Pedro. "Modelling and closed-loop system identification of a quadrotor-based aerial manipulator." *Journal of Physics: Conference Series*, vol. 1016.1, 2018.
- [12] J. E. Sierra, M. Santos. "Modelling engineering systems using analytical and neural techniques: Hybridization." *Neurocomputing*, 2018, pp. 70-83.
- [13] A. E. Ahmed, A. Hafez, A. N. Ouda, H. E. H. Ahmed, H. M. Abd-Elkader. "Modeling of a small unmanned aerial vehicle." *International Journal of Aerospace and Mechanical Engineering*, vol. 9, 3, 2015.
- [14] L. Ljung, *System Identification*, System Identification Theory for The Users, 1999, p. 609.

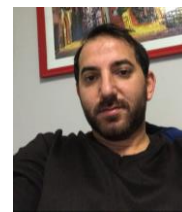
BIOGRAPHIES



EGEMEN BELGE received his B.Sc. and M.Sc. degrees both in Electrical Electronics Engineering from the University of Mersin in 2016 and 2018, respectively. He is currently Ph.D. student at Zonguldak Bülent Ecevit University, Zonguldak, since 2018 where he works as a research assistant.



HIZIR KAAN KABA received his B.Sc. degree in Mechatronics Engineering from the University of Işık in 2018. He is currently M.Sc. student at Zonguldak Bülent Ecevit University.



AHMET PARLAK received his B.Sc. degree in Electrical Electronics Engineering from the University of Zonguldak Bülent Ecevit University in 2016. He is currently M.Sc. student at Zonguldak Bülent Ecevit University.



AYTAÇ ALTAN received his B.Sc. and M.Sc. degrees in the department of Electrical Electronics Engineering from Anadolu University in 2004 and 2006, respectively. He is received his Ph.D. degree in the department of Electrical Electronics Engineering from Zonguldak Bülent Ecevit University in 2018. He is currently a lecturer at the department of Electrical Electronics Engineering at the Zonguldak Bülent Ecevit University in Turkey. His research interests include system identification, signal processing, image processing, model based control and robotic systems.



RIFAT HACIOĞLU received his B.Sc. degree in the department of Electrical Electronics Engineering from Dokuz Eylül University in 1993. He is received his M.Sc. and Ph.D., degrees in the department of Electrical and Computer Engineering from Illinois Institute of Technology, Chicago, USA, in 1996 and 2002, respectively. He is currently an Associate Professor at the department of Electrical Electronics Engineering at the Zonguldak Bülent Ecevit University in Turkey. His research interests are signal processing, control applications, linear and nonlinear system identification, parametric model estimation approaches, microprocessor based control systems, image identification problems.

Voltage Control of Cuk Converter with PI and Fuzzy Logic Controller in Continuous Current Mode

M. YILMAZ, M. F. ÇORAPSIZ and M. R. ÇORAPSIZ


Abstract— In today's energy systems, many equipment operates with Direct Current (DC) voltage. However, it is not always possible to obtain the voltage level required for the operation of these equipment from standard power supplies. For this reason, DC-DC converters are used to achieve the desired voltage values for equipment with different DC voltage levels. These converters are divided into three general categories, named Buck, Boost and Buck-Boost. The most preferred converter is the Cuk converter with low output ripple voltage, which can operate in both buck and boost modes. In this study, a detailed analysis of the Cuk converter, which is frequently used in Photovoltaic (PV) Panels was performed and different control methods of the output voltage were proposed. While performing this analysis, the dynamic model of the Cuk converter was created in which, Proportional-Integral (PI) and Fuzzy Logic (FL) are used to control the output voltage of the Cuk converter. The performances of both controllers were compared with respect to performance parameters such as steady state error, settling time and rise time. When the results obtained were evaluated as a whole, it was observed that FLC achieved the desired reference with less rise and settling time. In this study, modeling and controller applications of Cuk converter are realized by using MATLAB / SIMULINK program.

Index Terms— DC-DC Converter, Cuk Converter, Fuzzy Logic Controller, PI Controller, Voltage Control


I. INTRODUCTION

DC-DC converters are circuit topologies that convert the direct current voltage into a DC voltage at the desired


MEHMET YILMAZ, is with Department of Electrical-Electronics Engineering of Ataturk University, Erzurum, Turkey, (e-mail: mehmet.yilmaz@atauni.edu.tr).

 <https://orcid.org/0000-0001-7624-4245>

MUHAMMED FATİH ÇORAPSIZ, is with Department of Electrical-Electronics Engineering of Atatürk University, Erzurum, Turkey, (e-mail: corapsiz@atauni.edu.tr).

 <https://orcid.org/0000-0001-5692-8367>

MUHAMMED REŞİT ÇORAPSIZ, is with Department of Electric and Energy, Bayburt University, Bayburt, Turkey, (e-mail: rcorapsiz@bayburt.edu.tr).

 <https://orcid.org/0000-0001-5477-5299>

Manuscript received January 16, 2019; accepted April 2, 2020.
DOI: [10.17694/bajece.660025](https://doi.org/10.17694/bajece.660025)

voltage level [1]. The inputs of these converters are generally non-adjustable DC voltages, obtained by rectifying the line voltage. DC-DC converters are used to convert this non-adjustable voltage to the desired voltage level using appropriate switching techniques [2]. The DC converter is similar to the DC equivalent of a transformer with adjustable AC conversion ratio can be changed. It is used to achieve the desired voltage levels by decreasing or increasing the DC source voltage at the input of the DC transformers as well as by adjusting the voltage to the desired values when the transformers AC are applied [3]. The most important characteristics of these converters are their high efficiency and fast dynamic response. They are usually controlled by the DC-Pulse Width Modulation (DC-PWM) method [1]. The block diagram of the DC-DC converters is shown in Figure 1.

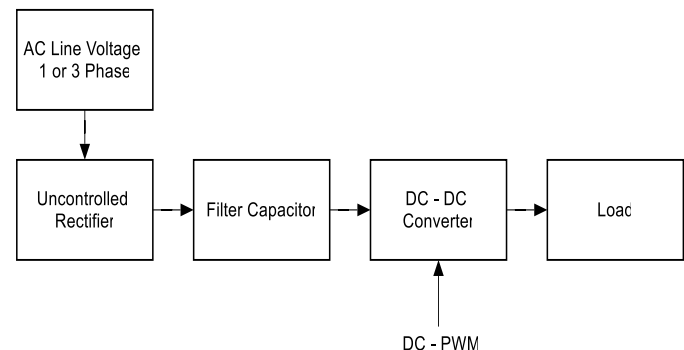


Fig. 1 DC-DC converter block diagram

Generally, DC-DC converters are used in DC motor drive applications, switched power supplies, marine cranes, electric cars, power factor correction applications and PWM based photovoltaic systems [1-3]. DC-DC converters are classified according to circuit topologies as buck converter, boost converter, buck-boost converter, SEPIC converter and Cuk converter etc. [3]. Buck and boost converter are the basic ones among those converters. The Cuk and buck-boost converter is obtained by the cascade connection of the buck and boost converter.

There are many studies on the Cuk converter in the literature. For example, Gupta and Lakshmi [4] designed a PI-controlled converter for photovoltaic panels. State-space modeling technique was used for continuous case modeling and Ziegler-Nichols method was used to determine PI parameters. The authors showed that, the output voltage reaches the desired

reference voltage after approximately 0.2 sec. Simulation studies were performed in MATLAB / SIMULINK program. Rakshit and Maity [5] designed a Cuk converter with closed loop fuzzy logic control. They applied PI and PID controller approaches to the same Cuk converter circuit structure and compared the obtained results. Settling time was 0.6 sec. for PID controller, 0.4 sec. for PI controller and 0.05 sec. for FLC. When the overshoot was compared, it was observed that the FLC controller performed better than the PI and PID controller. Boaretto et al. [6] modeled the Cuk converter for both continuous and discontinuous operating states with PWM switching methods and then compared both models. They suggested that the PWM switching method is suitable for continuous operation. Mohamed Assaf et al. [7] performed simulation studies of DC-DC converters. A dynamic analysis of the DC-DC converters was performed. Using the state equations obtained, the authors applied cascade controllers to the aforementioned converters. They observed that the settling time of the Cuk converter was 0.357 sec. and the output voltage percent was 1.96%. They used MATLAB / SIMULINK program in simulation studies. Fernão Pires et al. [8] designed a new non-isolated DC-DC converter topology for PV. This new topology was created by combining conventional DC-DC Cuk and boost converter circuits in which a single switch is used. The DC-DC converter they designed has higher static voltage gain than the conventional boost converter. This converter was designed using a fixed frequency PWM technique that can be associated with the MPPT algorithm. Besides Mohamed M. Algazar et al. [9] performed FLC for MPPT for PV, a new control method for Cuk converter was proposed. They studied this method under variable temperature and isolation conditions. As a result of the study, they suggested that the system with MPPT using FLC increases the efficiency of energy production from PV panels. Dileep and Singh [10] studied the selection of non-isolated DC-DC converters for PV systems. Consequently, comparative information about the characteristics of different isolated non-isolated DC-DC converters was presented. In addition, the authors investigated various research studies on MPPT PV-based DC-DC converters. They observed that the selection of the DC-DC converter had a significant effect on the overall performance of PV systems. Modeling and stability analysis of the closed loop current mode controlled Cuk converter was conducted by Kamran Mehran et al. [11]. Modeling and stability analysis were performed using Takagi-Sugeno (TS) fuzzy algorithm. Julio Cezar dos Santos de Moraes et al. [12] conducted a Cuk converter PV AC module with switched inductor structure. In order to develop a PV AC module, a high static gain Cuk converter structure was proposed with switched inductors. There are also several topologies in the literature based on the traditional CUK converter [13-15].

II. CUK CONVERTER AND DYNAMIC ANALYSIS

Cuk converters are electronic circuits that transfer the DC input voltage to the output at desired voltage levels. It was discovered by Slobodan M. Cuk [16]. Cuk converters were obtained by sequentially connecting the boost and buck converters. The most distinctive aspect of Cuk converters is

the use of capacitors for energy transfer [17]. The basic circuit diagram of the Cuk converter is shown in Figure 2.

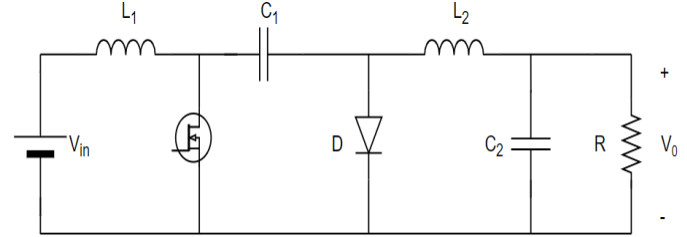


Fig. 2 Cuk converter basic circuit

The current and voltage waveforms for continuous current in a permanent state are shown respectively in Figure 3 and Figure 4 [3].

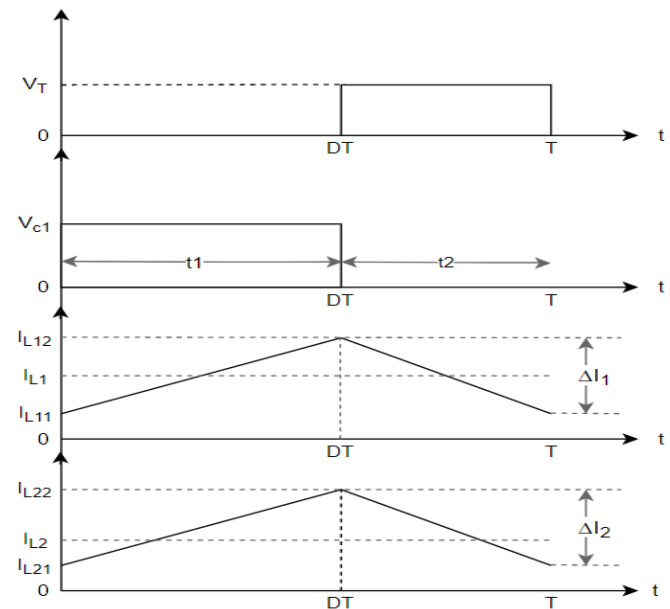


Fig. 3 Cuk converter waveforms

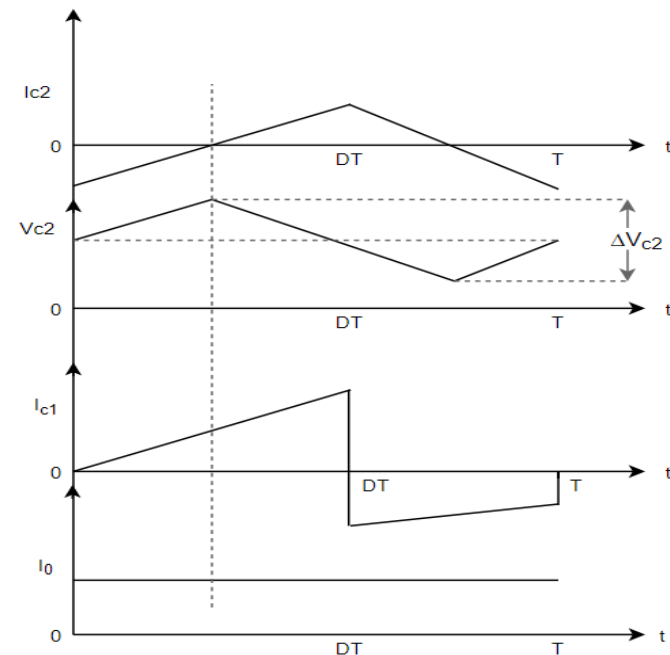


Fig. 4 Cuk converter waveforms

Before beginning, we give the nomenclature used in the study here:

- V_0 Output Voltage, V
 V_{in} Input Voltage, V
 V_{c1} Average Capacitor Voltage, V
 t_1 Time of switching element is closed, s
 t_2 Time of switching element is open, s
 T Switching Period, s
 D Duty Cycle
 ΔI Ripple Current, A

Assuming that the inductor current of inductance L_1 increases linearly from I_{L11} to I_{L12} during t_1 ,

$$V_{in} = L_1 \frac{I_{L12} - I_{L11}}{t_1} = L_1 \frac{\Delta I_1}{t_1} \quad (1)$$

Time expression from Equation (1),

$$t_1 = L_1 \frac{\Delta I_1}{V_{in}} \quad (2)$$

If V_{in} voltage is applied to input, capacitor C_1 starts to charge. If capacitor C_1 decreases the inductor current of inductance L_1 linearly from I_{L12} to I_{L11} during t_2 ,

$$V_{in} - V_{c1} = -L_1 \frac{I_{L12} - I_{L11}}{t_2} = -L_1 \frac{\Delta I_1}{t_2} \quad (3)$$

Time expression from Equation (3),

$$t_2 = -L_1 \frac{\Delta I_1}{V_{in} - V_{c1}} \quad (4)$$

Equation (2) and (4) in the ΔI_1 are equalized and if $t_1=DT$, $t_2=(1-D)T$ is the average capacitor voltage,

$$V_{c1} = \frac{V_{in}}{1-D} \quad (5)$$

Assuming that the inductor current of inductance L_2 increases linearly from I_{L21} to I_{L22} during t_1 ,

$$V_{c1} + V_0 = L_2 \frac{I_{L22} - I_{L21}}{t_1} = L_2 \frac{\Delta I_2}{t_1} \quad (6)$$

Time expression from Equation (6),

$$t_1 = L_2 \frac{\Delta I_2}{V_{c1} + V_0} \quad (7)$$

If the inductor current of inductance L_2 decreases linearly from I_{L22} to I_{L21} during t_2 ,

$$V_0 = -L_2 \frac{\Delta I_2}{t_2} \quad (8)$$

Time expression from Equation (8),

$$t_2 = -L_2 \frac{\Delta I_2}{V_0} \quad (9)$$

Equation (7) and (9) in the ΔI_2 are equalized and if $t_1=DT$, $t_2=(1-D)T$ is the average capacitor voltage,

$$V_{c1} = -\frac{V_0}{D} \quad (10)$$

Since the equation (5) and (10) are equivalent to each other, the average output voltage,

$$V_0 = -D \frac{V_{in}}{1-D} \quad (11)$$

As with other converters, the dynamic analysis of the Cuk converters is examined for two separate cases, namely the switching element is open and when it is closed. The equivalent circuit is shown in Figure 5 when the switching element is closed, and the equivalent circuit is shown in Figure 6 when the switching element is open.

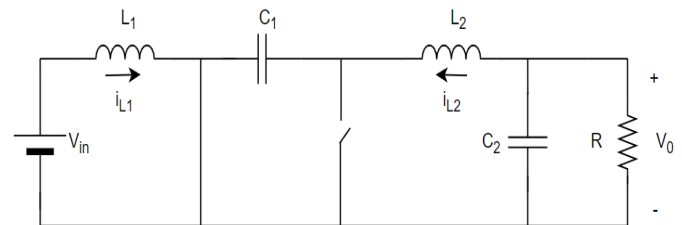


Fig. 5 Switching element is closed

When the switching element is closed, the circuit is examined in two parts. In the first case, the current from the source provides energy to the inductance. In the second case, the capacitor C_1 leads the diode to open and is discharged via capacitor C_2 , resistor R and inductance L_2 . The circuit model for the closed state of the switching element is obtained as shown in equations (12), (13), (14) and (15).

$$\frac{di_{L1}}{dt} = \frac{1}{L_1} V_{in} \quad (12)$$

$$\frac{dV_{c1}}{dt} = \frac{1}{C_1} (-i_{L2}) \quad (13)$$

$$\frac{di_{L2}}{dt} = \frac{1}{L_2}(-V_0 + V_{C1}) \tag{14}$$

$$\frac{dV_0}{dt} = \frac{1}{C_2}(i_{L2} - \frac{V_0}{R}) \tag{15}$$

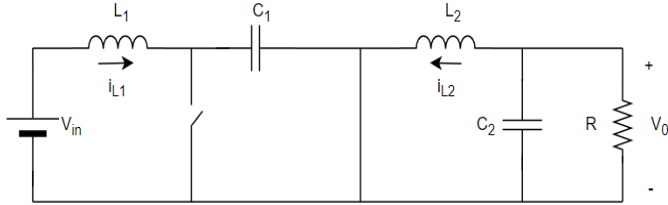


Fig. 6 Switching element is open

The current will flow through the diode when the switching element is open. At this time, the capacitor C_1 is charged via the source and inductance. The capacitor C_1 provides energy to the L_2 inductance, the C_2 capacitor and the R resistance. The circuit model for the open state of the switching element is obtained as shown in equations (16), (17), (18) and (19).

$$\frac{di_{L1}}{dt} = \frac{1}{L_1}(V_{in} - V_{C1}) \tag{16}$$

$$\frac{dV_{C1}}{dt} = \frac{1}{C_1}i_{L1} \tag{17}$$

$$\frac{di_{L2}}{dt} = \frac{1}{L_2}(-V_0) \tag{18}$$

$$\frac{dV_0}{dt} = \frac{1}{C_2}(i_{L2} - \frac{V_0}{R}) \tag{19}$$

III. PICONTRROLLER

PI controller is obtained by integrating proportional (P) and integral controller (I). It is the most commonly used controller type in industrial control systems due to its simplistic structure. [18]. A new control signal is generated by the PI controller based on the error value between the output signal and the reference signal. Then, the generated control signal is sent to the system and the operations are repeated until the steady state error is minimized. The transfer function of the PI controller is given in equation (20). The block diagram of the system controlled with the PI controller is shown in Figure 7.

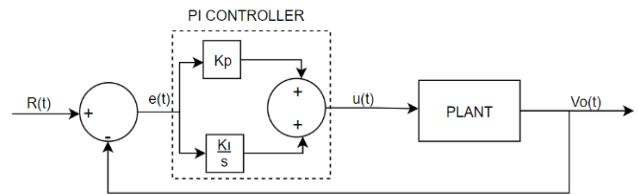


Fig. 7 PI Controller

$$u(t) = K_p e(t) + K_i \int e(t)dt \tag{20}$$

Where in K_i presents the integral gain and K_p is used for the proportional gain.

The controller is tuned optimally by using a trial and error method, and gains are found as $K_p = 0,04$ and $K_i = 4$.

IV. FUZZY LOGIC CONTROLLER

The most prominent feature of fuzzy control, which has become a major competitor to classical controllers in the field of control in recent years, is that it saves the designer from mathematical operations [19]. Fuzzy logic-based controllers can now be found in almost every area, ranged from automobile braking systems, washing machines, freezers to product quality control systems in factories [20]. In the classical types of controls (PI, PD, PID), a number of mathematical expressions need to be analyzed to design the controller. Although this process is simple for linear systems, it requires solution of difficult mathematical operations in nonlinear systems. However, when designing a FLC for any linear or non-linear system, there is no need to analyze the mathematical expressions. When using FLC, the components of the controller are prepared in systemically based on verbal expressions rather than a system of mathematical expressions [21].

A. Components of Fuzzy Logic Controller

The fuzzy logic controller, which was first used by Mamdani in 1974 [22], consists of basically three components: fuzzification, rule base and defuzzification. The FLC basic block diagram is shown in Figure 8.

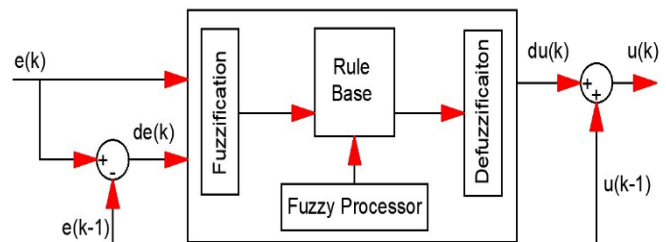


Fig. 8 Basic block of fuzzy logic controller

According to Fig. 8, the error (e) and change of error (de) refers to inputs of the system and, the control signal (u) refers to the output of the system. The relationship between error, change of error and output signal can be expressed as follows;

$$e(k) = V_{ref}(k) - V_{out}(k) \tag{21}$$

$$de(k) = e(k) - e(k - 1) \tag{22}$$

$$u(k) = du(k) + u(k - 1) \tag{23}$$

B. Fuzzification

The FLC has two definite inputs: error (e) and variation (de) are in the absolute number space. These precise entries are transferred to the rule base by converting them into fuzzy values with a degree of membership (μ) ranging from “0” to “1” in the blur unit. Triangle, trapezoid, bell, gaussian, cauchy, sinusoidal or sigmoid membership functions are used in the selection of the membership function. In this study, trapezoidal and triangular membership functions are used for five fuzzy levels.

C. Rule Base

In the rule base, the fuzzy rules, that are made by area experts in advance, are executed and fuzzy values are generated consequently.

If e is NS and de is PS then du is ZZ

In this rule base, “e” represents error, “de” represents error change and “du” represents degree of exit membership. In addition, the rules are defined as Negative Big (NB), Negative Small (NS), Zero (ZZ), Positive Small (PS) and Positive Big (PB). The operation of fuzzy control rules can be summarized as follows [23].

1. If the output of the system is lower and farther than the given reference point, that is, there is a large error in the positive direction, the controller must increase the output voltage rapidly.
2. If the output of the system is lower but close to the given reference point, that is, there is a small error in the positive direction, the controller should increase the output slowly.
3. If the output of the system is exactly at the given reference point, a meaning that there is no error, the controller should not interfere with the output.
4. If the output of the system is higher but close to the given reference point, that is, if the error is small in the negative direction, the controller should reduce the output slowly.
5. If the output of the system is higher and for away compered to the given reference point, that is, if the error is large in the negative direction, the controller should reduce the output rapidly.

D. Defuzzification

Finally, the fuzzy values are converted to exact values in a rinsing unit, just as in the input, and this exact value is sent to the output named the rinsing output (du). The control mark (u) is obtained by adding the previous value of the output to the defuzzification output. Membership functions and rule base used in this study are shown in Figure 9 and Table 1,

respectively. In the defuzzification process, there are different methods such as the center of the areas, the average of the maxima, Sugeno, Tsukamoto. When the central method of the areas is used in the defuzzification process;

$$u = \frac{\sum_{k=NB}^{PB} \mu_k(u_k) u_k}{\sum_{k=NB}^{PB} \mu_k(u_k)} \tag{24}$$

In Equation (24), k represents the active fuzzy set at the output and u represents the controller output. In addition, μ_k(u_k) is the degree of membership obtained from the active rule for fuzzy output, and u_k is the absolute output value with the largest membership in the active output fuzzy set in the same rule. In this study, Mamdani fuzzy inference system model type was used. In addition to, a limiter was used to keep the error values in the range of -1 to 1 at the input of the fuzzy logic controller. The control signal (u) obtained from the fuzzy logic controller was compared with a carrier signal similar to the sawtooth and, then the duty ratio of the controlled switch was determined.

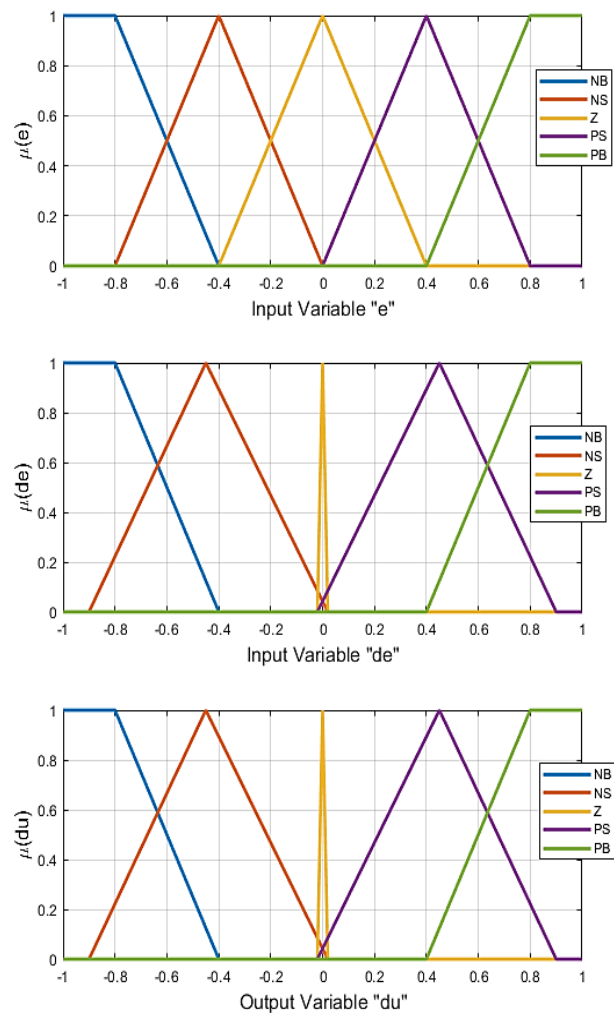


Fig. 9 Membership Functions

TABLE I
RULE BASE

		error				
		NB	NS	ZZ	PS	PB
de	NB	NB	NB	NB	NS	ZZ
	NS	NB	NB	NS	ZZ	PS
	ZZ	NB	NS	ZZ	PS	PB
	PS	NS	ZZ	PS	PB	PB
	PB	ZZ	PS	PB	PB	PB

V. SIMULATION RESULTS

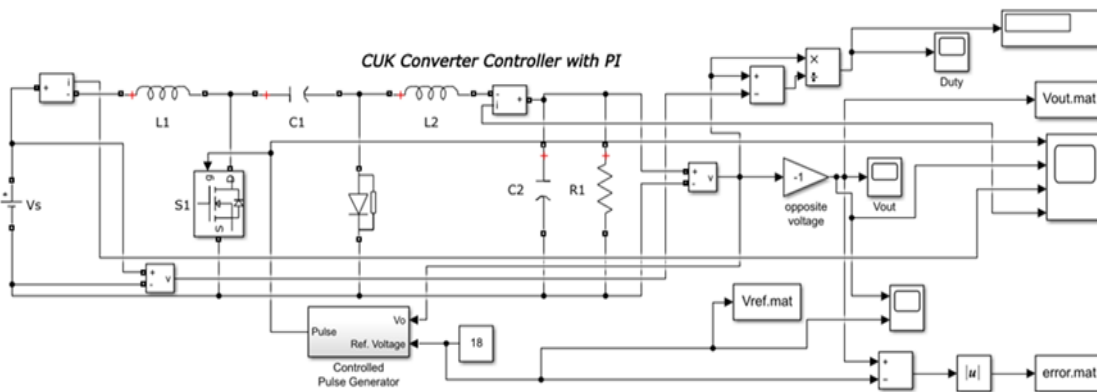


Fig. 10 Cuk converter simulation model controlled by PI controller

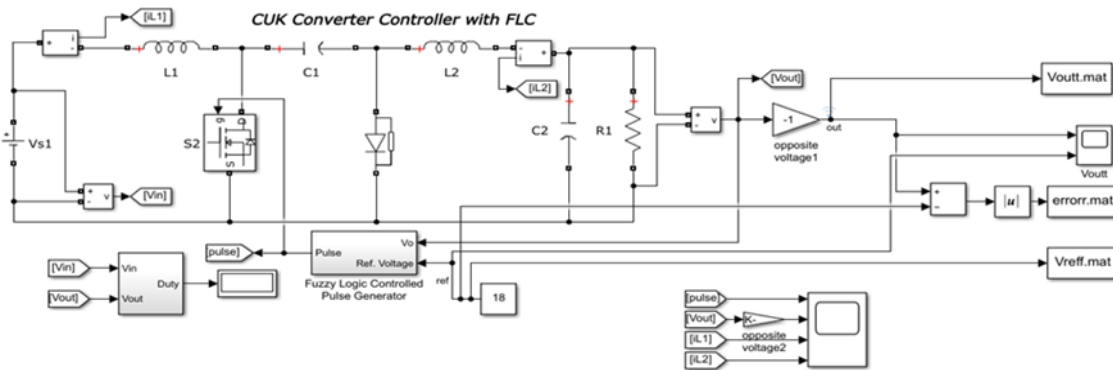


Fig. 11 Cuk converter simulation model controlled by FL controller

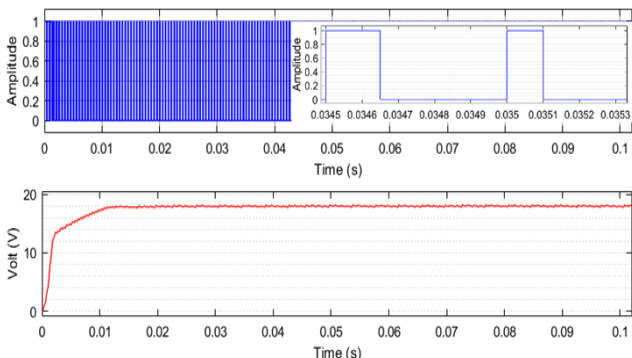


Fig. 12 Change of output voltage and duty period with PI controller

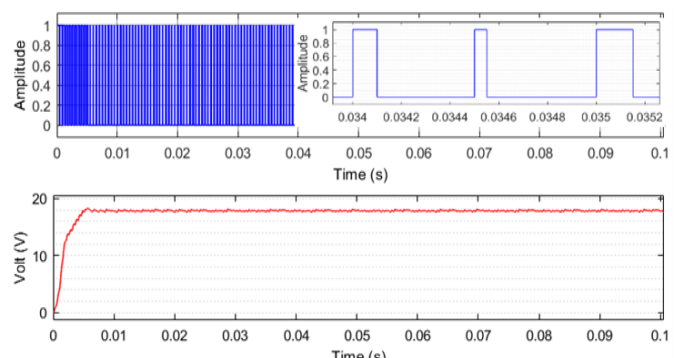


Fig. 13 Change of output voltage and duty period with FL controller

In this study, PI and FL controller is applied to Cuk converter circuit to obtain the desired output voltage. Table 2 shows the values of the parameters used in the Cuk converter simulation model. From the simulation results, the performance parameters such as steady state error, settling time and rise time of the controllers were obtained and a given in Table 3. Figure 10 shows the simulation model using the PI controller, and Figure 11 shows the simulation model using the FLC.

TABLE II
CUK CONVERTER CIRCUIT PARAMETERS

Input Voltage	12V
Output Voltage	-18V
Inductor Values L_1 & L_2	100 μ H & 100 μ H
Capacitor Values C_1 & C_2	150 μ F & 3300 μ F
Switching Frequency	2 kHz
Load Resistance	20 Ω

The change of output voltage and duty period with PI controller is applied to Cuk converter and shown in Figure 12 and the change of output voltage and duty period with FL controller is presented in Figure 13. In addition, if the PI controller is applied, the output voltage error graph is shown in Figure 14. Similarly, the error plot of the voltage at the converter output in FL controller is shown in Figure 15. A visual comparison of the output voltages obtained when both controllers are applied to the Cuk converter circuit is presented in Figure 16.

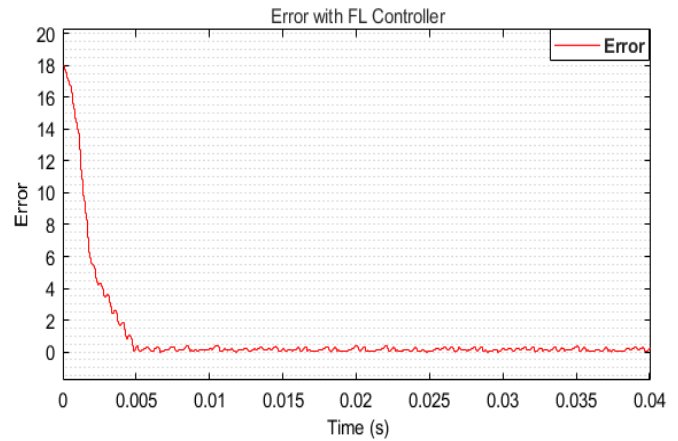


Fig. 15 The output voltage error controlled with FLC

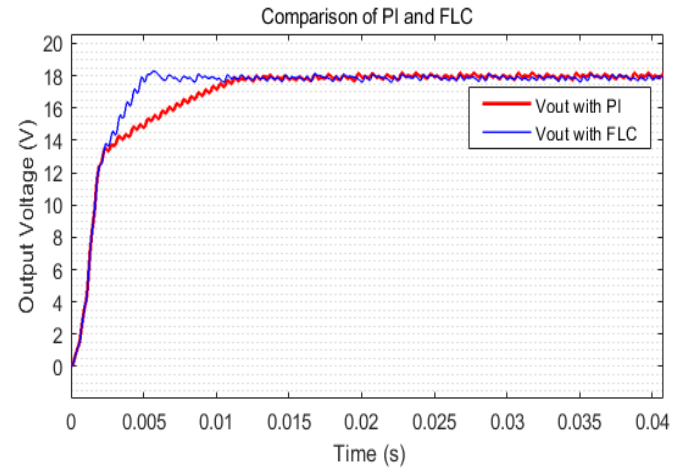


Fig. 16 Comparison of output voltages when both controllers are applied to the Cuk converter circuit

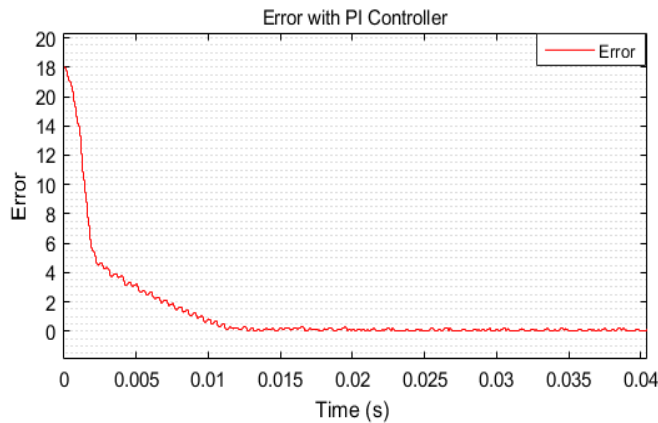


Fig. 14 The output voltage error controlled with PI controller

In Figure 12, it was observed that the Cuk converter output voltage reached a reference value of 18V after 10 ms without overshoot. In Figure 13, it is observed that the Cuk converter output voltage reaches a reference value of 18V after 5 ms. In Figure 14, it was observed that the error reached 0 after 10 ms when the PI control method was applied and when the FL control method was applied in Figure 15 the error reached 0 after 5 ms. In Figure 16, it is observed that FL control method gives better results than PI control method.

TABLE III
COMPARISON OF CONTROLLER

CONTROLLER	PI	FLC
Error_RMSE	0.7442	0.7235
Rise Time (ms)	6.55	3.6
Settling Time (ms)	10	5

VI. CONCLUSIONS

In this study, dynamic analysis and modeling of Cuk converter is done by using MATLAB / SIMULINK. PI and FL controller are used to follow the desired reference value. PI controller parameters $K_p = 0,04$ and $K_i = 4$ are found. The output voltage follows the reference after about 10 ms when the PI controller is used, and the reference follows after 5 ms when the FLC is used. While the PI controller is used, the ripple is %2,6 and the FLC is %2,5. As a future research, Cuk converter controlled in real time with PI, PID and FLC with PV panels can be analyzed.

REFERENCES

- [1] Bodur, H., (2012). Power Electronics. İstanbul: Birsen Education.
- [2] Mohan, N., Undeland, T. M., & Robbins, W. P. (2003). Power electronics: converters, applications, and design. John Wiley & sons.
- [3] Rashid M. H. R. (2015). Power Electronics Devices, Circuits and Applications. London: Pearson Education.
- [4] Gupta, Yelamarthi & P. Sri. (2014). Analysis and Design of CUK Converter using PI Controller for PV Application. International Journal for Scientific Research & Technology. 2. 2321-613.
- [5] Rakshit, Saptarshi & Maity, Jayabrata. (2018). Fuzzy Logic Controlled Cuk Converter. 0771-0775. 10.1109/ICCSP.2018.8524168.
- [6] Boaretto, Fernanda & Junior, João & Marca, Ygor & Santos Dias de Moraes, Paulo Mario Dos & Kirsten, André. (2018). Small-Signal Modelling of the Cuk Converter. 10.13140/RG.2.2.17307.16164.
- [7] Assaf, Mohamed & Seshsachalam, D. & Chandra, Dinesh & Tripathi, Ramesh. (2005). DC-DC converters via matlab/simulink. 464-471.
- [8] Pires, V. & Foito, Daniel & Baptista, F.R.B. & Silva, Fernando. (2016). A photovoltaic generator system with a DC/DC converter based on an integrated Boost-Cuk topology. Solar Energy. 136. 1-9. 10.1016/j.solener.2016.06.063.
- [9] Alazar, Mohamed & AL-monier, Hamdy & EL-halim, Hamdy & Salem, Mohamed. (2012). Maximum power point tracking using fuzzy logic control. International Journal of Electrical Power & Energy Systems. 39. 21-28. 10.1016/j.ijepes.2011.12.006.
- [10] G, Dileep & Singh, S.N.. (2017). Selection of non-isolated DC-DC converters for solar photovoltaic system. Renewable and Sustainable Energy Reviews. 76. 1230-1247. 10.1016/j.rser.2017.03.130.
- [11] Mehran, Kamyar & Giaouris, Damian & Zahawi, Bouchaib. (2009). Modeling and stability analysis of closed loop current-mode controlled Cuk converter using Takagi-Sugeno fuzzy approach. IFAC Proceedings Volumes. 42(7):223-228
- [12] Cezar, Julio & Luiz, Juliano & Gules, Roger. (2018). Photovoltaic AC-Module Based on a Cuk Converter with a Switched-Inductor Structure. IEEE Transactions on Industrial Electronics. PP. 1-1. 10.1109/TIE.2018.2856202.
- [13] Tiwari Neeraj, Bhagwan Das D. MPPT controller for photo voltaic systems using Cuk DC/DC convertor. International Journal of Advanced Technology and Engineering Research (IJATER).
- [14] Alazar MohamedM, et al. Maximum power point tracking using fuzzy logic control. Int J Electr Power Energy Syst 2012; 39(1): 21-8.
- [15] Chen Zengshi. PI and sliding mode control of a Cuk converter. IEEE Trans PowerElectron 2012; 27(8): 3695- 703.
- [16] J. A. M. Bleijns and J. A. Gow, "Fast maximum power point control of current-fed DC-DC converter for photovoltaic arrays", Electornic Letters, Vol. 37, No. 1, January 2001, pp. 5-6.
- [17] Singh MD., (2008). Power Electronics, Tata Mc Graw-Hill Education.
- [18] Corapsiz, M. F., & Erenturk, K. (2015). Trajectory tracking control and contouring performance of three-dimensional CNC. IEEE Transactions on Industrial Electronics, 63(4), 2212-2220.
- [19] Çorapsız, M. R.. "Performance Analysis of Speed Control of PMDC Motor using Fuzzy Logic Controller". Eastern Anatolian Journal of Science 3 / 2 (Kasım 2017): 16-29.
- [20] Çorapsız, M. R., Reduction of commutation torque ripples in brushless direct current motors, Karadeniz Technical University, Graduate Institute of Natural and Applied Sciences, Trabzon, 2018.
- [21] Kahveci, H., The implementation of an electronic differential system based on fuzzy logic for direct driven electric vehicles, Karadeniz Technical University, Graduate Institute of Natural and Applied Sciences, Trabzon, 2013.

- [22] Mamdani, E.H. ve Assilian, S., An Experiment in Linguistic Synthesis with A Fuzzy Logic Controller, International Journal of Man-Machine Studies, 7,1 (1975) 1-13.
- [23] DURANAY, Zeynep Bala, GULDEMİR, Hanifi . "Study of Fuzzy Logic Control of De-Dc Buck Converter". Firat University Turkish Journal of Science and Technology 12 / 2 (Ekim 2017): 23-31.

BIOGRAPHIES



MEHMET YILMAZ was born in Trabzon, Turkey, in 1991. He received the first B.S. and the M.S. degree from Karadeniz Technical University, Trabzon, Turkey, in 2014, 2018, respectively, all in electrical- electronics engineering. Since 2018, he studies Ph.D. in Electrical-Electronics Engineering at the Ataturk University. He has been a member of the Chamber of Electrical Engineers in Turkey. Currently, he is a Research Assistant, Ataturk University, Erzurum, Turkey. His research interests include renewable energy, power electronics and electric drive systems, control of electric machinery, electric vehicles, smart grid and power management in electric vehicles.



M. FATİH ÇORAPSIZ was born in Erzurum, Turkey, in 1981. He received the first B.S. degree from Firat University, Elazığ, Turkey, in 2003, and the M.S. and Ph.D. degrees and the second B.S. degree from Ataturk University, Erzurum, Turkey, in 2009, 2014 and 2016, respectively, all in electrical engineering. He has been a member of the Chamber of Electrical Engineers in Turkey.

Currently, he is a Assistant Professor with the Department of Electrical and Electronics Engineering, College of Engineering, Ataturk University, Erzurum, Turkey. His research interests include theory of mechatronic and robotic systems, DC-DC converters, and motor drive circuits, with a focus on observation and estimation-based control.

M. REŞİT ÇORAPSIZ was born in Erzurum, Turkey, in 1984. He received the first B.S. degree from Firat University, Elazığ, Turkey, and the second B.S. degree from Ataturk University, Erzurum, Turkey, and the M.S. degree from Karadeniz Technical University, Trabzon, Turkey, in 2009, 2015, 2018, respectively, all in electrical engineering. Since 2018, he studies Ph.D. in electrical engineering at the Karadeniz



Technical University. He has been a member of the Chamber of Electrical Engineers in Turkey.

Currently, he is an Instructor with the Department of Electrical and Energy, Vocational School of Technical Science, Bayburt University, Bayburt, Turkey. His research interests include power electronics and electric drive systems, control of electric machinery, electric vehicles, smart grid and power management in electric vehicles.

Determination of the roll-off value in the air-gapped inductor of a DC-DC boost converter circuit with FEA parametric simulations

P. ARIKAN, S. BALCI and F. BATTAL


Abstract—The electromagnetic behavior of the inductors used as passive circuit elements directly affects the electrical and mechanical performance of the power electronics circuits. In general, when using inductor core structures with/without air-gap length in the classical design process, the dynamic effects of the inductance value are not considered in the design stage. However, the inductance value may change during the operation of the circuit due to electrical and magnetic parameters of the inductor, and this change is called roll-off value of the inductance. In this study, the roll-off value has been determined graphically and numerically based on mechanical parameters (such as air-gap length) and electrical parameters (such as winding turns and DC current amplitude) for an air-gapped ferrite E core designed with finite element analysis (FEA) software. Thus, not only the inductance value has been calculated in the design stage but also the roll-off value during the operation of the circuit has been reported with the parametric simulation studies.

Index Terms—FEA parametric simulations, gapped-core, inductor parameters, roll-off value.


I. INTRODUCTION

NOWADAYS, advanced power electronics application area is growing with the tendency to use renewable energy sources and they are frequently used in daily life with current applications such as electric vehicles [1]. In this context, with the advances in semiconductor technology, both power electronics circuit topologies are developing, and


PINAR ARIKAN, is with Department of Engineering Sciences, Karamanoğlu Mehmetbey University, Karaman, Turkey, (e-mail: pnryrky@gmail.com).

 <https://orcid.org/0000-0002-7974-9289>

SELAMİ BALCI, is with Department of Electrical and Electronics Engineering, Karamanoğlu Mehmetbey University, Karaman, Turkey, (e-mail: sbalci@kmu.edu.tr).

 <https://orcid.org/0000-0002-3922-4824>

FUNDA BATTAL, is with Department of Electronics and Automation Nevşehir HacıBektaş Veli University, Nevşehir, Turkey, (e-mail: fundacodur@gmail.com).

 <https://orcid.org/0000-0002-7233-2843>

Manuscript received December 23, 2019; accepted Mar 3, 2020.

DOI: [10.17694/bajece.664044](https://doi.org/10.17694/bajece.664044)

innovations are provided in magnetic circuit elements in these topologies [2]. The design concept of the magnetic circuit elements has been revised in order to provide smaller size and less power losses [3] with the development of the soft magnetic materials. Especially in the modern design approach, the most suitable design can be reached shortly before the prototype production with the development of software that enables electromagnetic modeling by finite element method, and the revolution in solver performance in computer hardware.

The inductors used in DC-DC power converter circuits are known as DC inductors and are exposed to the high frequency ripple on the DC current. The flux density in the inductor core appears to be triangular ripple wave depending on the peak value of the current, and the inductance value is dynamically similar to this ripple wave [4-6]. In the design of the air-gapped inductor, the analysis of bad effects such as the fringing flux effects during the determination of the air-gap length can be expressed mathematically in classical design approach. However, for the effects of the air-gap parts on the dynamic inductance and roll-off values of the inductor, electromagnetic modeling facilitates the design process [7].

The software used for electromagnetic modeling with finite element analysis (FEA) is often used to determine the electromagnetic, thermal and cooling performance of the inductor before prototype production. In context, Dang *et al.* [8] present the electromagnetic modeling, FEA simulation and design stage of a high-power inductor for battery charge system in order to reduce the core volume for electric vehicle. For the thermal coupled FEA, Du *et al.* [9] present in order to determine an inductor on both its electromagnetic and thermal behavior. Munguia *et al.* [10] explain how designers can take advantage of the useful features of electromagnetic FEA software to quickly model nonlinear behavior inductors and determine their performance. In addition, software developed in recent years, such as Ansys-Electronics, enables the analysis of both electromagnetic components and power electronics circuits together for the inductor design in power electronics circuits.

In this study, roll-off value of the ferrite core inductor has been determined with the parametric FEA software based on mechanical and electrical parameters for air-gapped E core designed for a DC-DC boost converter circuit. The graphs of the incremental inductance changes based on parametric DC current have been obtained with Ansys-Electronics Desktop

2019R3, and roll-off values of the designed inductor have been determined. In addition, flux distributions in the inductor core have been visualized and saturation effect has been examined. Thus, not only the inductance value calculated during the design phase, but also the dynamic inductance value during the operation of the circuit has been reported with the simulation studies.

II. THEORETICAL ANALYSIS OF THE INDUCTANCE VALUE IN INDUCTORS WITH AIR-GAPPED CORE

In order to obtain the desired inductance value in air gap inductor design, Eq.(1) can be written according to the initial permeability value for effective permeability (μ_{eff}) [11-12].

$$\mu_{eff} = f(\mu_i, H_c, B_{sat}) \quad (1)$$

The magnetic field intensity (H_c) and saturation flux density (B_{sat}) values and initial permeability (μ_i) values of the core material are very important, and this information can be reached from the core material datasheet. The basic parameters and the magnetic equivalent circuit for a given air-gapped core are given in Fig. 1. Here, the magnetomotor force NI as flux and the reluctances in the path of flux are core and air gap resistance elements as \mathfrak{R}_c and \mathfrak{R}_g , respectively [13].

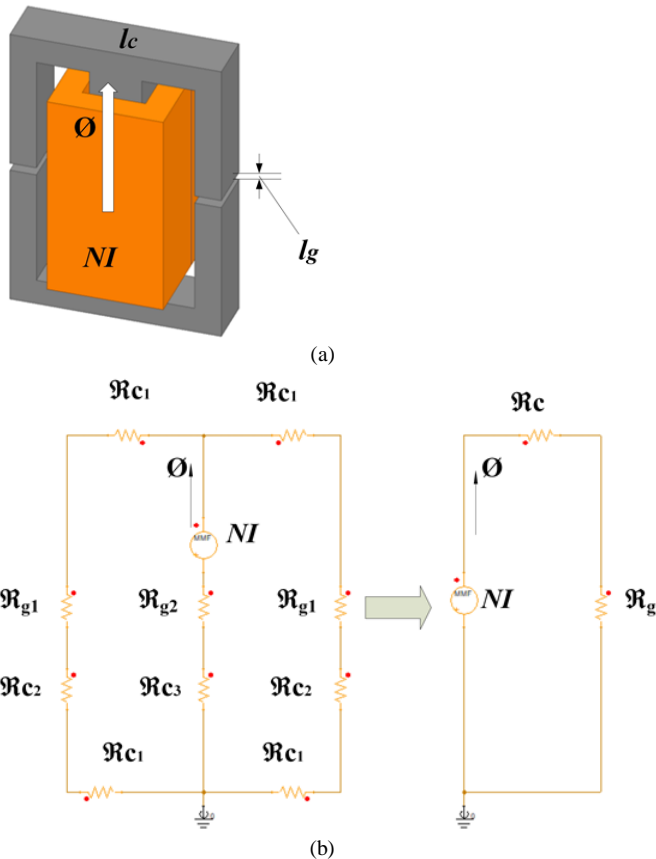


Fig.1. For the air gapped inductor, (a) basic magnetic parameters and (b) equivalent magnetic circuit [14]

As the ferrite N87 ferrite core material [15] can be seen in the B-H curve given in Fig. 2, the saturation is delayed by the air gaps in the core in large current (NI_{sat2}) inductor designs to provide a soft saturation flux characteristic. However, although saturation can be delayed by air gaps, sharp saturation occurs only after a certain current value in materials such as ferrite, amorphous, nanocrystalline and Si-Fe. In powder materials such as Kool M μ , MPP and XFlux, there is no need to determine the air gap length in the core structures and a soft saturation occurs with the distributed air gap feature [16-17].

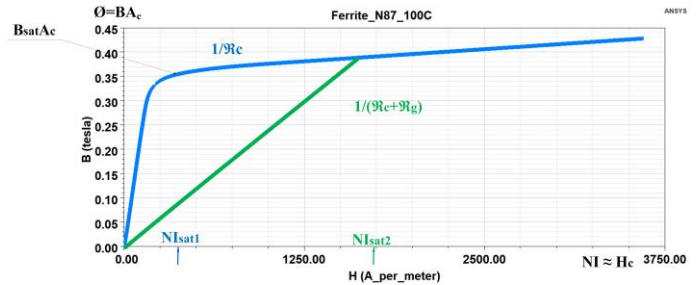


Fig.2. B-H curve for ferrite core material, and soft saturation effect with air-gapped core structure [11, 15]

According to Ampere's law, the relationship between air gap effect and closed-circuit magnetic field can be expressed in Eq.(2) [11].

$$\oint_{closed-path} H \cdot dl = F_c + F_g \quad (2)$$

The magnetomotor force can be modeled as an electrical circuit with the Amper-Winding (NI) value of the inductor and the reluctance values in the air-gapped inductor core (Fig. 1). Thus, Eq.(3) can be written for the flux (Φ) circulating in the core [11].

$$N \cdot I = \Phi(\mathfrak{R}_c + \mathfrak{R}_g) \quad (3)$$

In order to calculate the core and air gap reluctance values, in Eqs. (4-5), since the permeability of the air is $\mu_0 = 4\pi 10^{-7}$ H/m, very high reluctance occurs in the air gap parts of the inductor core. Since the core permeability μ_c is very large compared to the gap, the reluctance of the core remains at very small values [16-17].

$$\mathfrak{R}_c = \frac{l_c}{\mu_c A_c} \quad (4)$$

$$\mathfrak{R}_g = \frac{l_g}{\mu_0 A_g} \quad (5)$$

The inductors are energy storage magnetic circuit elements and the stored energy (W) can be expressed as in Eq. (6) [18].

$$W = \frac{1}{2} LI^2 = \frac{1}{2} \oint_v \vec{H} \cdot \vec{B} dv \quad (6)$$

where: v - is the volume of the inductor core.

The magnetic field strength is equal to the energy calculated by the closed-loop inductance and the square of the current as H and flux density B . The Faraday equation given in Eq.(7) can be written to determine the flux value in the core [11].

$$V(t) = -n \frac{d\Phi}{dt} = -L \frac{di}{dt} \quad (7)$$

From this, Eq.(8) can be written proportional to the square of the number of turns for the induced voltage due to the equivalent reluctance [11].

$$V(t) = \frac{N^2}{\mathfrak{R}_c + \mathfrak{R}_g} \cdot \frac{di}{dt} \quad (8)$$

Thus, for the inductance value and saturation flux value in air-gapped inductor design, Eq.(9) and Eq.(10) can be given [19].

$$L = \frac{N^2}{\mathfrak{R}_c + \mathfrak{R}_g} \quad (9)$$

$$\Phi_{sat} = B_{sat} \cdot A_c \quad (10)$$

In this situation, the current value which can cause saturation of the air-gapped inductor core can be determined by Eq. (11) [11].

$$I_{sat} = \frac{B_{sat} \cdot A_c}{N} (\mathfrak{R}_c + \mathfrak{R}_g) \quad (11)$$

Thus, the effective permeability value of the core structures designed with the air gap can be determined by the mean length of the core, the permeability of the inductor core and the air gap length as given by Eq.(12) [7, 20].

$$\mu_{eff} = \frac{l_c}{\frac{l_c}{\mu_c} + l_g} \quad (12)$$

Furthermore, the required air gap length Eq.(13) may be calculated in the inductor designed with air gap for a given core cross sectional area (A_c) and the desired inductance value (L) [7].

$$l_g = \left(\frac{0.4\pi N^2 A_c (10^{-8})}{L} \right) - \left(\frac{l_c}{\mu_c} \right), [cm] \quad (13)$$

However, depending on the air gap length, there are deviations in the inductance value due to the fringing flux

effect. It is therefore more convenient to use multiple air openings and distributed air-gap parts instead of leaving one large air-gap part in the core. The core window height (G) or winding length has a great effect on the fringe flux. Accordingly, the fringing flux factor (F) can be calculated by Eq.(14) [20].

$$F = \left(1 + \frac{l_g}{\sqrt{A_c}} \ln \frac{2G}{l_g} \right) \quad (14)$$

Since the fringing flux effect reduces the equivalent reluctance value in air-gapped core structures, an inductance value is higher than the inductance value calculated in Eq.8. Therefore, the inductance value in the proportion to the fringing flux factor increases as shown in Eq.(15) [7, 20].

$$L = F \left(\frac{0.4\pi N^2 A_c \cdot 10^{-8}}{l_g + \frac{l_c}{\mu_c}} \right) \quad (15)$$

III. INDUCTOR DESIGN PARAMETERS WITH THE AIR-GAPPED FERRITE CORE

The variables that should be considered in the design process as electrical and mechanical variables in air gap inductors for power electronics circuits are given in Table 1. In order to determine the dynamic inductance value according to this table, a number of specific parameters such as DC-DC boost converter switching frequency and inductor current ripple, which are designed for a certain power value, were utilized.

TABLE I
Technical specifications for DC-DC boost converter inductor [2]

Power Density	5 kW
Input Voltage (V_i)	200 V
Output Voltage (V_o)	400 V
Load Current (I_o)	12.5 A
Inductance (L)	1 mH
Inductor current (I_L)	25 A/ 22.5A-27,5A
Ripple Magnitude (ΔI_L)	5 A
Switching Frequency (f_s)	20 kHz
Duty Ratio (D)	0.5
Core Material	Ferrit 3C90
Core Structure	E1006028

In order to maintain the continuous conduction mode (CCM) [21-22] in DC-DC boost converter circuits, the critical inductance value of inductors should not fall below the critical value in the equation given by Eq.(16) [23].

$$L_c = L = \frac{D(1-D)R_L}{2f_s} \quad (16)$$

where: D - is defined as the switching duty ratio;
 R_L - is load resistance value;
 (f_s) - is the switching frequency.

For this purpose, before the inductor prototype, parametric DC current test can be performed to determine the inductance stability. Thus, the inductance stability or roll-off values can be obtained. Although the roll-off values for core materials are specific in the manufacturer's datasheets, inductance stability has nonlinear behavior and the inductance stability may vary from classical mathematical methods to the operation of the circuit. If the inductance value starts to collapse well below the desired current value in the design stage, the electrical and mechanical parameters must be changed. For example, increasing the air gap length and changing the number of turns. However, such revisions are both costly and unnecessarily prolong the design process. In such cases, 3D electromagnetic modeling enables the inductor to be modeled in a virtual environment close to real conditions.

The theoretical flux distribution of the DC-DC boost converter inductor for the dynamic inductance in the core can be defined by the dynamic minor loop B-H curve as shown in Fig. 3.

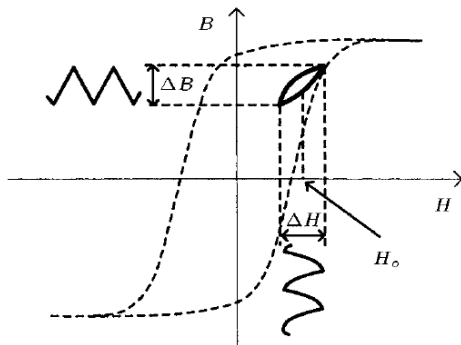


Fig.3. Dynamic minor loop of the inductor flux density [11, 23-24]

In this curve, for the flux density ripple (ΔB), Eq.(17) can be explained under ripple superimposed DC excitation. The DC current (I_o) and the number of turns (N) multiplied by the average magnetic circuit length (l_c) and Eq.(18), the magnetic field strength fluctuation (ΔH) was obtained [24-25].

$$\Delta B = \frac{1}{NA_c} \int_0^{DT_s} v_L dt \tag{17}$$

$$\Delta H = \frac{NI_o}{l_c} \tag{18}$$

Where D is the duty ratio of the power electronics circuit, and T_s is the switching signal period. Thus the DT_s is the conduction period of the power switches.

IV. PARAMETRIC SIMULATION STUDIES WITH FEA SOFTWARE

The parametric simulation studies have been performed with linear steps using FEA software according to electrical (N and I_{dc}) and mechanical variables (l_g) given in Table 2. Thus, the dynamic inductance value of an inductor designed with E1006028 core can be easily determined. In addition, in order to ensure inductance stability, the air gap length can be optimized mechanically in order to prevent saturation at a certain number of turns and DC current value.

TABLE II
 Electrical and mechanical parameters of air gap

Parameters	Range	Linear Step
l_g	0-3 mm	0.50 mm
N	50-80 turns	5 turns
I_{dc}	0-40 A	2 A

According to these values, parametric simulation studies were carried out with Ansys-Electronics Desktop 2019.R3 software in Maxwell 3D section. For this purpose, E1006028 core material was modeled in three dimensions according to datasheet content [15] and B-H values of ferrite core material were defined in the software library.

A three-dimensional graph based on the parametric DC current increase and winding number values over a given air gap length provides useful information for the designers to determine the inductance value (Fig. 4). According to this graph, DC current value winding number and saturation effect can be seen clearly. The green zones in the graph can be defined as acceptable roll-off values (60%) for the desired inductance value. Thus, the parametric inductance values have been determined with the 3D electromagnetic modeling and more realistic conditions [26].

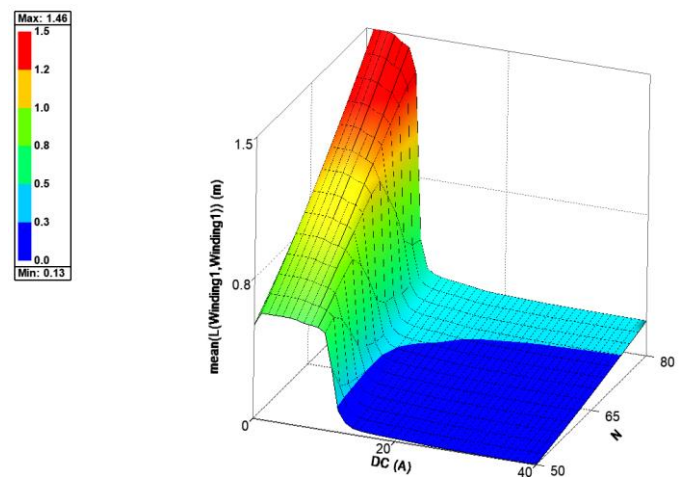


Fig.4. Parametric 3D inductance change graph of designed ferrite core with air gap for maximum 1.5 mH

The flux linkage graph, which is defined as winding flux in inductors, shows the electromagnetic behavior of the airgap designed core. According to this graph given in Fig. 5, the green areas for acceptable roll-off values were determined as 0.01-0.015 Weber/m². This condition, when visualized as the useful flux circulating in the core, does not cause any saturation for ferrite core material with about 0.2 Tesla. For

ferrite core material, the saturation flux value is given as 0.36 T in the datasheet [15], and the 0.2 T flux value for the 20 kHz switching frequency is very useful.

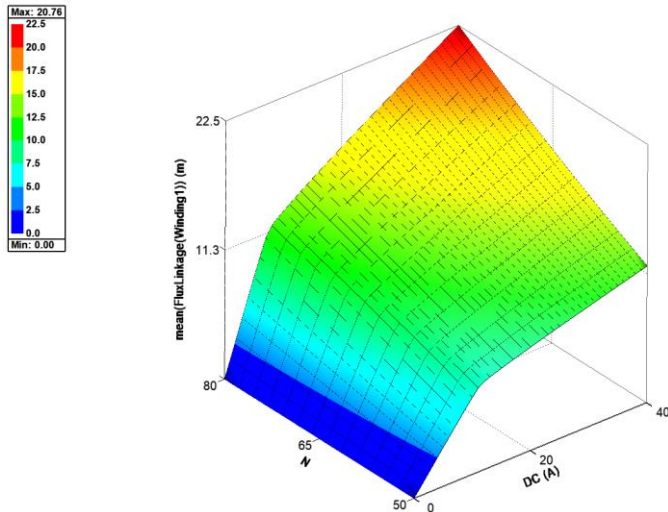


Fig.5. Parametric 3D flux linkage graph of ferrite core designed with the air gap for maximum 22.5 mWeber/m²

When the air gap length is changed, the number of turns of the inductor windings is also changed to achieve the desired inductance value. For example, in the case of 1.5 mm air-gap in l_{g2} , the number of turns is 57 turns, whereas in the case of $l_{g1} = 3$ mm air-gap, 72 turns are sufficient according to the simulation studies. These values are obtained in parametric 3D graphics. Thus, as can be seen in the roll-off graphs given in Figure 6, the dynamic inductance values have been determined according to the air gap lengths l_{g1} and l_{g2} . After 10 A with dc current increase, roll-off in $l_{g2} = 1.5$ mm air gap core decreases to 40% and inductance value decreases too much. However, in the case of $l_{g1} = 3$ mm air gap, the roll-off has been determined as 60% and the inductance value is stable even after a current of 10 A DC, thus providing a soft saturation effect.

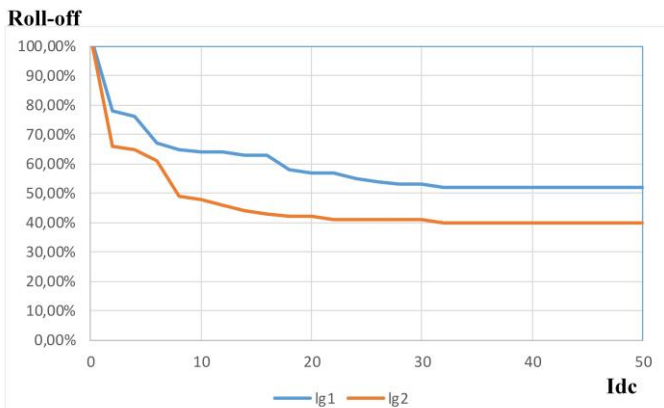


Fig.6. Roll-off values due to DC current increase for ferrite core designed with different air gap length

V. POWER ELECTRONICS CIRCUIT PERFORMANCE ANALYSIS WITH FEA CO-SIMULATION

The link in the simulation circuit Ansys-Twin Builder software given in Fig. 7 is defined with Maxwell 3D software,

for the power electronics circuit performance test of the inductor electromagnetic modeling and inductance stability and roll-off values are determined in dynamic behavior. Thus, both the realistic behavior of the designed inductor and the performance of the power electronics circuit can be determined.

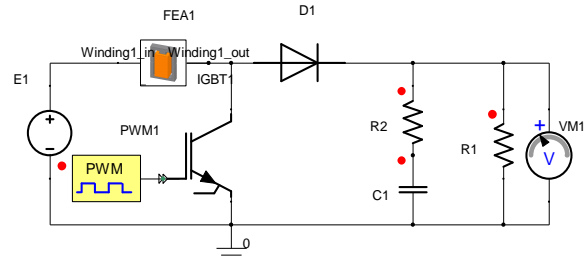


Fig.7. The co-simulation circuits of DC-DC Boost Converter with air-gapped inductor

The inductor current graph obtained according to the co-simulation results is shown in Fig. 8 with the ripple value given in Table 1. Thus, the performance of the designed inductor is revealed, and a continuous transmission mode can be provided.

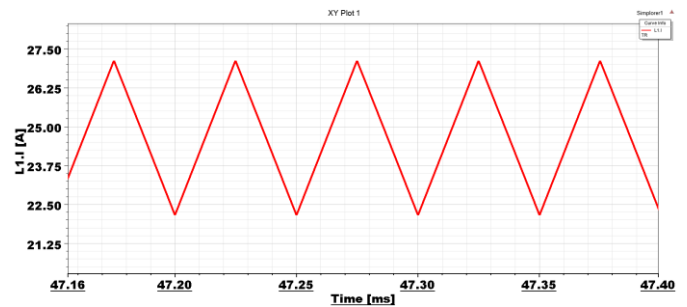


Fig.8. Inductor current waveform and ripple magnitude

The dynamic inductance graph of the modeled inductor is given in Fig. 9. Thus, approximately 60% of the roll-off of the inductor in the power electronics circuit is also confirmed with the co-simulation.

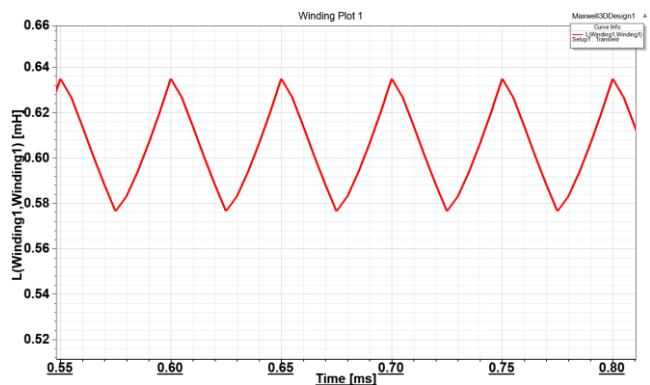


Fig.9. Dynamic behavior of the inductor in the power electronics circuit

The flux distribution in the core of the inductor designed with 3 mm air gap and tested for electromagnetic behavior in the power electronics circuit is given in Fig. 10. Thus, the flux value did not approach the saturation flux value and was displayed as approximately 0.2 T thanks to the air gaps.

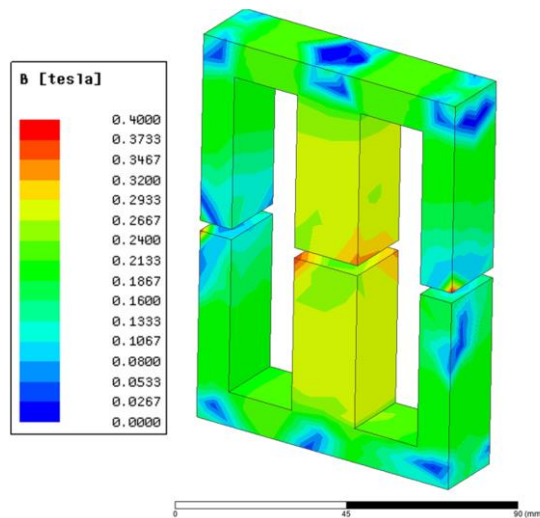


Fig.10. Flux distribution in air-gap inductor core for maximum 0.4 T

Finally, the output voltage performance of the power electronics circuit designed as air gap inductors remains within acceptable limits in terms of the ripple value as shown in Fig. 11.

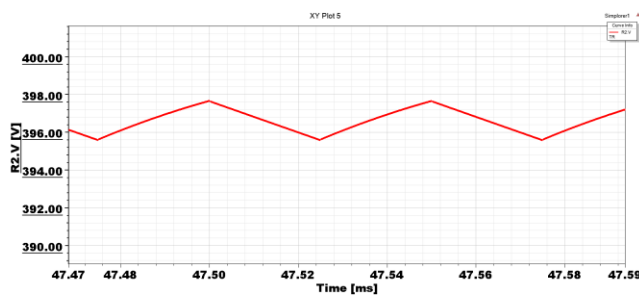


Fig.11. The output voltage waveform of converter

VI. CONCLUSION

The inductors as passive circuit elements of DC-DC converter circuits are extremely important in terms of providing continuous conduction mode with the critical inductance value. In the classical design approach, the dynamic behavior of air-gapped inductors is determined according to both electrical and mechanical parameters. Also, before the prototype production, roll-off values can be seen with DC incremental tests. In this study, the parametric simulation and electromagnetic modeling of the inductors depending on number of turns, DC current and the air-gap length have been realized using FEA software. Thus, the dynamic analysis of the inductance and performance analysis in the power electronics circuit has been reported. As a result, if the inductance roll-off values are not within the acceptable limits, the most suitable value of the air-gap length can be easily determined in FEA modeling and the more realistic dynamic inductance behaviors can be seen with the co-simulations for DC-DC boost converter circuit.

The importance of this study explains the determination of air-gapped core inductor behavior before the prototype realization studies according to the theoretical information

given for the inductor design and 3D electromagnetic modeling studies. As future studies, performance investigation can be done with experimental studies of an E1006028 core inductor. Also, the inductor discussed in this study is for medium frequency and high power applications. For this reason, the power electronics circuit is modeled for medium frequency (20 kHz) and kW power levels.

REFERENCES

- [1] M. Ashraf, N. Masoum. "High efficiency boost converter with variable output voltage using a self-reference comparator." *AEU - International Journal of Electronics and Communications*, vol. 68. 11, 2014, pp 1058-1064.
- [2] S. Balci. "A CFD simulation of the liquid-cooled pipe conductors for the high-power and high frequency power electronic circuits." *Measurement* 147, 2019, doi.org/10.1016/j.measurement.2019.106885.
- [3] D. IjikeOna, G. YunTian, R. Sutthaweekul, S. MohsenNaqvi. "Design and optimisation of mutual inductance based pulsed eddy current probe." *Measurement*, 144, 2019, pp 402-409.
- [4] S. Balci. "A Comparative Simulations on the Electromagnetic and Mechanical Effects of the Various Inductor Core Forms for DC-DC Converter Circuits." *Düzce University Journal of Science & Technology*, vol. 7, 2019, pp 1130-1139.
- [5] M. H. Sarul, G. Yıldırım, R. Gülgün. "Measurement of the inductance of a coil with core at different currents by a dc chopper." *Electrical Engineering*, vol. 82. 5, 2000, pp 273-277.
- [6] S. Dwari, S. Jayawant, T. Beechner, S. K. Miller, A. Mathew, M. Chen, J. Riehl, J. Sun. "Dynamics Characterization of Coupled-Inductor Boost DC-DC Converters." *IEEE COMPEL Workshop, Rensselaer Polytechnic Institute, Troy, NY, USA*, 16-19, 2006, pp 264-269.
- [7] İ. Sefa, S. Balci, N. Altin. "Design of the AC Line Reactors with Finite Element Method." *Journal of Polytechnic*, vol. 18. 4, 2015, pp 257-267.
- [8] Z. Dang, J. A. Abu Qahouq, "Modeling and Design Guidelines of High-Density Power Inductor for Battery Power Unit." *IEEE Applied Power Electronics Conference and Exposition (APEC)*. Long Beach, CA, USA, 2016.
- [9] Y. Du, T. Jiang, Z. Xu, "Coupled Electromagnetic and Thermal Simulation Method for a DC-DC Converter Inductor." *IEEE Southeast Con 2018*. St. Petersburg, FL, USA, 2018.
- [10] D. Munguia, G. Healy, "Design of High Power Planar Magnetics for a 1.8KW Phase Shifted Full Bridge Converter Using Advance FEA Electromagnetics Tools." *PCIM Europe 2019*, Nuremberg, Germany, 2019.
- [11] B.G. You, J.S. Kim, B.K. Lee, G.B. Choi, D.W. Yoo, "Optimization of Powder Core Inductors of Buck-Boost Converters for Hybrid Electric Vehicles." *Journal of Electrical Engineering & Technology* vol. 6. 4, pp 527-534, 2011.
- [12] H. C. Sartori, J. E. Baggio, J. R. Pinheiro, "A Comparative Design of an Optimized Boost Inductor Taking into Account Three Magnetic Materials Technologies: Volume, Cost and Efficiency Analysis." *10th IEEE/IAS International Conference on Industry Applications*. Fortaleza, Brazil, 2012.
- [13] M. S. Perdigão, S. F. Ferreira, M. Martins, A. S. Mendes, J. M. Alonso, "Finite Element Analysis of a Variable Inductor for an RSCC based LED Lamp Driver." *IEEE Industry Applications Society Annual Meeting*. Addison, TX, USA, 2015.
- [14] Z. Dang, J. A. Abu Qahouq. "Permanent-Magnet Coupled Power Inductor for Multiphase DC-DC Power Converters." *IEEE Transactions on Industrial Electronics*, vol. 64. 3, 2017.
- [15] Ferrit E1006028 core datasheet. <http://ferroxcube.home.pl/prod/assets/e1006028.pdf>.
- [16] S. Balci, N. Altin, S. Özdemir, İ. Sefa, "FEM Based Parametric Analysis of AC Line Reactors." *IEEE (POWERENG) 4th International Conference on Power Engineering, Energy and Electrical Drives, Istanbul, 2013*, pp 1328-1333.
- [17] İ. Sefa, S. Balci, N. Altin, S. Ozdemir, "Comprehensive analysis of inductors for an interleaved buck converter." *IEEE 15th International Power Electronics and Motion Control Conference (EPE/PEMC 2012)*, Serbia, 2012.

- [18] A.W. Lotfi, M. A. Wilkowski. "Issues and Advances in High-Frequency Magnetics for Switching Power Supplies." PROCEEDINGS OF THE IEEE, vol. 89, 6, 2001.
- [19] V. Leonavičius, M. Duffy, U. Boeke, S. C. Ó Mathúna. "Comparison of Realization Techniques for PFC Inductor Operating in Discontinuous Conduction Mode." IEEE Transactions on Power Electronics, vol. 19, 2, 2004.
- [20] M.K. Kazimierczuk, *High-frequency magnetic components*, Second Edition, Wiley, Ohio, USA, 2014.
- [21] M. K. Kazimierczuk, L. A. Starman. "Dynamic Performance of PWM DC-DC Boost Converter with Input Voltage Feedforward Control. IEEE Transactions on Circuits and Systems—I: Fundamental Theory and Applications." vol. 46, 12, 1999.
- [22] P. Mattavelli. "Digital Control of dc-dc Boost Converters with Inductor Current Estimation." Nineteenth Annual IEEE Applied Power Electronics Conference and Exposition, APEC '04. Anaheim, CA, USA, 2004.
- [23] M. Rashid, *Power Electronics Devices, Circuits, And Applications* Fourth Edition. NJ, USA: Pearson Education, 2014.
- [24] S. Iyasu, T. Shimizu, K. Ishii, "A Novel Inductor Loss Calculation Method on Power Converters Based on Dynamic Minor Loop." IEEE European Conference on Power Electronics and Applications. Dresden, Germany, 2005.
- [25] W. K. Mo, K. M. Paasch, M. Sachmann, "Optimal inductor winding geometries for minimizing winding loss in gapped inductor designs." 19th European Conference on Power Electronics and Applications (EPE'17 ECCE Europe). Warsaw, Poland, 2017.
- [26] A. Stadler, T. Stolzke, and C. Gulden, "Nonlinear Power Inductors for Large Current Crest Factors, Hindawi Publishing Corporation, Journal of Engineering", Article ID 687581, 6 pages, <http://dx.doi.org/10.1155/2013/687581>, 2013.



FUNDA BATTAL She received the B.S. and M.S. degrees in electrical education from the Gazi University, Ankara, in 2008 and 2011, respectively, and the Ph.D. degree in electrical and electronic engineering from Gazi University, Ankara, in 2019. Since 2011, she is a lecturer Electronics and Automation Department, Nevsehir Haci Bektas Veli University. Her research interests include transformer, core material, and transformer vibration.

BIOGRAPHIES



PINAR ARIKAN She received the B.S. degree in electrical and electronics engineering from Karadeniz Technical University, Trabzon, in 2017. She is currently a M.S. student in Karamanoğlu Mehmetbey University, Graduate School of Natural and Applied Sciences. Her research interests the design of coupled inductor for DC-DC converter circuits, and the electromagnetic 2D/3D modeling with FEA software.



SELAMI BALCI He received the B.S., M.S. and the Ph.D. degrees in electrical education from the Gazi University, Ankara, in 1997, 2010 and 2016, respectively. Since 2018, he is currently an assistant professor at the faculty of engineering, department of electrical and electronics engineering, Karamanoğlu Mehmetbey University. His research interests include medium frequency power transformer, transformer design with the soft magnetic core material, transformer thermal analysis, inductor design, electromagnetic modeling with FEA software, and performance analysis of the power electronics circuits.

The Spectral and Statistical Analysis of Gait Dynamics in ALS Disease

O. AKGUN, A. AKAN and O. YILMAZ


Abstract— Amyotrophic lateral sclerosis (ALS) disease, also known as motor neuron disease, is a disease resulting from loss of motor nerve cells in the spinal cord and brain stem region at central nervous system. Researchers can't find the reason of ALS for certain, however there are a wide variety of risk factors in respect of this disease. This disease has more than one risk factor. Researchers believe that it is resulted from a virus which leads to abnormality in immune system, spoils the structure of DNA and functioning of enzyme system, exhibiting neurotoxic properties. The signals coming to a single arm or leg muscle from upper and lower motor neurons are highly determinative in diagnosis of the disease, although there is not a specific test for diagnosing the ALS disease for certain. Doctors still conduct many tests even though the main symptoms of ALS are the signals coming to muscles. The developments related to gait analysis are used an auxiliary factor in diagnosis and analysis of ALS diseases. In this study, gait signals from control individuals and ALS diseases have been recorded by means of sensors sensitive to the strength under the foot. These signals' time-amplitude, amplitude spectrum, phase spectrum, average value and variance changes have been analysed. As a result of these inspections, differences of ALS signals from control signals have been identified.

Index Terms— ALS Disease, Gait Test, Spectral Analysis, Statistical Analysis.


I. INTRODUCTION

WALKING IS ADVANCING the body for the purpose of being able to go from one place to another. Though walking seems to be a very simple part of the life, actually it is an extremely chain of complex movements. For being able to walk without tiring for a long time without getting tired; brain, spinal cord, peripheral nerves, muscles, bones and joints


OMER AKGUN, is with Department of Computer Engineering University of Marmara, Istanbul, Turkey, (e-mail: oakgun@marmara.edu.tr).

 <https://orcid.org/0000-0003-3486-2197>

AYDIN AKAN, is with Department of Biomedical Engineering University of Izmir Katip Celebi, (e-mail: aydin.akan@ikcu.edu.tr).

 <https://orcid.org/0000-0001-8894-5794>

OZGUR YILMAZ, is with Department of Computer Education & Instructional Technology İstanbul University-Cerrahpaşa, Istanbul, Turkey, (e-mail: oyilmazistanbul.edu.tr).

 <https://orcid.org/0000-0002-1122-4558>

Manuscript received January 25, 2020; accepted March 18, 2020.
DOI: [10.17694/bajece.679926](https://doi.org/10.17694/bajece.679926)

should work together, joint movement, time and strength of contraction should be sufficient [1,2].

The complexity of normal walking and the difficulty of assessment with naked eyes have compelled scientists to develop detailed and reliable examination methods. Scientific assessment of human motion dates back to the era of Aristotle. Aristotle has found in B.C. 350 that joint movements are done by muscle constraints, after a few centuries Galen (A.D. 131-201) has suggested that muscle contraction is managed by nerves. During the Renaissance the x period Galilei's student Borelli, and in the 19th century Marey have performed dynamic studies regarding the human gait [3,4]. Eadweard Muybridge, the famous photographer has contributed significantly to the gait analysis in the late nineteenth century. A next stage in development of gait analysis is the conversion of photo images into numerical variables by Braune and Fisher in 1895. In 1930s, Eberhart and Inman have repeated these experiments in more detail by using a perforated disk rotating in front of the lens of the camera. The researchers who have put scientific gait analysis into clinical use as we know it today by dint of their studies began in 1950s are Verne Inman and Jacquelin Perry [1,3]. Gait analysis systems intended for clinical use have been developed by virtue of rapid development of personal computers began in 1980s, and they have been put up for sale and come into use in many countries around the world.

In today's context, the gait analysis is the numerical evaluation, identification and interpretation of the walking. Although many gait problems can be detected through visual inspections of experienced physicians, the gait analysis technology is required in order to interpret the problem numerically, to re-evaluate after saving it and to demonstrate the effectiveness of the treatment done in an objective manner. The gait of the patient is firstly assessed by looking with eyes and video recordings in modern gait analysis laboratories. Afterwards the motion data is transferred to the computer by means of transmitter or reflectors attached to appropriate points on the patient's body, and also ground reaction strength variations which are measured while pressing on a force platform mounted on the ground is installed into computer. In addition to these data dynamic electromyography and energy consumption measurements are also made in advanced laboratories. All these information are converted into digital data via custom software. Finally, it is diagnosed by evaluating the data together with clinical condition of the patient [3,4].

The studies within the scope of this article regarding the numerical assessment of neurodegenerative diseases and gait

dynamics have been done by Hausdorff and et.al. in the years of 1997 and 2000. These researches were usually based on analyses associated with gait rhythm [5,6].

Meanwhile this study is based on kinetic analysis of gait dynamics. Kinetic analysis is an examination of forces forming the movement (ground reaction forces, joist moments, joist strengths). The only data that can be measured in kinetic analysis is the ground reaction force vector (YTKV). YTKV is measured with plates sensitive to pressure which is measuring the total force applied to the ground by foot and called force plate. Transducers, which have been placed at both sides, front and rear and inside of the platform, measure the components of the load coming on the force plate surface at every three planes and transfer this data to computer [7,9].

II. DATA COLLECTION SYSTEM

60-seconds gait signals have been recorded by utilizing piezoelectric sensors sensitive to the force under the foot, by means of sensors sensitive to strength from 14 Normal and 13 ALS patients. The signals have been transferred to computer environment by converting into 18000 pcs series x(n) discrete signals (Fig. 1). [1].

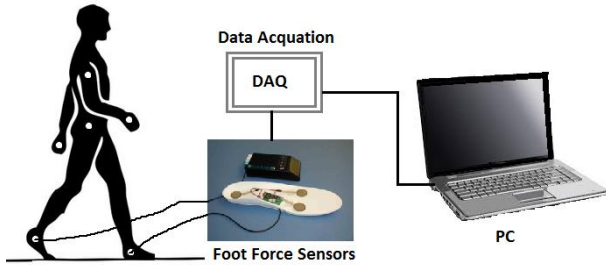


Fig.1. Data Collection System

III. TIME-AMPLITUDE ANALYSIS

The amplitude variations of the signal which will be analysed with time-amplitude analysis can be examined within time. It can be regarded as the most basic analysis. Variable electrical signals are a function of the time and frequency. Therefore, amplitude variations should be examined before making analyses of electrical signals on time and frequency plane. The change of these x(n) signs depending on time was shown in Fig.2.

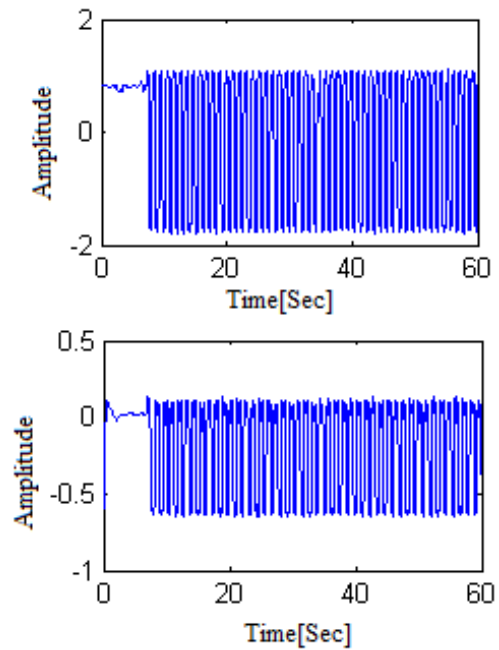


Fig.2. Control (upper) and ALS (lower) gait signals time-amplitude graphic

TABLE I
MAX. AND MIN. OF GAIT SIGNALS

Patient No	Control		ALS	
	Max	Min	Max	Min
1	0.3	-0.6	0	-0.6
2	0.2	-0.6	0.2	-0.5
3	0	-0.6	0.1	-0.6
4	0	-0.6	0	-0.5
5	0.7	-2	0.1	-0.6
6	0.2	-2	-0.1	-0.6
7	0.1	-0.6	0.1	-0.6
8	0.5	-2	0	-1.6
9	1	-1.7	0.2	-0.6
10	0.8	-1.8	0	-0.6
11	0.4	-2	0.1	-0.6
12	0.2	-0.6	0	-0.6
13	0.5	-2	0	-0.6
14	0.3	-0.6		

ALS signals observed in Fig.2 and Table 1 take the 0,0.1 and 0.2 values for max. and -0.6 and -0.5 values for min. It is seen that there are deviations in control signals.

IV. AMPLITUDE AND PHASE SPECTRUM ANALYSIS

The analysis made on the frequency plane is called spectrum analysis. The energy levels are specified of the signals in respect of different frequencies forming any signal in spectrum analysis. Electrical and physical information, which are useful for the test of electrical and mechanic

systems are obtained with this analysis.

The discrete Fourier transformation, which is defined with the expression of

$$X[k] = \sum_{n=0}^{N-1} x(n)e^{-j(2\pi/N)kn}, k = 0,1,\dots,N-1 \quad (1)$$

was applied to $x(n)$ signs in order to find amplitude and phase spectrums of $x(n)$ gait signals.

The amplitude spectrums of control and ALS gait signals are seen in Fig.3 [10,12].

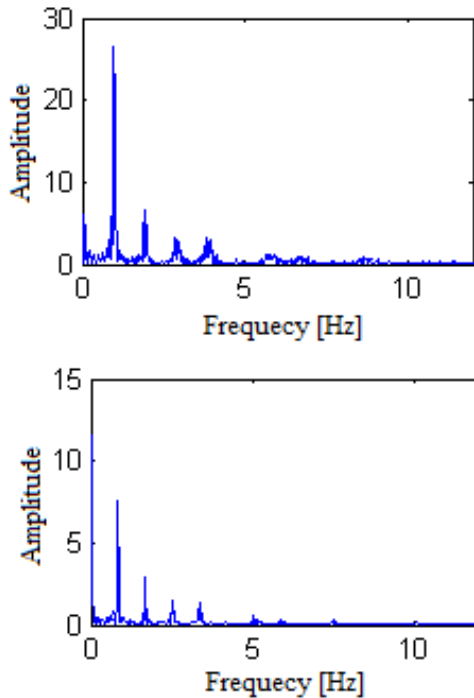


Fig.3. The amplitude spectrums of control (upper) and ALS (lower) signals

TABLE II
MAX. SPECTRAL AMPLITUDE VALUES

PatientNo	Control	ALS
1	10.2	3.6
2	12.4	8.3
3	7.8	9.6
4	6.1	4.2
5	39.2	4.7
6	21.5	4
7	79	5.7
8	26.3	8.7
9	26.6	9.9
10	30.2	8.6
11	22.9	7.6
12	8.7	9
13	21.6	6.5
14	73	

When Figure 3 and Table 2 are scrutinized, it is noteworthy

that spectral amplitude values of ALS signals around 1Hz are smaller than those of control signals. Meanwhile in Figure 3, the phase spectrums of control and ALS gait signals have been given in Fig.3.

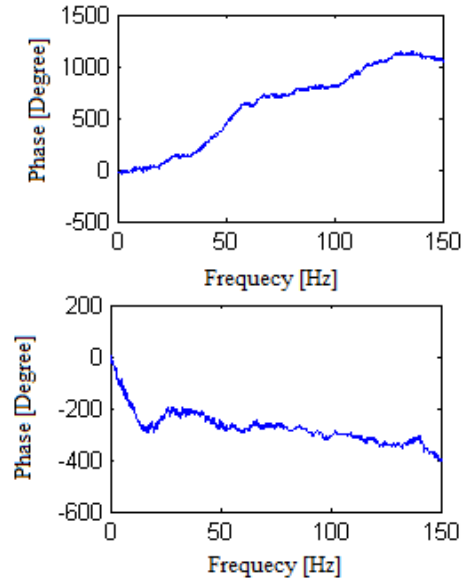


Fig.4. The phase spectrums of control (upper) and ALS (lower) signals

It is seen that the inclinations are smaller in spectral graphics in Fig.4 and of ALS signals in Table 3.

TABLE III
THE MAX. AND MIN. DEGREE VALUES REACHED BY CURVES IN PHASE SPECTRA

Patient No	Control		ALS	
	Max	Min	Max	Min
1	1134	-180	1339	-48
2	7	-459	1600	-235
3	100	-114	670	-35
4	534	-35	829	0
5	12	-500	645	-32
6	311	-46	802	-41
7	3049	-20	350	-28
8	1221	0	395	-84
9	1072	-31	261	-91
10	289	-102	529	-136
11	1952	-93	0	-400
12	1081	-200	17	-42
13	455	-154	204	-89
14	603	-74		

V. STATISTICAL ANALYSIS

One of the special values that can be defined the distribution of X gait signals is the expected values of signals, and the other one is the variance. The expected values of X gait signals are defined by expression of

$$E(X) = \sum_{n=1}^{18000} x_n P(X = x_n) \quad (2)$$

Consider the average of X as. In this case, the variance of X is calculated as the total of [7-13].

$$\sigma^2 = E(X - \mu)^2 = \sum_{n=1}^{18000} (x_n - \mu)^2 P(X = x_n) \quad (3)$$

TABLE IV
AVERAGE VALUES AND VARIANCES OF CONTROL AND ALS GAIT SIGNALS

Patient No	Control Mean	ALS Mean	Control Variance	ALS Variance
1	-0.062	-0.323	0.121	0.038
2	-0.128	-0.018	0.137	0.083
3	-0.308	-0.234	0.069	0.088
4	-0.323	-0.151	0.069	0.043
5	-0.461	-0.234	1.240	0.055
6	-0.858	-0.341	0.816	0.040
7	-0.259	-0.197	0.080	0.081
8	-0.682	-1.055	1.001	0.168
9	0.00007	-0.211	1.428	0.113
10	-0.267	-0.281	1.385	0.064
11	-0.694	-0.191	0.958	0.094
12	-0.117	-0.268	0.129	0.072
13	-0.528	-0.236	1.040	0.079
14	-0.089		0.139	

When the Table 4 is examined; the average values of ALS gait signals take values usually around -0.2. Meanwhile it has been observed in an inspection according to variance values that ALS signals were taking quite small values compared to control signals.

VI. CONCLUSION

Today there are over 30000 people living with ALS disease. This disease causes progressive degeneration of motor neurons in the body affecting the voluntary muscle control. At the end, those affected loses their livelihood, independence and eventually their lives. Occurrence of upper and lower motor nerve cells symptoms together are characteristic. For individuals suffering this disease, the upper motor neuron symptoms are excessive reflexes, stiffness, rising upward of thumb and reflex increase at the chin, and the lower motor neuron symptoms are weakness, muscle weakness, cramps and

twitches. The disease varies from person to person. In this study, gait dynamics test was done in determination of ALS disease. The findings obtained from test results are rather satisfactory. The results of analyses can be enumerated as below.

- ALS gait signals take certain max. values in time-amplitude analysis (0,0.1,0.2).
- The amplitudes of ALS gait signals in the amplitude spectrum are notable with small values.
- The phase curves of ALS gait signals in the phase spectrum manifest themselves with small inclinations.
- ALS signals are diagnosed with average values at around -0.2. Moreover their variances have been identified as a distinguishing feature specific to ALS signals with pretty small values.

REFERENCES

- [1] O. Akgun, H. Demir, A. Akan, , "Detection of ALS Disease Using AR Models of Gait Dynamics", Electrical, Electronic and Computer Engineering Symposium (ELECO), 26-30 November 2008, Bursa Turkey.
- [2] J. Perry , *Gait Analysis: Normal and Pathological Function*, Slack Incorporated Thorofare, NJ, 1992.
- [3] S. Yalçın, N. Özaras , *Gait Analysis*, Avrupa Publications, Istanbul, 2001 (inTurkish).
- [4] M. Whittle , *Gait Analysis: an introduction*, Butterworth Heinemann, Oxford, 1993.
- [5] J.M. Hausdorff, S.L. Mitchell, R. Firtion, C.K. Peng, M.E. Cudkowicz, J.Y. Wei, A.L. Goldberger, "Altered fractal dynamics of gait: reduced stride-interval correlations with aging and Huntington's disease", *J. Applied Physiology*, Vol. 82, 1997, pp.262-269.
- [6] J.M. Hausdorff, A. Lertratanakul, M.E. Cudkowicz, A.L. Peterson, D. Kaliton, A.L. Goldberger, "Dynamic markers of altered gait rhythm in amyotrophic lateral sclerosis" *J. Applied Physiology*, Vol. 88, 2000, pp.2045-2053.
- [7] J.R. Gage, "The clinical use of kinetics for the evaluation of pathological gait in ce-rebral palsy", *Instr Course Lect.*, Vol 44, 1995, pp.507-15.
- [8] F. Akdeniz, *Probability and Statistics*, Nobel Medical Publications, Istanbul, 2004.
- [9] T. Liu, Y. Inoue, K. Shibata, K. Shiojima, "Three-dimensional gait analysis system with mobile force plates and motion sensors", *Ubiquitous Robots and Ambient Intelligence (URAI)*, 2011 8th International Conference on 2011, pp. 107-110.
- [10] S. Tallapragada, L.V.C. Srinivas, "Marker less view independent gait analysis using DFT", *Electronics Computer Technology (ICECT)*, 2011 3rd International Conference on Vol. 3 , 2011 , pp. 11-13.
- [11] J.H. Yoo, M.S. Nixon, C.J. Harris, "Model-driven statistical analysis of human gait motion", *Image Processing. 2002. Proceedings. 2002 International Conference on Vol. 1, 2002*, pp.285-288.
- [12] E. Onal, J. Dikun, "Short-Time Fourier Transform for Different Impulse Measurements". *Balkan Journal of Electrical & Computer Engineering*, Vol.1, No.1, 2013, p.44-47.
- [13] Y. Wu, S. Krishnan, "Statistical Analysis of Gait Rhythm in Patients With Parkinson's Disease", *Neural Systems and Rehabilitation Engineering*, IEEE Transactions on Vol. 18. 2, 2010 , pp. 150-158.

BIOGRAPHIES



OMER AKGUN is an assistant professor at the, Department of Computer Engineering, Technology Faculty, Marmara University. He received his first Ph.D. in the Communication Engineering in 2009 from Yildiz Technical University and the second Ph.D. in the Electronic and Communication Education Department in 2011 from Marmara University. His current research interests are signal processing, biomedical signal processing, signal modelling and communication systems.



AYDIN AKAN received his Ph.D. degree in Electrical Engineering from the University of Pittsburgh, Pittsburgh, PA, USA, in 1996. He has been with the Department of Electrical and Electronics Engineering, Istanbul University, between 1996 and 2017. Currently he holds a Professor position at the department of Biomedical Engineering, Izmir Katip Celebi University, since 2017. His research interests are non-stationary signal processing, time–frequency signal analysis methods and their applications to wireless communications and biomedical engineering. Dr. Akan is a senior member of the IEEE Signal Processing Society and editorial board member of the Digital Signal Processing Journal.



OZGUR YILMAZ, graduated with a bachelor's degree in 2000 from Department of Electrical Education, Marmara University, Turkey. He received M.S. and PhD. degrees in Electrical Education from the Institute for Graduate Studies in Pure and Applied Sciences, Marmara University, in 2003 and 2010, with his dissertation titled “Web Supported Design, Teaching, and Evaluation of the Electromechanical Systems Course.” He has worked as an Assistant Professor at İstanbul University-Cerrahpaşa. He has published numerous papers on mechatronics education, distance education and game-based learning. His current research focuses on game-based learning and distance education

Artificial Intelligence Based Game Levelling

Y. SARICA and M. CETIN*

Abstract— The applications of artificial intelligence (AI), which is a comprehensive information technology, have been closely related to game technologies. Today, artificial intelligence-based game development applications are increasing their popularity day by day. In this study, the levelling process of a 2-dimensional (2D) platform game has been investigated. The game developed and called “Renga” has a basic gameplay. Game data has been processed through an artificial neural network (ANN), *k*-nearest neighbour, decision and random tree algorithms and deep learning model that is trained with gameplay and user information. The classification process with the output data provides results for the next game level. In this way, the most effective playability impression that the developers offer to the game users has been created according to game. Furthermore, the variety of difficulty calculated with dynamic data by the user is provided by Renga, in which new sections/levels are created with user-specific assets. Thus, the most efficient gaming experience has been transferred to the users.

Index Terms— Artificial Intelligence, Difficulty Adjustment, Content Generation, *k*-Nearest Neighbor, Random Forest, Artificial Neural Networks.

I. INTRODUCTION


AMONG the platforms, different types of games that can attract the attention of their target users are increasing and improving. Video games that contribute to this development are very successful in environmental design and transferring the realistic behavior of Non-Player Characters (NPC) [1]. In video games, which have become an important part of the entertainment industry, the goal is to optimize players' experience rather than creating the most difficult game [2-4]. A good video game has a long learning curve. Therefore, it is possible for the artificial intelligence-based-algorithm to update itself in accordance with the interaction of people who learn the game. Research on the content production of games helps to develop better quality or interesting games. The player's expectation is that the difficulty of the game matches his personal gaming skills [5,

6]. Therefore, a proper content generation depends on the algorithm producing a meaningful output with player performance. Depending on the algorithm, the difficulty level of the game should increase as the player's abilities develop [7].


Designing quality content for millions of players in the game industry has become the goal of game developers. In [8], a procedural level generator based on interactive evolutionary algorithm has been introduced for the platform game. Another automatic level generator based on Darwin's theory of Natural selection was proposed in [9]. In [10], a Rhythm-based approach was proposed for 2D Platform content generation. In [11], dynamic difficulty adjustment is provided by Polymorph approach which changes game difficulty depending on player performance. In the literature, there are various studies that are designed by machine learning (genetic algorithm, artificial neural network, support vector machine) [12-16], probabilistically techniques [17], Procedural Content Generation (PCG) [18-20] for dynamic game levelling. The difficulties caused by the behavior of NPC objects can be supported by artificial intelligence applications to maximize a user's enjoyment. The use of artificial intelligence in games is diversifying and deepening day by day. In [21], Yannanakis and Togelius have pointed out the following matters on the subject: behavior learning of non-player characters, exploration and planning applications, player modeling, artificial intelligence game competence applications, methodological content production, stories that may vary, the formation of realistic environmental movements, artificial intelligence game design, use of commercial games, general in-game artificial intelligence.

In artificial intelligence based studies proposed for dynamic game leveling, instead of randomly determining the course of the game, levelling is done taking into account the player's characteristics (aggression, courage, intelligence and cooperation) and the appropriate reaction is selected depending on these designs. Neural networks successfully applied to various games can be considered as a means of updating the artificial intelligence system. As an artificial intelligence system that can evolve over time, many game developers often do not use genetic algorithms because they need too much CPU power and are too slow to produce useful results [22]. Finite State Machines (FSMs) have often used by game developers because it is easy to implement, test, modify, and personalize [22]. Increasingly, there has been a tendency towards Fuzzy State Machines (FuSMs) since the use of fuzzy logic allows the recognition of non-binary conditions [22]. Demasi and Cruz used the genetic algorithm technique to protect the most appropriate agents according to the player's game performance [23]. In addition, there are several studies

YUNUS SARICA, is with Department of Computer Engineering, University of Pamukkale University, Denizli/Turkey,(e-mail: yunusssarica@gmail.com).

 <https://orcid.org/0000-0002-1969-9005>

MERIC CETIN, is with Department of Computer Engineering, University of Pamukkale University, Denizli/Turkey,(e-mail: mcetin@pau.edu.tr).

 <https://orcid.org/0000-0002-7871-4850>,

*Corresponding Author.

Manuscript received November 05, 2019; accepted April 04, 2020.

DOI: [10.17694/bajece.642973](https://doi.org/10.17694/bajece.642973)

based on Gaussian Mixture Module with dynamic scripting [24, 25].

In this study, game levels for *Renga*, which was developed as a 2D platform game, have been identified based on artificial intelligence. In this context, perception of difficulty in infinite games has been examined, balanced the player performance of *Renga* game and then different game levels have been designed according to player ability. As a result of processing the game data with artificial intelligence methods (artificial neural network, k-nearest neighbor, decision and random tree algorithms, deep learning model), it is provided to produce game-specific procedurally appropriate content and meaningful outputs specific to the player. Algorithmically variable game content has been generated by PCG, which significantly has reduced the time and cost of development processes for game developers. In this way, the most effective playability impression that the developers offer to the game users has been created according to *Renga*. Furthermore, the variety of difficulty calculated with dynamic data by the user is provided by *Renga*, in which new sections/levels are created with user-specific assets. Thus, the most efficient gaming experience has been transferred to the users.

II. PLATFORM GAMES AND *RENGA*

In this section, the process and technological developments of platform games that have been on the agenda since the beginning of video game history are presented. Platform games are the first phase of video game history. Bertie the Brain is known as the first industrial game developed in the 1950s. This process continued eight years later with William Higinbotham's Tennis for Two. Pong, as the first realization of this game, has become one of the platform games [26]. Then, game technologies have developed into different categories. Productions such as Super Mario, Contra and Metal Slug have not lost their popularity despite the emergence of 3D games, and then, a new era of video games began with platform games such as Limbo, Inside, Ori and the Blind Forest. Many of these games have a special place today because they are widely distributed and can make their voices heard through digital gaming platforms [27].

After all these similarities and examples that have continued its genre, it was decided to perform this study with platform games. In the comparative tests, it has been paid attention that the developed game (*Renga*) can work in harmony with the selected artificial intelligence method.

Considering the existence of common objects, platform games can generally consist of a character, enemy, interactive objects and various platform groups. Although the aim of the selected character is to terminate the game parts, it is also important to be able to produce the appropriate level according to the game performance. One of the goals of designed game is to help the developer to design the level by providing sufficient data as soon as possible while building the sections. In this study, a 2D platform game (*Renga*) is designed which is simple to use and diversifies game assets. The developed game has been also introduced globally to the literature. In

Renga methodological content production is based on an artificial intelligence algorithm and the identification data obtained in this way contribute to the change of in-game dynamics. In addition, a new data set was created by processing the data collected over the game. Game design is based on basic steps for artificial intelligence methods used in the study. The degrees of difficulty in the game depend on the numerical properties of the game assets. The main characteristics of the interaction of this process are:

- Simple gameplay,
- Interaction between assets,
- Updating the character structure,
- Contribution to the game score
- Academic study
- To provide entertainment to the global users

Renga, which is compatible with phones with Android operating system, is developed by Unity platform. *Renga* is an infinite game that can be played until the player makes a mistake. The game engine Unity is widely used in such games because it supports 3D modeling and animation design as well as 2D game development [28]. Each field in *Renga* consists of a scene whose content is full of game objects ("sprite", "texture", "prefab" and "object"). The game has 4 scenes actively: "Menu", "howToPlay", "gameplay" and "end" scenes. In the "menu" scene, the data to be tested is collected. Users have access to the code file of any scene in Unity through information on randomness value, horizontal speed, gravity, jump value and distance between obstacles. The game process takes place on the "gameplay" scene. A gameplay scene for the *Renga* is illustrated in Fig.1. Only *Renga* is available as game asset in this scene. Background, floor, "Aheng" and "Rengec" are dynamically composed.

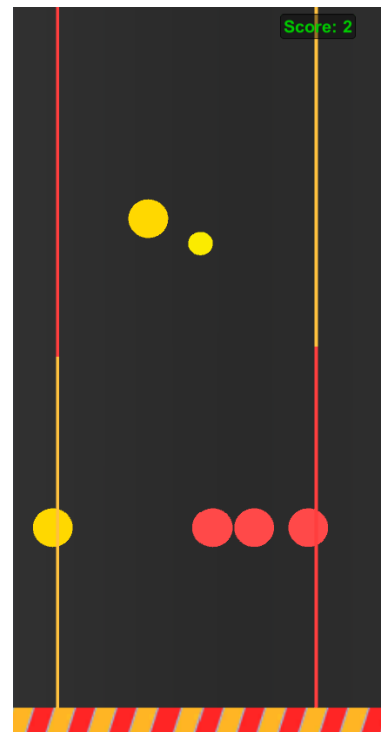


Fig.1. *Renga* gameplay scene

III. ARTIFICIAL INTELLIGENCE BASED GAME LEVELING

In this section, the effects of *Renga* game for the user, how the data is collected and the results obtained with the technology and methods used in this process are explained.

A. Data Collection

In order to control the game, the collected data is used to create a level, which affects the type and difficulty of the level. The statistical characteristics such as how often the player jumped, died, how much he caught *Renga* cannot be directly controlled by the game because of they depend on the player's skill and playing style. In the tests, the game levels for a particular player have been dynamically generated as:

1. A first level is generated with random parameters.
2. Game features are recorded.
3. Using recorded game features, a new adapted level is generated based on player experience.

The tests in this study are based on the approximately 1500 game session played by the 15 player (each user played the game at least 5 times with a randomly generated initial value). In the simulations, 80% of the input data set was used as training data. Sample game feature records of any user according to random parameters (seed, speed, gravity, jump, distance) are given in Table I.

TABLE I
RANDOM GAME SAMPLES FOR ANY USER

PHONE_ID	P1	P2	P3	P4	P5	P6	P7
ABCDEF9876	50	1.0	12	4.0	1.8	0	0
ABCDEF9876	50	1.0	10	4.0	1.8	0	0
ABCDEF9876	50	1.0	14	4.0	1.8	0	0
ABCDEF9876	50	1.0	12	5.5	1.8	0	0
ABCDEF9876	50	1.0	12	2.5	1.8	0	0

*P₁= seed value, P₂= speed value, P₃= gravity value, P₄= jump value, P₅= distance value, P₆= static value, P₇= score.

The "PlayerPrefs" feature provides access to game information within Unity. This data, which was added in the database before, was obtained with `datas.php` file. The web service is provided with the `WWWForm()` object specified in Unity. The data is then transferred to the server, "`http://gameonyou.tk/`". *Renga* is available to users in the Google Play Store without academic features. Therefore, the data is primarily hosted on the local server created with `xampp`. In addition, the application file for the academic version externally is shared with users via the address, "`http://gameonyou.tk/Renga.apk`". In the file dump named `dbConnData.php`, records are processed by connecting to the database with the `pdo` structure.

First of all, when the user runs the game for the first time, the in-game data is collected and inserted to the database.

Gameplay data, except for fixed values, were examined by increasing or decreasing each value at a certain rate, except for a general data type. If any row given in Table 1 is considered; as the values changed, the other values remained constant and this situation continued with combinations. These records are explained in detail as follows:

Seed: Seed is a fixed value and set to 50. All users' test data were used with this seed value in artificial intelligence processes. Users who have completed all test values will start playing different sections with a different seed value.

Speed: The default value of 1.0 is the feature that determines the speed of *Renga* during gameplay. With the values of 0.8 and 1.2, it was aimed to observe the gameplay status of the users. Reaching the Rengecs is necessary for the advancement of *Renga*.

Gravity: The default value of 10.0 allows vertical movement of *Renga* in a constant direction during gameplay. Interactions of 12.0 and 8.0 values in users were examined. If *Renga* hits the ground due to this value, the game will end. The user protects *Renga* from the ground by touching the screen during the gameplay.

Distance: The default is 1.6. This variable specifies the value of the distance between the rails. It is differentiated with 1.0 and 3.0 values. It is inversely proportional to the speed value and can change the perception of difficulty.

Jump: The default value is 4. The effect of the differences with the binary values of 5.5 and 2.5 was observed in the users. The user can increase the height of the vertical moving *Renga* by the magnitude of this value.

The importance of jump and gravity values was considered when passing over or under the distinction points of Rengecs. Distinction points where *Renga* cannot pass are prevented from occurring. When enough data (recorded at 5 different times) is collected from the user, the in-game values will change. After all these collected data, the results of the related user are classified as easy, medium or difficult in designed artificial intelligence models. Since the initial data collected from the user changes after the new level proposed by the artificial intelligence model, the user will be able to continue to play the game more efficiently.

B. Examination of Renga Data

The data collected by the *Renga* game is sufficient for user-based difficulty control and adjustment. According to the input set given as an example in Table 2, it will be appropriate to mention the data that creates the algorithm inputs and the features of these data before explaining the artificial intelligence model to be created.

Randomness: It can be defined as the renovation value. All assets in the game are created according to a random value. The user plays at least 5 games for each data of the same randomness value. According to Eq. (1), it is possible for a user to present a minimum of 120 result outputs in the test environment through the designed artificial intelligence models.

$$4! = 24 \rightarrow 24 * 5 = 120 \quad (1)$$

In this way, new data can be created for difficulty control using the in-game values in the different seed values.

TABLE II
SAMPLE GAME DATA

PHONE_ID	K1	K2	K3	K4	K5	K6
ABCDEF123	42	1.5	7	2.5	3	6
ABCDEF123	42	1.5	7	2.5	3	9
ABCDEF123	42	1.5	7	2.5	3	9
ABCDEF123	42	1.5	7	2.5	3	10
ABCDEF123	42	1.5	7	2.5	3	4

*K₁= randomness, K₂= horizontal speed, K₃= gravity value, K₄= jump value, K₅= distance between obstacles, K₆= score.

Horizontal speed: The path taken from the beginning of the game by the “K₂” values in Table 2. According to this table;

$$K_2 = 0,016 * 1,5 = 0,024 \quad (2)$$

In Eq. (2), 0.024 units of horizontal position changes occur for each frame per second. Increasing horizontal speed, which is one of the inputs in the artificial intelligence model, will increase the perception of difficulty in the game cause the user to be in easy class.

Gravity: It is vertical movement value, created with the “Rigidbody” feature, assigned to the *Renga* object from the beginning of the game. According to the gravity values given in Table 2 and equation (3), 0.112 unit changes occur for each frame.

$$K_3 = 0,016 * 7 = 0,112 \quad (3)$$

Increasing this value is inversely proportional to the jump value. There should be a balance between these values. Increasing this value is inversely proportional to the jump value. Increasing the value increases the difficulty, and the fact that it is too low makes the game impossible.

Jump: This variable value of action which user interacts with. Each time the user touches the screen, *Renga* will move vertically upwards by 2.5 units. During this movement, gravity also continues the downward movement. The accelerated *Renga* starts to move down again vertically. If the value is too high, even the lowest gain will lead to an impossible game.

Distance between obstacles: It is the distance unit between Rengecs. There is a 3-unit distance between the Rengecs, which are proportioned by two color segments (yellow and red) according to their seed value. Increasing this value makes the game easier, while decreasing it can significantly differentiate the perception of difficulty in the user experience.

Score: This value determines the player's highest gain in a game process. It also represents the number of Rengec that

could be passed. It is the most important factor in the classification of difficulties for users. These data are differentiated in the results section.

During the *Renga* design, many different technical fields have been used besides the package programs. In the following sections, used artificial intelligence algorithms and findings are presented.

C. Methods

In this study, the recursively used data was tested in various artificial intelligence models such as k-nearest neighbor, random forest, deep learning and artificial neural networks in the category of classification algorithms and analyzed through confusion matrix. Success and error coefficients were determined based on the variables within each method and were detailed for future studies.

1) k-Nearest Neighbor

The k-nearest neighbor method performs the classification process based on similarities where the data is clustered in cartesian plane. According to the selected value of k, which cluster the data belongs to is determined [29]. The pseudo-code expression for this algorithm is given as follows:

```
% X: training data, Y: class labels of X, x: unknown sample
➤ Determine parameter k: the number of nearest neighbor
➤ Classify (X, Y, x)
➤ for i = 1 to m do
    • Compute distances d(Xi,x) between the query-instance and all training data.
➤ end for
➤ Compute set I containing indices for the k smallest distances d(Xi,x).
➤ return majority of the category label for {YI}
```

2) Random Forest

Random forests, which are commonly used in classification or regression processes, consist of a large number of individual decision trees [30]. Random forest is an efficient tree algorithm that can model game data collected from different users in a compatible way. The pseudo-code for this algorithm is given as follows [31]:

```
➤ Generate c classifiers
➤ for i = 1 to c do
    • Randomly sample the training data D with replacement to produce Di
    • Create a root node, Ni containing Di.
    • Call BuildTree(Ni)
➤ end for
BuildTree(N):
➤ if N contains instances of only one class then
    • return
➤ else
```

- *Randomly select* x
- *Select the feature* F *with the highest information gain to split on*
- *Create* f *child nodes of* N , N_1, \dots, N_f , *where* F *has* f *possible values* (F_1, \dots, F_f)
- *for* $i = 1$ *to* f *do*
 - *Set the contents of* N_i *to* D_i , *where* D_i *is all instances in* N *that match*
 - F_i
 - *Call* $\text{BuildTree}(N_i)$
- *end for*
- *end if*

3) Deep Learning

Deep learning is a subfield of machine learning framework that involves one or more layered artificial neural networks. The difference from artificial neural networks is the ability to perform more operations as given in [32, 33]. In this study, deep learning algorithms were used to see the contribution of newly added game-data during training.

4) Artificial Neural Network

Artificial Neural Network is a computational model inspired by biological neural networks. An artificial neural network may comprise a plurality of neurons arranged in a series of layers. The input layer receives various forms of information. These data are passed through one or more hidden layers and converted to the output unit where the input can be used. An example of pseudo-code for back-propagation algorithm in training ANN is given as [34]:

% X : training data of size $m \times n$, y : class labels of X , w : the weights for respective layers, l : the number of layers

- % $D^{(l)}_{ij}$ the error for all i, j , $t^{(l)}_{ij} = 0$, for all i, j .
- *for* $i = 1$ *to* m *do*
 - $a^l = \text{feedforward}(x^{(i)}, w)$
 - $d^l = a(L) - y(i)$
 - $t^{(l)}_{ij} = t^{(l)}_{ij} + a^l_j * t^{(l+1)}_i$
- *if* $j \neq 0$ *then*
 - $D^{(l)}_{ij} = (1/m) * t^{(l)}_{ij} + \lambda w^{(l)}_{ij}$
- *else*
 - $D^{(l)}_{ij} = (1/m) * t^{(l)}_{ij}$, *where* $(d/dw^{(l)}_{ij}) * j(w) = D^{(l)}_{ij}$
- *end if*
- *end for*

IV. SIMULATION RESULTS

In this study, it is assumed that all users have played the game for the first time. The data set used for *Renga* consists of data from multiple users. Although the values for the application are the same, differences in points affect the result. The main reason is that the records are classified as *hard*, *medium* and *easy* according to the data content called “score”. As a result of the tests, it was found that the most important factors affecting the score were gravity and speed variables. In the experimental tests, the simulations were tried on the artificial intelligence methods mentioned above with many

different parameters and the most successful results were recorded. The results provide the level identification information for the *Renga* game. This data can be used to provide a more efficient level and in-game dynamics to the players. In the tables below, the best performance rates obtained for artificial intelligence methods are given.

TABLE III
PERFORMANCE RATES IN DEEP LEARNING

	Hard	Medium	Easy	Consistency
Hard (Estimated)	9	20	0	% 31.03
Medium (Estimated)	12	9	10	% 29.03
Easy (Estimated)	92	151	345	% 58.67
Singular Performance	% 7.96	% 5.00	% 97.18	

In deep learning tests, data has been trained according to the model of 2 hidden layer network. The best results for the tests with different activation functions were obtained in *maxout* function. The performance of the deep learning algorithm according to the learning rate (0.6) in the test data is %56.02.

TABLE IV
PERFORMANCE RATES IN k -NEAREST NEIGHBOR

	Hard	Medium	Easy	Consistency
Hard (Estimated)	42	50	0	% 45.65
Medium (Estimated)	78	154	0	% 66.38
Easy (Estimated)	0	0	0	% 0.0
Singular Performance	% 35.00	% 75.49	% 0.0	

Many tests have been performed with different k values for the k -nearest neighbor method which classifies according to similarities in the data. According to the simulation results, the best performance of the algorithm with the related learning rate was obtained for $k = 5$. The performance of the k -nearest neighbor algorithm according to the learning rate (0.8) in the test data is %60.49. It is generally considered that the k -nn method should not be used in inconsistent data because it has different gains for variable user types.

TABLE V
PERFORMANCE RATES IN RANDOM FOREST

	Hard	Medium	Easy	Consistency
Hard (Estimated)	4	1	0	% 80.00
Medium (Estimated)	13	40	17	% 57.14
Easy (Estimated)	27	48	174	% 69.88
Singular Performance	% 9.09	% 44.94	% 91.10	

As shown in Table 5, the change of algorithmic parameters for random forest did not bring an effective difference to the results. For this reason, the samples were differentiated and the findings were separated according to linear-complex samples. The performance of the random forest algorithm according to the learning rate (0.8) in the test data is %63.89.

The process that leads to the measurement of the level of the game allows the new values to be generated by the obtained values to be tested on the same users and thus to determine the new perception of difficulty. The decision tree structures in the random forest algorithm according to gravity and speed variables, which is one of the most important factors affecting the score, are presented in Figure 2 and Figure 3.

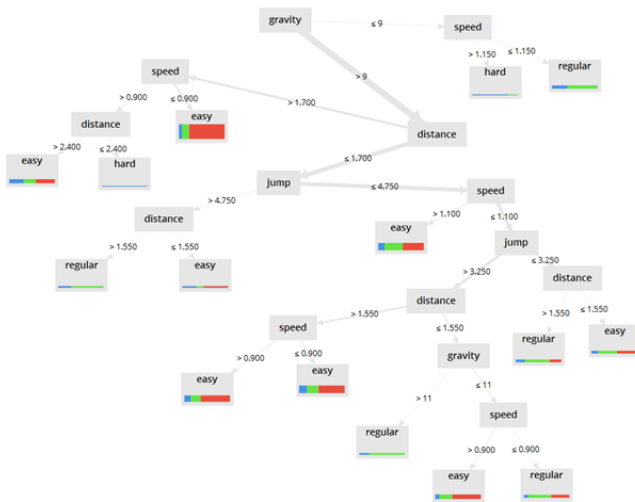


Fig.2. Decision tree method via gravity

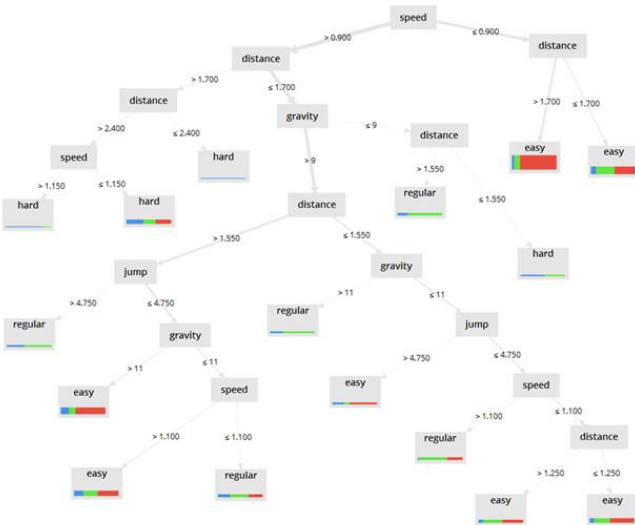


Fig.3. Decision tree method via horizontal speed

TABLE VI
PERFORMANCE RATES IN ARTIFICIAL NEURAL NETWORK

	Hard	Medium	Easy	Consistency
Hard (Estimated)	0	0	0	% 0.0
Medium (Estimated)	15	28	12	% 50.91
Easy (Estimated)	29	61	179	% 66.54
Singular Performance	% 0.00	% 31.46	% 93.72	

The performance of the ANN algorithm according to the learning rate (0.8) in the test data is %63.58. In the

experimental simulations, one hidden layer, one input layer and one output layer have been used for the ANN algorithm. The best results for the tests with different activation functions were obtained in sigmoid function for ANN implementation. According to the tests, ANN method is the second most efficient algorithm. However, it was found that the hard tag could not be predicted by the data sets mentioned.

V. CONCLUSION

In this paper, a level identification procedure has been performed with different methods and different parameters. This classification process has been tested on the users of *Renga*, an infinite game. The proposed structure provides an online gaming adaptation mechanism that can be used to effectively optimize the player experience. It is seen that the best method used in the study is the random forest algorithm. Furthermore, the random forest algorithm provided the best classification distribution. Besides the success rate, it is seen that it constitutes an efficient method for classification. The models of decision trees it offers are guiding and can provide high benefits. In a data set classification in which users' game identities are included, it is anticipated that new studies may emerge. In this way, the game can be played over certain basic users with certain values and data collection can make the leveling work more efficient. According to the results, the in-game values offered to the users gave positive results and it was observed that the game pleasure of the users was satisfactory. A short way have been shown to developers to present this experience, which saves time for improvements. The evaluations that can be made through this game can give an idea to be applied to other platform games. In the later stages of the studies, it was paved the way for its application in platform games with different genres and rich content.

ACKNOWLEDGMENT

This study was supported by Scientific Research Coordination Unit of Pamukkale University under the project number 2018FEBE003.

REFERENCES

- [1] Y. Sarica "Game Levelling with Artificial Intelligence." Master Degree Thesis, Pamukkale University, The Graduate School of Natural and Applied Science, 2019
- [2] A. J. Baldwin. "Balancing act: the effect of dynamic difficulty adjustment in competitive multiplayer video games", 2016.
- [3] Y. Zhang, S. He, J. Wang, Y. Gao, J. Yang, X. Yu, L. Sha. "Optimizing player's satisfaction through DDA of game AI by UCT for the Game Dead-End". In Natural Computation, Sixth International Conference on, Vol. 8, 2010, pp. 4161-4165.
- [4] J. P. Gee. "What video games have to teach us about learning and literacy". Computers in Entertainment, 1(1), 2003, 20-20.
- [5] M. Csikszentmihalyi. "Flow and the psychology of discovery and invention". Harper Perennial, New York, 1997, 39.
- [6] R. Hunnicke. "The case for dynamic difficulty adjustment in games". In Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology. 2005, pp. 429-433.
- [7] J. Sinclair. "Feedback control for exergames". Theses: Doctorates and Masters, 2011
- [8] M. Kerssemakers, J. Tuxen, J. Togelius, G. N. Yannakakis. "A procedural procedural level generator generator". In 2012 IEEE

- Conference on Computational Intelligence and Games, 2012, pp. 335-341.
- [9] F. Mourato, M. P. dos Santos, F. Birra. "Automatic level generation for platform videogames using genetic algorithms". In Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology, 2011, p. 8.
- [10] G. Smith, M. Treanor, J. Whitehead, M. Mateas, (). Rhythm-based level generation for 2D platformers. In Proceedings of the 4th International Conference on Foundations of Digital Games, 2009, pp. 175-182).
- [11] M. Jennings-Teats, G. Smith, N. Wardrip-Fruin. "Polymorph: dynamic difficulty adjustment through level generation". In Proceedings of the 2010 Workshop on Procedural Content Generation in Games 2010, p. 11.
- [12] F. Mourato, M. P. dos Santos, F. Birra. "Automatic level generation for platform videogames using genetic algorithms". In Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology 2011, p. 8.
- [13] L. Ferreira, C. Toledo. "A search-based approach for generating angry birds levels". In Computational intelligence and games, 2014.
- [14] L. Galway, D. Charles, M. Black. "Machine learning in digital games: a survey". Artificial Intelligence Review, 29(2), 2008, 123-161.
- [15] P. Spronck, I. Sprinkhuizen-Kuyper, E. Postma. "Online adaptation of game opponent AI in simulation and in practice". In Proceedings of the 4th International Conference on Intelligent Games and Simulation, 2003, pp. 93-100.
- [16] D. Johnson, J. Wiles. "Computer games with intelligence". In Fuzzy Systems. The 10th IEEE International Conference on, Vol. 3, 2001, pp. 1355-1358.
- [17] M. Persson. "Infinite Mario bros". 2008, Online Game.
- [18] W. Baghdadi, F. S. Eddin, R. Al-Omari, Z. Alhalawani, M. Shaker, N. Shaker. "A procedural method for automatic generation of spelunky levels". In European Conference on the Applications of Evolutionary Computation, 2015, pp. 305-317.
- [19] G. Smith, M. Treanor, J. Whitehead, M. Mateas. "Rhythm-based level generation for 2D platformers". In Proceedings of the 4th International Conference on Foundations of Digital Games, 2009, pp. 175-182.
- [20] V. der Linden, R. R. Lopes, R. Bidarra, "Designing procedurally generated levels", In Proceedings of the second workshop on Artificial Intelligence in the Game Design Process, 2013.
- [21] G. N. Yannakakis, J. Togelius. "A panorama of artificial and computational intelligence in games". IEEE Transactions on Computational Intelligence and AI in Games, 7(4), 2014, 317-335.
- [22] S. Woodcock, J. E. Laird, D. Pottinger, "Game AI: The state of the industry". Game Developer Magazine, 8,c2000.
- [23] P. Spronck, I. Sprinkhuizen-Kuyper, E. Postma. "Difficulty scaling of game AI". In Proceedings of the 5th International Conference on Intelligent Games and Simulation, 2004, pp. 33-37.
- [24] S. Lee, K. Jung. "Dynamic game level design using gaussian mixture model". In Pacific Rim International Conference on Artificial Intelligence, 2006, pp. 955-959.
- [25] P. Spronck, M. Ponsen, I. Sprinkhuizen-Kuyper, E. Postma. "Adaptive game AI with dynamic scripting". Machine Learning, 63(3), 2006, 217-248.
- [26] S. L. Kent. "The Ultimate History of Video Games: From Pong to Pokemon-The Story Behind the Craze That Touched Our Lives and Changed the World", 2001, New York: Three Rivers Press.
- [27] J. Togelius, S. Karakovskiy, J. Koutmik, J. Schmidhuber. "Super Mario Evolution. In Computational Intelligence and Games", IEEE Symposium, CIG 2009, pp. 156-161.
- [28] J. K. Haas. "A History of the Unity Game Engine", 2014.
- [29] B. Tay, J. K. Hyun, S. Oh. "A machine learning approach for specification of spinal cord injuries using fractional anisotropy values obtained from diffusion tensor images". Computational and mathematical methods in medicine, 2014.
- [30] L. Breiman. "Random forests". Machine learning, 45(1), 2001, 5-32.
- [31] N. Sirikulviriyaya, S. Sinthupinyo. "Integration of rules from a random forest". In International Conference on Information and Electronics Engineering, Vol. 6, 2011pp. 194-198.
- [32] Y. LeCun, Y. Bengio, G. Hinton. "Deep learning". Nature, 521(7553), 2015, 436-444.
- [33] X. Yao. "Evolving artificial neural networks". Proceedings of the IEEE, 87(9), 1999, 1423-1447.
- [34] H. Guo, H. Nguyen, D. A. Vu, X. N. Bui. "Forecasting mining capital cost for open-pit mining projects based on artificial neural network approach". Resources Policy, 101474, 2019.

BIOGRAPHIES



YUNUS SARICA was born in Van, Turkey, in 1993. He received the B.S. and M.S. degrees in computer engineering from the Pamukkale University, Denizli. Since 2015, he has been working as a computer engineer. Since 2015, he had been a Research Assistant with the Computer Engineering Department, Pamukkale University,

Denizli, through the M.S. degree. His research interests include game programming, artificial intelligence, desktop and web application programming. Beside these, he is interested in amateur theatre acting, voice-over and contemporary visual technologies.



MERIC CETIN received B.Sc in Electrical & Electronics Engineering from Pamukkale University in 2003; M.Sc. in Electrical & Electronics Engineering from Institute of Natural Sciences, Pamukkale University, in 2006. She obtained Ph.D. degree in Electrical & Electronics Engineering from Pamukkale University in

2015. She is currently an Assistant Professor at the Computer Engineering Department, Pamukkale University. Her research interests are in model predictive control, machine learning, and adaptive control with computational intelligence techniques. Dr. Cetin is a member of the European Embedded Control Institute and reviewer of several international journals.

Determination of the Ideal Protocol Stack for the Transmission of Health Data over 6LoWPAN IoT Networks

S. BİLGİLİ and A. K. DEMİR

Abstract — It is expected that almost every day electronic devices will be connected to the existing internet infrastructure in the context of Internet of Things (IoT). These devices will enable to sense and actuate the physical world. It is foreseen that miniaturized e-health devices will enable monitoring vital health of patients. There exist some studies on networking these e-health devices within the Internet. In this realm, several network protocols are being standardized. 6LoWPAN of IETF is one of these efforts where some set of protocols can be stacked over IEEE 802.15.4 radio. However, it is not clear that which ideal protocol stack for transmission of health data can be adopted well. The novelty of this work is that we studied determination of ideal protocol stack for transmitting health data over 6LoWPAN IoT networks. So then, we carried extensive simulations over Cooja simulator. The compelling results are presented in this work. The results show that 6LoWPAN IoT health networks can be used to serve vital health data of patients.

Index Terms—Internet of Things, 6LoWPAN, Medium Access Control, Radio Duty Cycling, CoAP.

I. INTRODUCTION

TANY of the current everyday electronic devices in use are able to exchange data autonomously. These devices can network among themselves as well as connect to an existing Internet network. The concept of the Internet of Things is the general name given to this structure [1], [2]. Devices within this scope are mostly low-cost and limited devices. IoT devices generally obtain data from their surroundings with the help of sensors on them. These devices send the data they receive to a center for processing or perform some other operations. Connecting each of these devices directly to the Internet will incur additional costs. To avoid this, generally, IoT devices primarily form a network within themselves. A

gateway node within the scope of this network acts as a bridge between the Internet and the IoT network.


The IoT concept is currently used in smart home/office systems, factory production lines, smart agricultural systems, smart city/traffic systems. In the future, it is envisaged that IoT systems will become more widespread and expand into more areas. The lightweight and portable IoT devices has being introduced IoT concept into the healthcare sector. Especially for health monitoring, their portability and low-power features have made IoT devices the preferred choice.

Although the IoT concept is currently in use in many areas, it is still a developing technology. Therefore, there are no emerged de facto standards yet. Although wireless technologies such as Wi-Fi, Bluetooth and GSM can also be used for the IoT environment, the power saving offered by 802.15.4 wireless technology is more suitable and common for IoT systems. For energy-saving IoT devices, hardware is also constrained to keep the energy consumption low. Therefore, instead of protocols with high system requirements, lightweight protocols are being proposed and standardized. One of these efforts is 6LoWPAN working group of IETF. 6LoWPAN seems like providing promising IoT network standards and protocols [3].


Typical 6LoWPAN standard network layers and protocols of IoT devices is given in **Hata! Başvuru kaynağı bulunamadı..** In this layered structure, the protocols in the transport, network and application layers are almost standardized and widely used. However, the protocols used in the data link layer needs to be selected according to application quality of service (QoS) requirements. Therefore, selection of ideal protocol stack for applications requires careful research.

In this work, careful determination of the ideal protocol stack for e-Health application is investigated. As is known, e-Health applications are sensitive to delays. Therefore, timely transmission and hand off e-Health data to a distant center are core QoS requirements of e-Health applications. Besides, an e-Health application may need different health sensors such as ECG, EMG, Blood Pressure, Body Temperature and Body Position etc. Each health sensor requires different maximum delay thresholds. This can vary from seconds to hours. Moreover, maximum delay thresholds may also vary for different patient categories such as critical, non-critical and follow-up patients. Thus, to achieve this, we investigated the performance of different protocol stack of MAC and RDC

SEDAT BİLGİLİ, is with Department of Computer Engineering Adana Alparslan Türkeş Science and Technology University, Adana, Turkey, (e-mail: sbilgili@atu.edu.tr).

 <https://orcid.org/0000-0001-6708-7209>

ALPER KAMİL DEMİR, is with Department of Computer Engineering Adana Alparslan Türkeş Science and Technology University, Adana, Turkey, (e-mail: akdemir@atu.edu.tr).

 <https://orcid.org/0000-0002-9256-0368>

Manuscript received November 07, 2019; accepted March 16, 2020.
DOI: [10.17694/bajece.644104](https://doi.org/10.17694/bajece.644104)

layer of an e-Health application for different patient categories so that patient’s vital health data can be monitored smoothly.

Transmission of e-Health data over IoT networks is fairly a new research subject. Thus, there does not exist satisfying results in the literature and applications. The novelty of this work is that e-Health data is transmitted with standard CoAP application layer protocol and the effect of performance belonging to MAC and RDC layer protocols, such as CSMA and contikiMAC, are explored with simulation of real network and hardware. Furthermore, the network topology and the number of patients (clients) equipped with e-Health sensors are varied with extensive simulations.

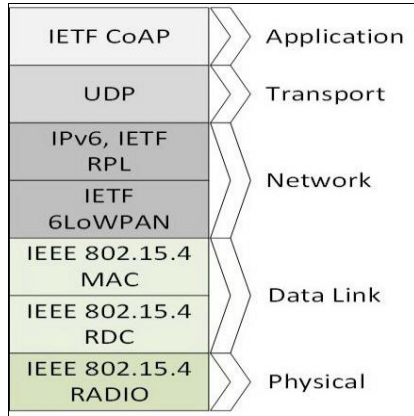


Fig.1. Network Layers in IoT

This paper is organized as follows: In section II we surveyed related works that are associated to subject of this paper. Our methods are provided in section III. Based on these methods, we investigated results in section IV. In section V, we provided conclusion and future work.

II. RELATED WORK

The transmission and processing of vital health data in IoT environments are emerging research subject that is being explored scarcely yet. The IoT for healthcare is broadly surveyed in [4], [5]. However, there does not exist a comprehensive survey of transmitting health data in 6LoWPAN IoT networks. Moreover, there does not exist concrete study on transmitting health data in 6LoWPAN IoT networks. We briefly talk about networking health data over 6LoWPAN IoT networks in this section.

Sphere framework [6] brings 6LoWPAN IoT standards for healthcare to the home. The initial outcomes exhibit good 6LoWPAN network performance to carry out healthcare data (99.97 percent average PDR).

IoT net platform [7] anticipates technological solutions for healthcare protection and services through trendy 6LoWPAN IoT networks. In [8], again, analysis of 6LoWPAN IoT network is provided for maternal healthcare. It is concluded that both CoAP and 6LoWPAN could be applied for healthcare monitoring. The research in [9] demonstrates that 6LoWPAN IoT network is able to be used for healthcare services with CoAP.

The initial research views that 6LoWPAN IoT networks are promising solution to be used in healthcare applications. However, as it is seen that the usage of 6LoWPAN IoT

networks in healthcare domain is an unexplored area. We extend this gap furtherly in this work.

III. METHODOLOGY

Within the scope of this section, we present the methodology of our work to transmit vital health data over a 6LoWPAN IoT network. The details are as follows:

A. Obtaining Health Data

Although performance evaluations of this work is executed on a simulation environment, real world health data is required to get realistic results. To achieve this, MySignals health sensor kit is used with various additional hardware, Arduino Uno microcontroller and Raspberry Pi microcomputer as seen in Fig. 1. MySignals health sensor kit is programmed with a simple firmware to get all sensor data. Various sensor data is collected from various persons to ensure collected data is realistic.

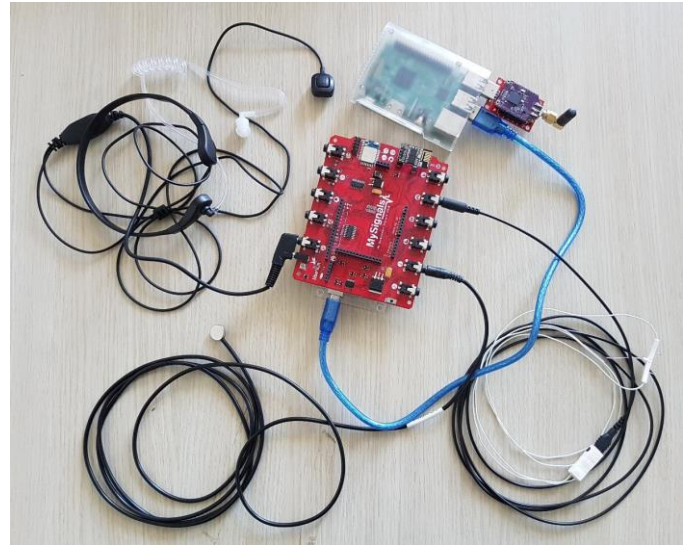


Fig. 1. eHealth Sensor Kit

B. Health Data Traffic Characteristics

Right after obtaining real-world health data, each sensor data is composed into a packet. Each health data is encapsulated into 64-byte packet. Each packet contains data from all available health sensors. Sensor data types and their length is given in Table I. In this work, although we used all available health sensors, there's 20 bytes free space in packet which can be used for patient information or extra sensors. These eHealth data packets are updated in memory dynamically every 300ms. The clients send CoAP CON messages to obtain eHealth data. Respectively, the server nodes send piggybacked CoAP ACK messages that contain eHealth data whenever they receive a CoAP CON message.

Table I Contents of eHealth Data Packet

ID	Sensor	Bytes
----	--------	-------

1	Body Position	6
2	Body Temperature	2
3	Snore Sensor	3
4	Galvanic Skin Response	8
5	Airflow Sensor	2
6	EMG Sensor	2
7	ECG Sensor	2
8	SPo2	4
9	Blood Pressure Sensor	6
10	Body Scale	2
11	Glucometer	6
12	Patient ID	1
	Total	44
	Free Space	20

C. Latency QoS Requirement of Health Data

We envisioned that each eHealth sensor data needs to be transmitted within a maximum latency deadline. Thus, we talked a couple of medical doctors to detect latency deadline of eHealth data. According to interviewed medical doctors, patients should be divided into three groups, critical, non-critical and follow-up patients, as they may have need different latency deadlines. As a result, maximum latency deadline values are given in Table II. These values represent the maximum latency tolerance of eHealth data generated by health sensors. For example, a critical patient's consecutive ECG sensor data needs to be transmitted within 1 minute. On the other hand, ECG sensor data demand maximum 60 minutes latency for non-critical and follow-up patients. Moreover, body position data requires maximum 30 minutes latency for critical and non-critical patients.

Table II eHealth Data Maximum Latency Deadlines

Sensor	Patient Category		
	Critical	Non-Critical	Follow-up
Body Position	30	30	15
Body Temperature	15	30	60
Snore Sensor	30	60	60
Galvanic Skin Response	60	60	60
Airflow Sensor	30	30	60
EMG Sensor	1	60	60
ECG Sensor	1	60	60
SPo2	1	60	60
Blood Pressure Sensor	30	60	60
Body Scale	60	60	60
Glucometer	60	60	60
Deadline 1: 1 minute			
Deadline 2: 15 minutes			
Deadline 3: 30 minutes			
Deadline 4: 60 minutes			

D. Simulation Environment

We used Cooja [10] network simulator that simulates multiple types of nodes and network software running on nodes. Cooja simulator mimics ContikiOS [11] operating system designed for IoT devices. Also, Cooja network simulator offers multiple measurement tools. WisMote is used as a hardware server and gateway node in simulations within IoT network. These server nodes generate sampled eHealth data traffic. Californium [12] is used as CoAP client running on real PC hardware and requiring eHealth data piggybacked in CoAP ACK messages. In other words, CoAP servers, running on simulated IoT network, are programmed with Erbium [13] implementation of CoAP, and CoAP clients, running on real PC, are programmed with Californium implementation of CoAP. The simulated IoT network topology is given in Fig. 2. The distance between each node is about 40 meters.

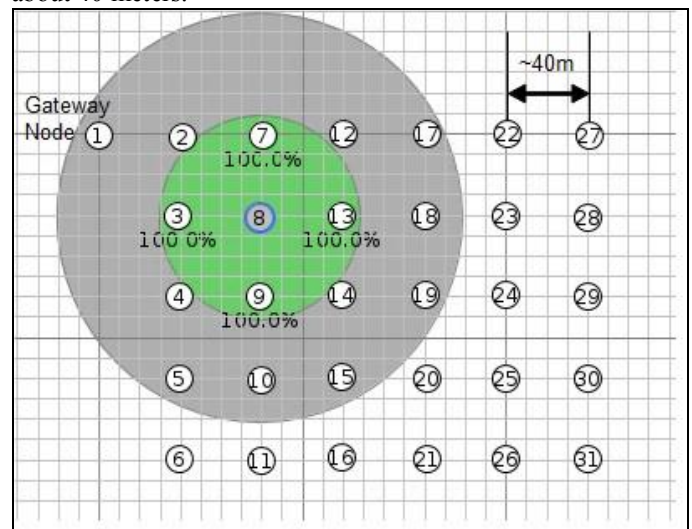


Fig. 2. IoT Network Topology in Simulations

CSMA and nullMAC protocols are varied at MAC layer. And, contikiMAC and nullRDC protocols are varied at RDC layer. Moreover, PDR values are ranged as 90, 95 and 100. This enabled us to simulate 12 different scenarios (2x2x3). The number of clients is varied from 2 to 30 stepped by 2 (2,4,6,...,30). The number of servers is always the same as the number of clients. The positions of servers are detected according to their relative position to the gateway node. Always, the closest nodes are selected as a server. Each client sends totally 100 CoAP CON requests. Each server replies with CoAP ACK message as soon as it gets CoAP CON request. The total elapsed time is to transmit 100 CoAP CON messages. According to CoAP standard, the default maximum re-transmission of lost CoAP CON packets is 4 times. One thing to keep in mind is that CoAP handles a default congestion control mechanism.

E. Performance Metrics

To inspect how well health data is transmitted in an IoT network, we need some measurable performance metrics. In this sense, we identified 3 different performance metrics, latency, energy efficiency and reliability. We give the definition of each performance metric below.

1) Latency

Latency metric defines the time elapsed between two consecutive CoAP ACK messages. As the subject of this work is health data, it means that the data must be transmitted within deadline time limits. These time limits are shown in Table II.

2) Energy Efficiency

Energy efficiency metric is used to represent the amount of energy consumed by a server node. This metric is calculated as the average of consumed energy of all server nodes. Indirectly, this metric shows the overall network life time.

3) Reliability

Reliability metric demonstrates the percentage of successfully received CoAP ACK messages. Higher reliability means that eHealth data is smoothly displayed at health center. The low reliability may result with low perceived health data.

IV. PERFORMANCE EVALUATION

In this section, we investigate the performance of 12 scenarios, given at Section III.D based on 3 different performance metrics.

A. Latency Evaluation

As stated earlier, health data should be transmitted within a certain period of time. Time limits for different health sensors and patient groups are shown in Table II. The graphs in this section show the performance indicators as well as the latency deadline indicators. Thus, it is possible to determine how many clients can be supported by the specified latency deadline in a given protocol stack. Please note that if there's no latency deadline indicator for a performance indicator, it means that all results in the graph comply with absent latency deadline.

Various simulations were performed for different protocol stacks. **Fig. 3** shows the latency graph with protocol stack of nullMAC and nullRDC in MAC and RDC layers respectively. As can be seen in the graph, all results are lower than deadline of 30 and 60 minutes (Hence, these deadline indicators are not shown in the graphs). All number of clients, other than 26 and 30, comply with deadline limit of 15 minutes. Therefore, the network can support up to 24 clients. For example, body temperature sensor data can be used safely for all patient groups up to 24 patients. However, for more critical data with deadline limit of 1 minute, only 8 clients can be serviced when PDR equals to 100.

Fig. 4 shows latency graph of CSMA and nullRDC protocols. As it is seen, deadline limits of 15, 30 and 60 minutes are missing in the graph because all latency values are lower than 15, 30 and 60 minutes. Therefore, only 1 minute latency indicator is given in the graph. For all PDR values, up to 16 clients can be monitored without any trouble. However, for example, when the client number is above 20 the network is not able to support health data with 1 minute latency deadline.

In **Fig. 5**, nullMAC and contikiMAC combinations are used for latency deadlines. Accordingly, for 30 minute latency, only up to 18 clients can be handled. Besides, for 15 minute latency, individually up to 8 clients can be supported. Nevertheless, for 1 minute latency, only 2 clients with PDR values 100, can be serviced.

CSMA and contikiMAC protocol combination latency graph which is given in **Fig. 6** shows that 24 clients can be monitored for latency deadline of 30 minutes. Based on the figure, for latency deadline of 15 minutes, there can be only 16 clients. More drastically, latency deadline of 1 minute can only support 4 clients.

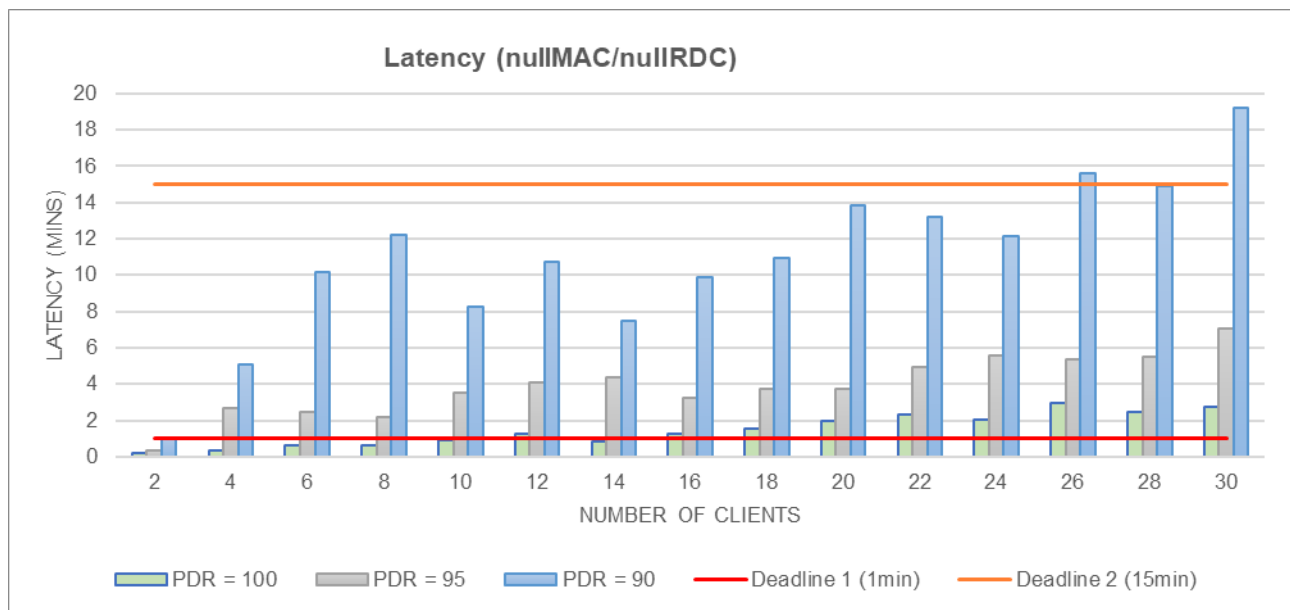


Fig. 3. Latency Graph for nullMAC/nullRDC Protocols

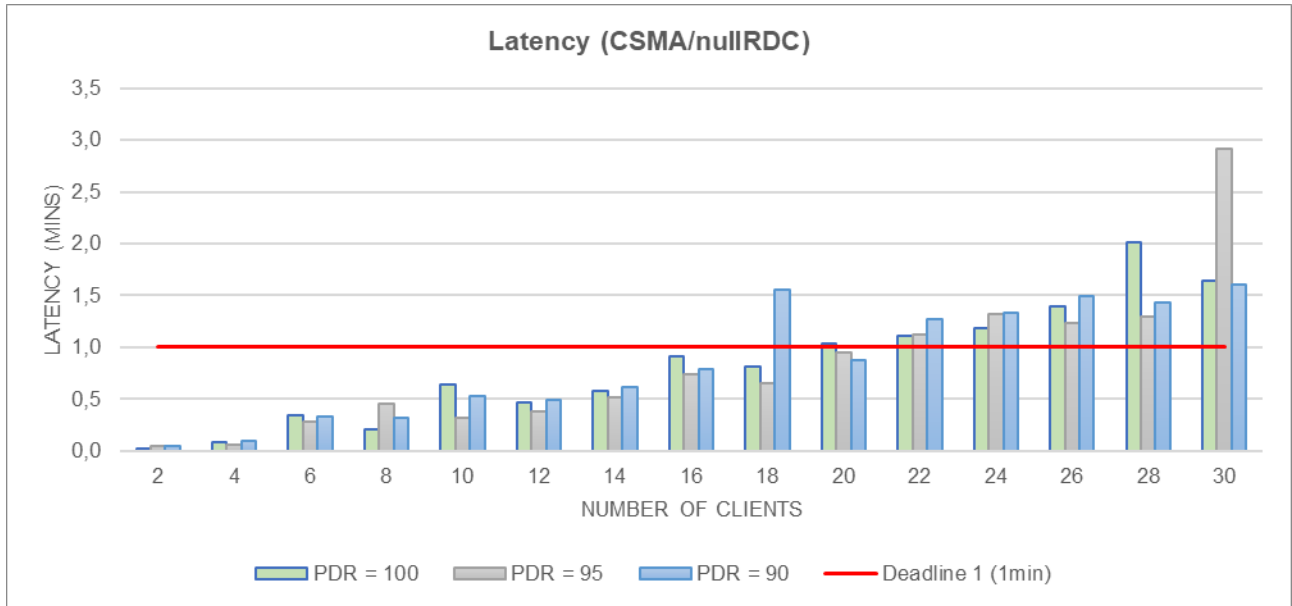


Fig. 4. Latency Graph for CSMA/nullRDC Protocols

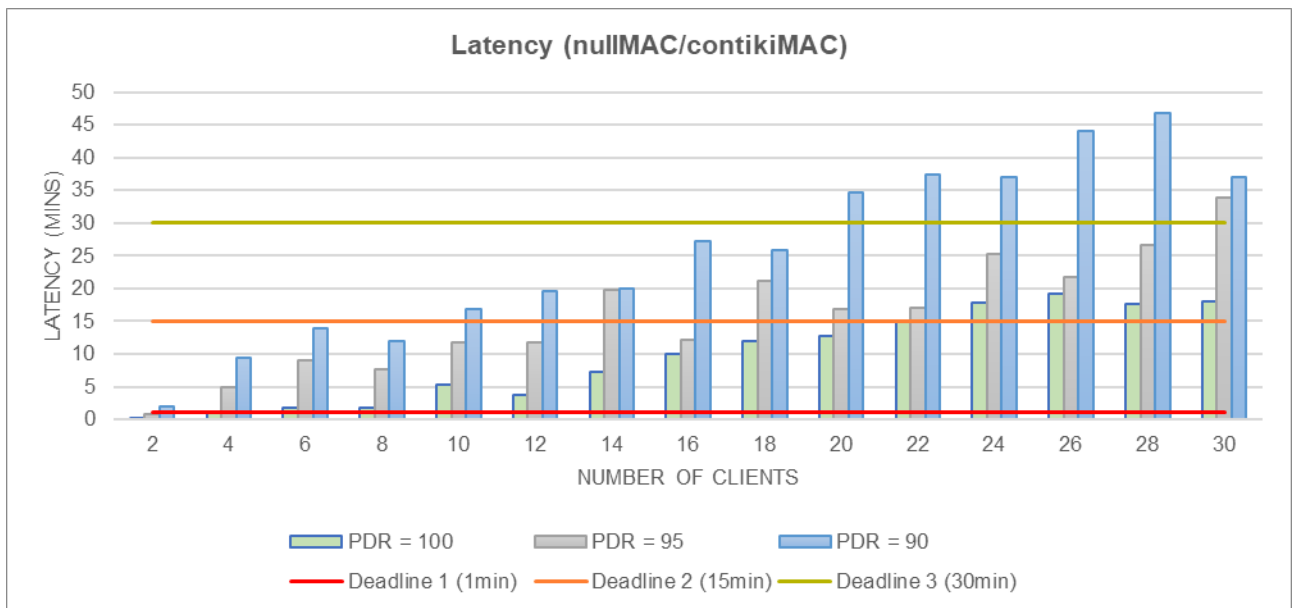


Fig. 5. Latency Graph for nullMAC/contikiMAC Protocols

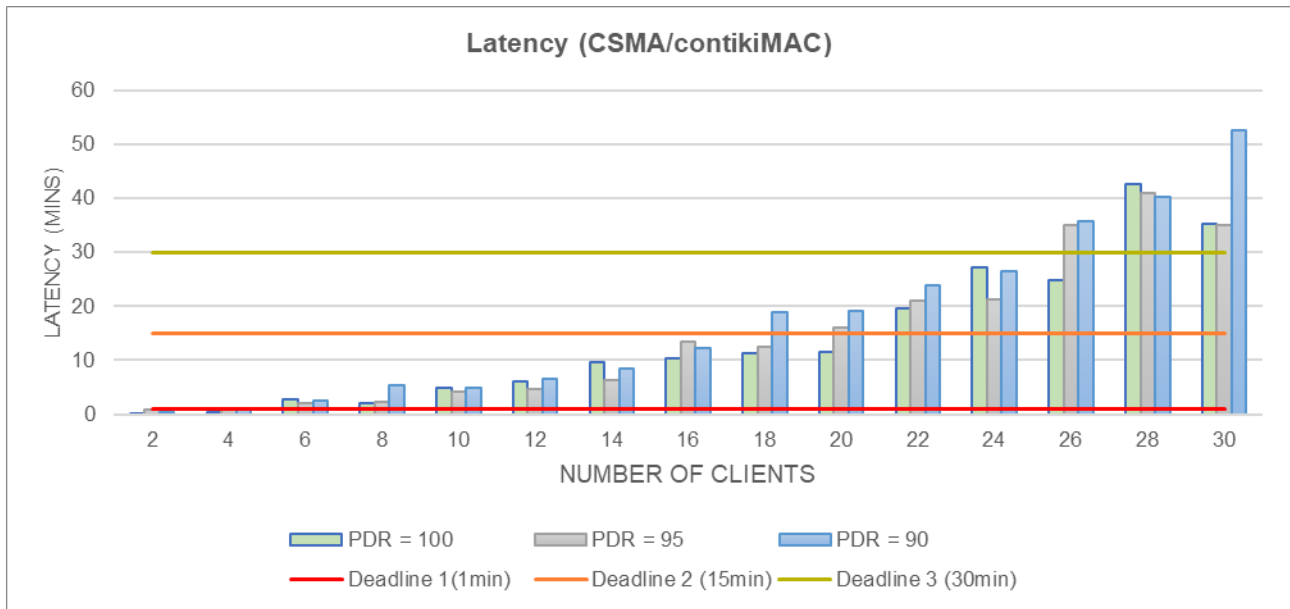


Fig. 6. Latency Graph for CSMA/contikiMAC Protocols

B. Energy Efficiency Evaluation

In this section, we investigate the energy efficiency of MAC and RDC layer protocol combinations.

Fig. 7 shows the energy consumption comparison of 4 protocol combinations when the PDR value is 100. Apparently, when nullRDC is used, the energy consumption significantly increases. Because while using nullRDC, 802.15.4 radio does not sleep when it is idle. Thus, it uses more energy. In contrast, contikiMAC sleeps the radio while it's idle. The effect of this mechanism can also be observed in the graph. Using CSMA or nullMAC protocol with nullRDC has almost the same energy consumption. However, while using contikiMAC protocol, selecting CSMA protocol causes

more energy consumption against selecting nullMAC protocol.

Energy consumption values while PDR value is 95 is given in **Fig. 8**. Again, selection of contikiMAC in RDC layer significantly decreases energy usage. The results show similar curve with previous graph. However, difference between nullMAC/nullRDC and CSMA/nullRDC combinations is much more distinctive when PDR is 95.

In **Fig. 9**, energy consumption values for PDR value 90 is presented. Based on the graph, again, it is clear that contikiMAC protocol provides more energy savings. As can be seen, for protocol combinations with nullRDC, energy consumption values get higher as the PDR value decreases.

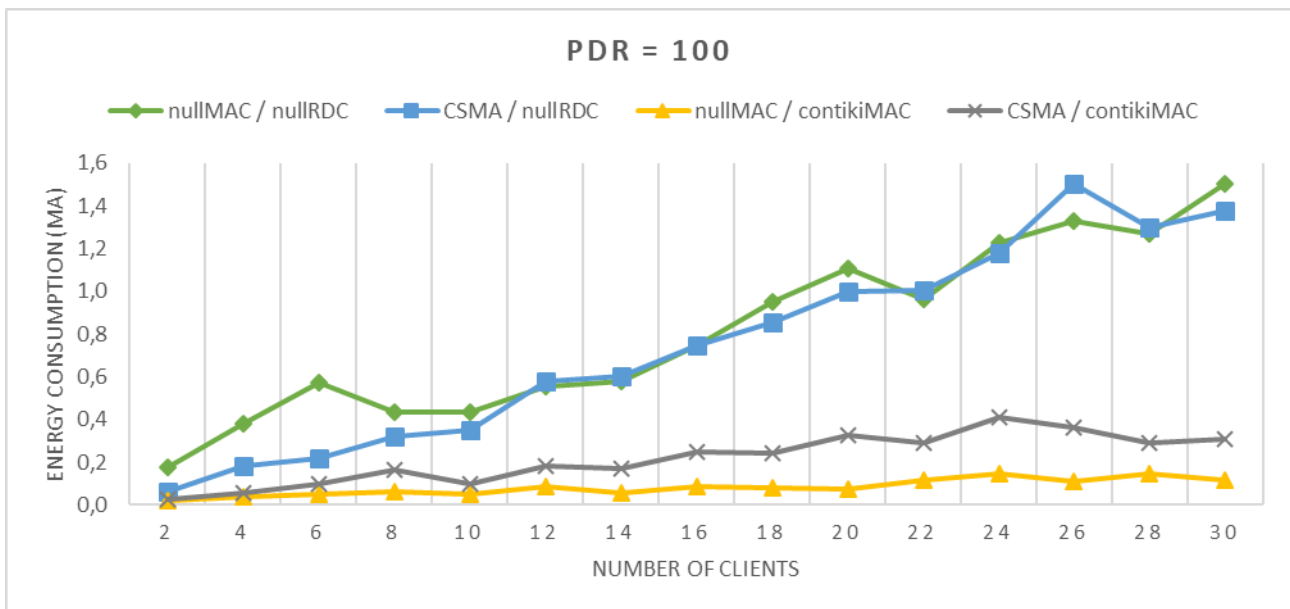


Fig. 7. Average Energy Consumption for PDR value = 100

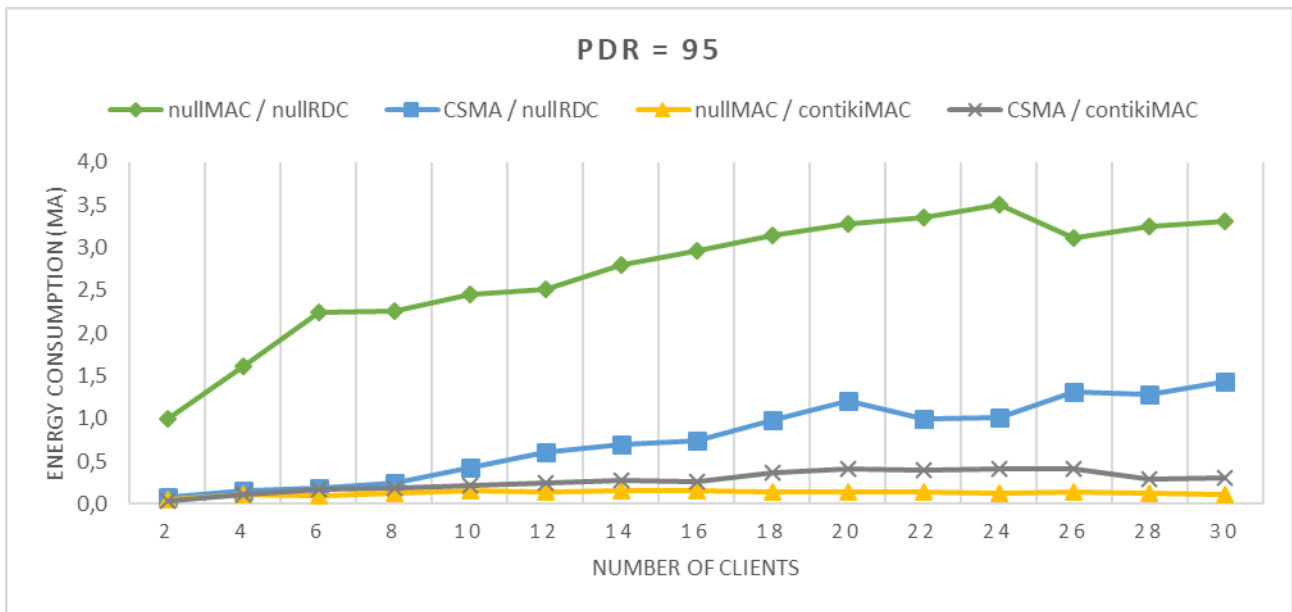


Fig. 8. Average Energy Consumption for PDR value = 95

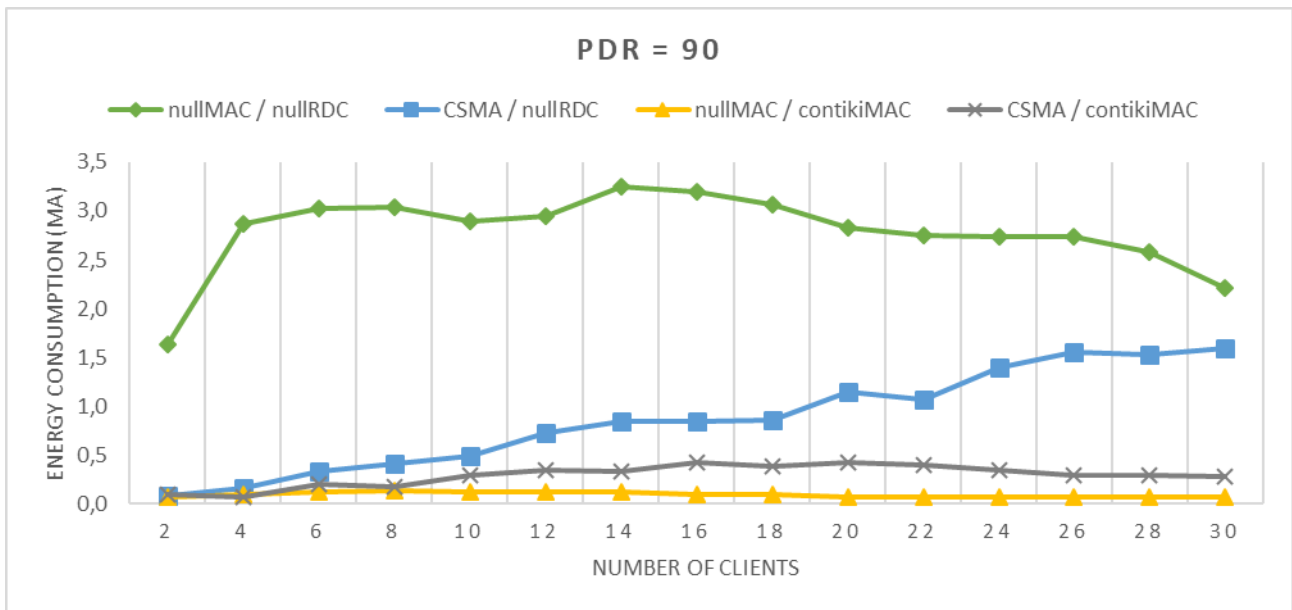


Fig. 9. Average Energy Consumption for PDR value = 90

C. Reliability Evaluation

In this section, reliability metric values are presented. These values show the rate of successfully transmitted health data packets. The four protocol combinations (nullMAC/nullRDC, CSMA/nullRDC, nullMAC/contikiMAC and CSMA/contikiMAC) are compared with different PDR values.

Reliability values when the PDR value is 100 is given in

Fig. 10. That graph shows that CSMA and nullRDC protocol combination gives better results. Besides, nullMAC and nullRDC protocol combination has results that are close to CSMA and nullRDC protocol combination. However, nullMAC and contikiMAC protocol combination cannot sustain reliability as the client number increases. CSMA and

contikiMAC protocol combination gives the worst reliability especially for high number of clients.

PDR value of 95 is given in

Fig. 11. As the client number increases, the gap between nullMAC/nullRDC and CSMA/nullRDC protocol combinations increases. CSMA/nullRDC gives the best performance compared to other three combinations. The graph shows that using CSMA protocol is helpful as CSMA maintains reliability high.

When PDR value equals 90, the same reliability performance is observed. Results are presented in Fig. 12. Once more, absence of CSMA protocol in protocol stack causes lower reliability results. Using CSMA is superior than using nullMAC cause CSMA protocol successfully

retransmits lost frames. Using CSMA protocol with contikiMAC protocol gives undesired results when the number of clients increases. This is due to the fact that contikiMAC

protocol sleeps radio to save energy, so CSMA protocol must wait for these sleep intervals.

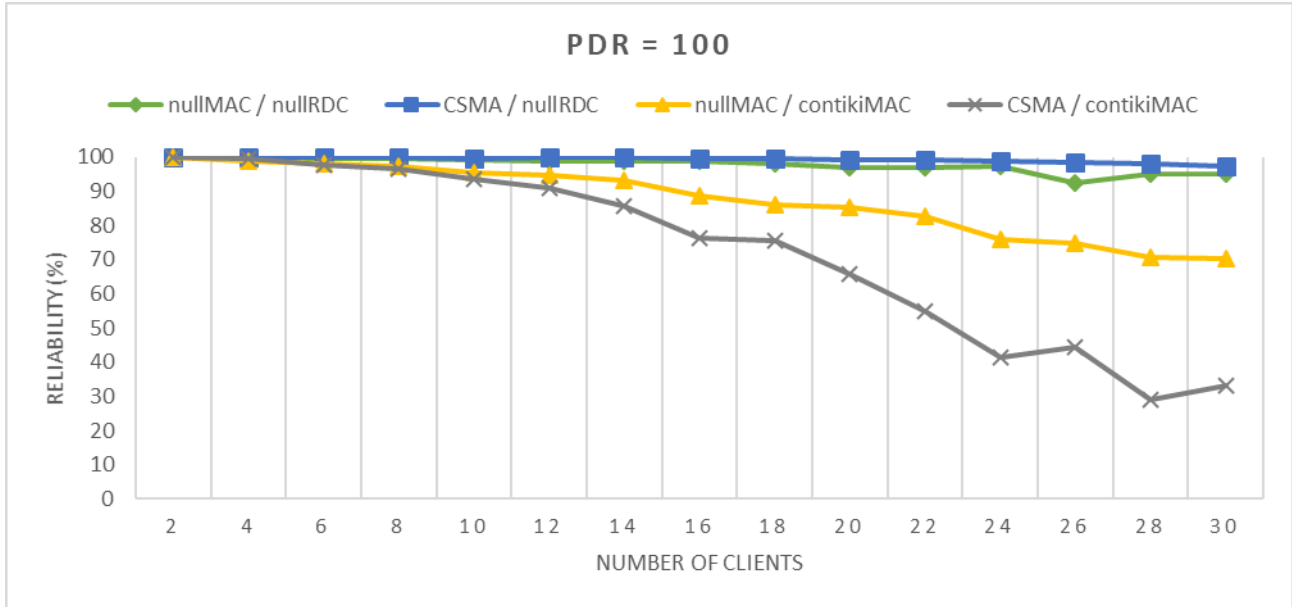


Fig. 10. Success Ratio/PDR value = 100

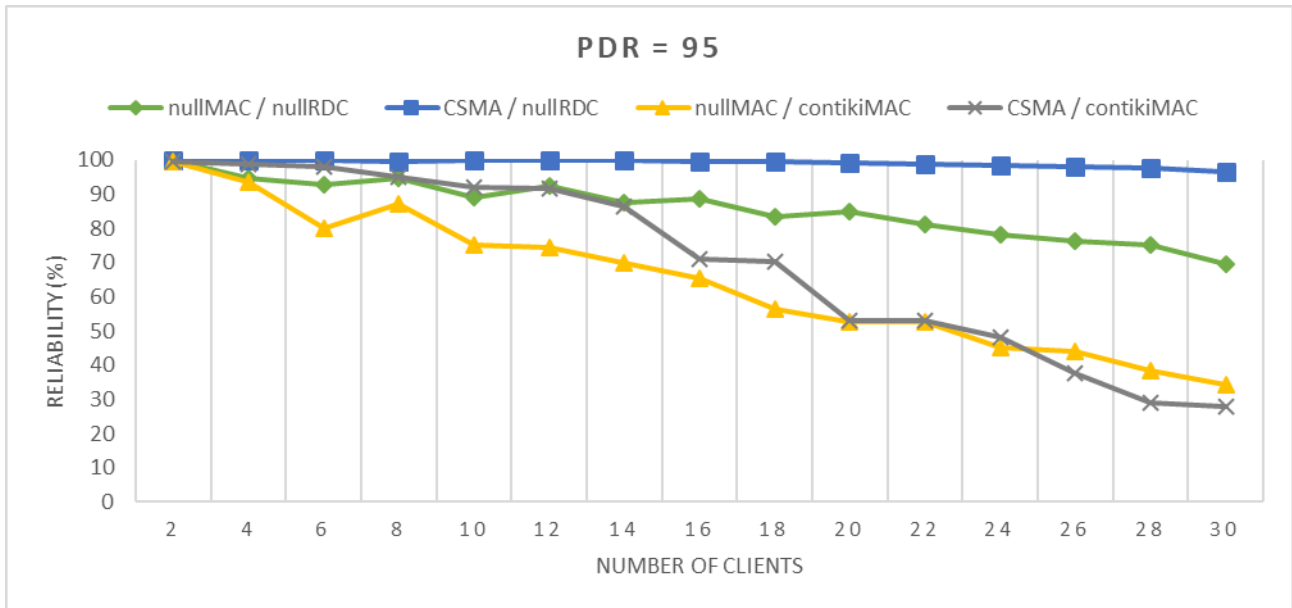


Fig. 11. Success Ratio/PDR value = 95

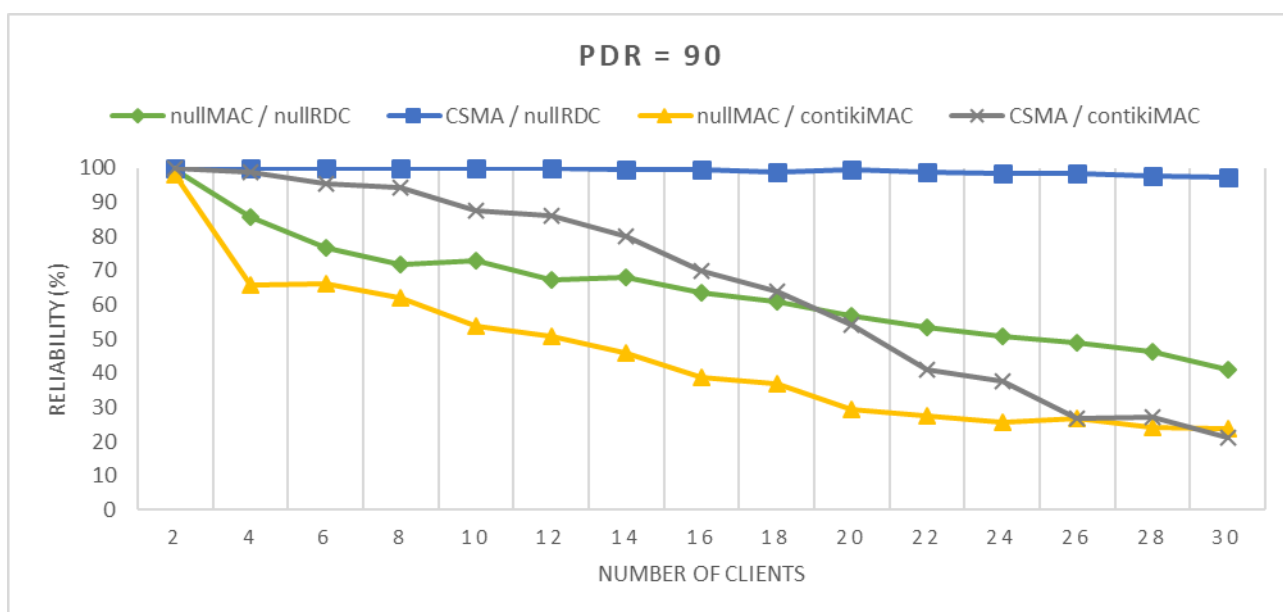


Fig. 12. Success Ratio/PDR value = 90

V. CONCLUSION AND FUTURE WORK

In this work, we investigate determination of ideal protocol stack for transmission of health data over 6LoWPAN IoT networks. Our extensive work shows that 6LoWPAN IoT networks are able to transmit vital health data of patients up to a certain point. Latency, reliability and energy efficiency graphs are analysed to determine the maximum patient number that can be supported in a 6LoWPAN IoT network.

Ideal protocol combination may differ based on patient category and health sensor. Although nullRDC protocol in RDC layer seems to show better performance, it results with inefficient energy consumption. Therefore, choosing ideal protocol combination highly depends on latency, reliability and energy efficiency performance results. For instance, even if CSMA/nullRDC combination can support up to 16 patients for critical health data in 6LoWPAN IoT network, energy consumption significantly gets higher.

In future, the more extensive simulations can be carried out by varying IoT network topology and PDR values. Routing protocol, RPL (especially RPL Objective Functions), is not investigated in this work to expose its effect on the performance. Moreover, only CoAP is used as an application layer protocol as it is the only candidate for now in IETF. Other being standardized protocols, such as MQTT, can also be considered. The least but not the last, CoAP promotes other congestion control mechanisms, such as CoCoA. Using different congestion control mechanisms to determine ideal protocol stack can also be inspected.

ACKNOWLEDGMENT

This work was supported by SIREN project funded by the Scientific and Technological Research Council of Turkey (TUBITAK) under Grant No 116E025 and by Tiktak project funded by Scientific Research Projects (BAP) Coordination of Adana Alparslan Turkes Science and Technology University under Grant No 16103002.

REFERENCES

- [1] L. Atzori, A. Iera, and G. Morabito, "The internet of things: A survey," *Comput. networks*, vol. 54, no. 15, pp. 2787–2805, 2010.
- [2] V. Gazis, "A Survey of Standards for Machine-to-Machine and the Internet of Things," *IEEE Commun. Surv. Tutorials*, vol. 19, no. 1, pp. 482–511, 2017.
- [3] Z. Shelby and C. Bormann, "Introduction, in 6LoWPAN: The Wireless Embedded Internet," in *6LoWPAN*, John Wiley & Sons, Ltd, 2009, pp. 1–25.
- [4] D. A. Gandhi and M. Ghosal, "Intelligent Healthcare Using IoT: A Extensive Survey," *Proc. Int. Conf. Inven. Commun. Comput. Technol. ICICCT 2018*, no. Icicct, pp. 800–802, 2018.
- [5] H. Ahmadi, G. Arji, L. Shahmoradi, R. Safdari, M. Nilashi, and M. Alizadeh, *The application of internet of things in healthcare: a systematic literature review and classification*, vol. 18, no. 4. Springer Berlin Heidelberg, 2018.
- [6] A. Elsts *et al.*, "Enabling Healthcare in Smart Homes: The SPHERE IoT Network Infrastructure," *IEEE Commun. Mag.*, vol. 56, no. 12, pp. 164–170, 2018.
- [7] E. L. Lydia, K. Shankar, M. Ilayaraja, and K. S. Kumar, "Technological Solutions for Health Care Protection and Services Through Internet Of Things(IoT)," *Int. J. Pure Appl. Math.*, vol. 118, no. 7 Special Issue, pp. 277–282, 2018.
- [8] K. Kabilan, N. Bhalaji, and S. Chithra, *Analysis of 6LoWPAN and CoAP protocols for maternal health care*, vol. 521. Springer Singapore, 2019.
- [9] J. Hwa Jung, D. Kyu Choi, J. In Kim, and S. Joo Koh, "Mobility management for healthcare services in coap-based iot networks," *Int. Conf. Inf. Netw.*, vol. 2019-Janua, pp. 7–12, 2019.
- [10] A. Sehgal, "Using the contiki cooja simulator," *Comput. Sci. Jacobs Univ. Bremen Campus Ring*, vol. 1, p. 28759, 2013.
- [11] A. Dunkels, O. Schmidt, N. Finne, J. Eriksson, F. Österlind, and N. T. M. Durvy, "The contiki os: The operating system for the internet of things," *Online*, <http://www.contikios.org>, vol. 605, 2011.
- [12] M. Kovatsch, M. Lanter, and Z. Shelby, "Californium: Scalable cloud services for the internet of things with coap," in *2014 International Conference on the Internet of Things (IOT)*, 2014, pp. 1–6.
- [13] M. Kovatsch, S. Duquenooy, and A. Dunkels, "A low-power CoAP for Contiki," *Proc. - 8th IEEE Int. Conf. Mob. Ad-hoc Sens. Syst. MASS 2011*, pp. 855–860, 2011.

BIOGRAPHIES



SEDAT BILGILI is a Research Assist. at Computer Engineering Department of Adana Alparslan Turkes Science and Technology University. He is also a PhD. student at department of Electrical and Electronics Engineering, Adana Alparslan Turkes Science and Technology University. He received his BS degree from Cukurova University,

Adana, Turkey in 2015. He worked at BSD Inc. as a Network Administrator between 2015–2016. He receives MSc. Degree from Adana Alparslan Turkes Science and Technology University, Adana, Turkey in 2019. His master thesis was on Healthcare in Internet of Things. He is interested in Internet of Things, Information Centric Networking, Embedded Systems and Computer Networks. He is also interested in Arduino, Raspberry Pi, general sensors and electronics as a hobbyist.



ALPER K. DEMIR is an Assist. Prof. at Computer Engineering Department of Adana Alparslan Turkes Science and Technology University. He received his BS degree from Hacettepe University, Ankara, Turkey and MS degree from University of Southern California (USC), Los Angeles, USA, and PhD degree from

Kocaeli University Kocaeli, Turkey in 1993, 1998 and 2013 respectively. Between 2009–2013, he worked at Huawei Telecommunications Inc. as a Senior Software and Research Engineer. Between 2001–2009 he worked at Kocaeli University, Computer Engineering Department as a Senior Instructor. He is interested in Computer Networks, Distributed Systems and Operating Systems in general. Currently, his research is focused on Internet of Things.

Determination of Leakage Inductance Percentage for Gapped Iron-Core Shunt-Reactors with M4 Steel as Core Material

A. DÖNÜK

Abstract—Leakage inductance component has a significant importance in total inductance value of GISR. Neglecting this component in the design phase, results in an expensive and bulky core structure. Variation of leakage inductance component in percentage is determined and presented as graphical curves for M4 steel by applying energy method. FEA are performed for various GISR with several operating voltages and temperature rise values to determine the leakage inductance component. A design tool with Matlab/Guide is also developed for analytical calculations to obtain the physical dimensions for FEA. Graphical curves introduced to the literature in this work provide manufacturers or design engineers to perform fast, reliable and economical GISR design with an alternative material and offer variety.

Index Terms— Air-gap, design, leakage inductance, M4 steel, shunt reactor

I. INTRODUCTION

SHUNT REACTORS are widely used in power grids for reactive compensation, harmonic filtration etc. They mostly have gapped iron core with distributed air-gaps along the limb. Equivalent circuit for a Gapped Iron-core Shunt-Reactor (GISR) is given in Fig. 1, where L_l is leakage, L_c is core and L_g is total air-gap inductance, whereas R_w , R_c and R_g are the resistances representing winding, iron and gap losses [1].

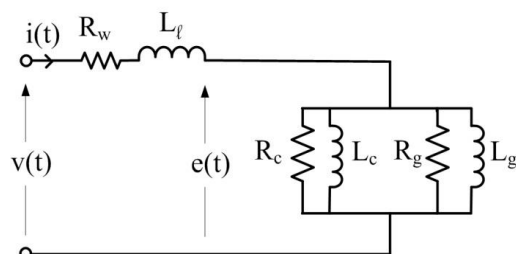



Fig.1. Equivalent Circuit for a GISR

A simple and practical analytical calculation for leakage inductance component of GISR does not exist in the literature;

ATILLA DÖNÜK is with Department of Electrical & Electronics Engineering of Aydın Adnan Menderes University, Aydın, Turkey, (e-mail: adonuk@adu.edu.tr).

 <https://orcid.org/0000-0001-9468-8456>

Manuscript received November 13, 2019; accepted April 23, 2020.
DOI: [10.17694/bajece.646625](https://doi.org/10.17694/bajece.646625)

therefore this component is mostly neglected in the design phase and analysis. Influence of core gap in design of current limiting transformers is presented in [2] but leakage inductance component is not considered. Leakage inductance is neglected in another shunt reactor design study [3]. Although the leakage fields are considered in [4] determination of leakage inductance is not studied. Leakage inductance component is usually neglected in the studies on design of GISR including influence of dimensional parameters, design optimization, stress characteristics, etc. [5-8]. Finite Element Analysis (FEA) is widely used in GISR studies [9-12]. Determination of inductance components including leakage inductance by FEA is performed in [1] and [12], where calculation is based on energy method. A family of graphical curves called nomographs representing the variation of leakage inductance component in percentage against reactive power in terms of four different operating voltages and three different temperature rise values are presented in [12,13] for M330-35 AP Non-Grain-Oriented (NGO) steel (will be called M33 steel hereafter) by FEA. The proposed design criteria in this study is minimum Present Value Cost (min PVC). Presented results show that ratio of leakage inductance in the total amount may be significant. Therefore, neglecting the leakage inductance component in the design phase will result in a bulky and costly core.

In this work, leakage inductance components for M120-27S Grain-Oriented (GRO) steel (will be called M4 steel hereafter) by the method proposed in [12, 13] are calculated and presented as graphical curves. A design tool for single-phase GISR in Matlab/Guide is developed for analytical calculations. As a result, physical dimensions of the GISR under design are obtained for FEA. FEA are performed for various GISR with several operating voltages and temperature rise values to determine the leakage inductance component.

Since such information in the literature exist only for M33 steel as core material, the results presented in this work are beneficial for the design of GISR with M4 steel as an alternative core material. In addition, the developed design tool provides a practical and accurate design of such reactors.

II. METHODOLOGY

A. Design Tool for Analytical Calculations

Leakage inductance percentage values are calculated by FEA, therefore physical dimensions of the GISR under design are

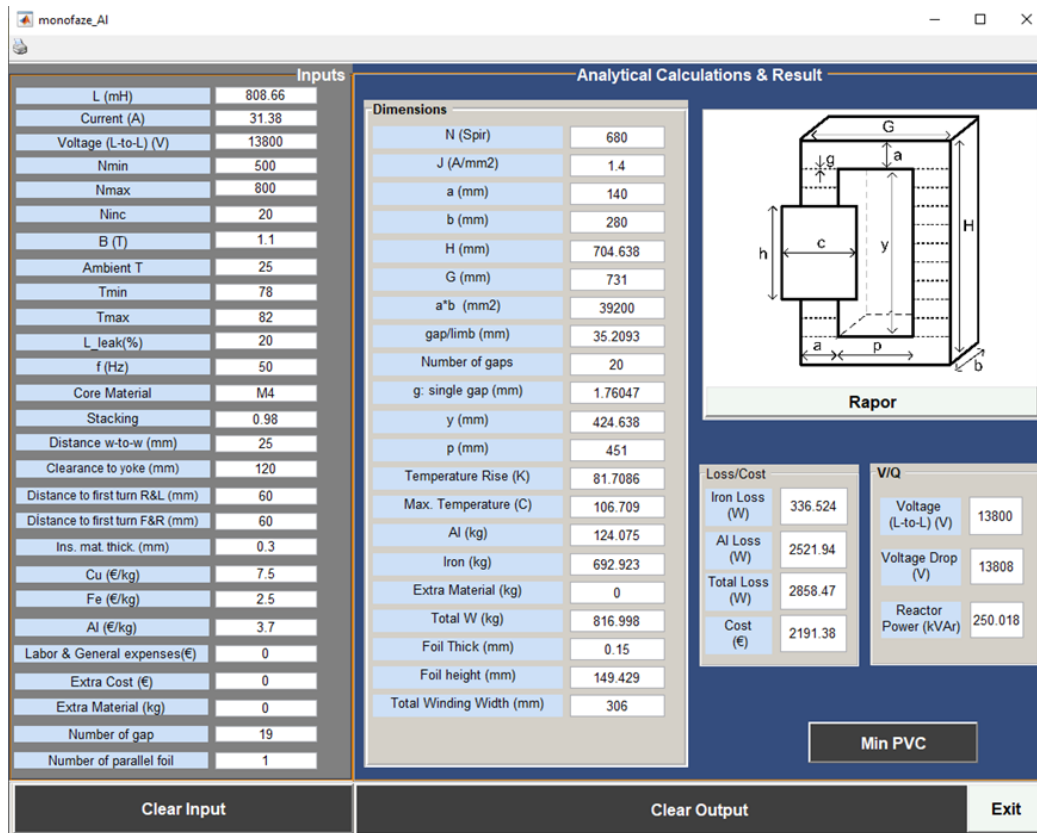


Fig.2. General overview of the design software

required for proper modelling in Maxwell 3D software. To obtain the physical dimensions, a design tool ware for analytical calculations is developed in Matlab/Guide environment as shown in Figure 2. The input data such as target inductance value, operating current and line-to-line voltage, minimum and maximum values of number of turns with incremental step, magnetic field density, ambient temperature, tolerances for temperature rise, leakage inductance percentage, operating frequency, type of core material, stacking factor, distances and clearances between coils and yoke, insulation material thickness, per-kg prices for core and winding materials, number of air-gaps and number of parallel foils for the GISR under design are to be entered by user. Extra data for labor and general expenses and extra cost of additional material can also be added.

Similar assumptions, for each set of operating voltages, such as

- Clearances
 - between yokes
 - between windings
 - between tube and windings
- insulation thickness
- tube thickness
- temperature rise values (60, 80 and 100 K with ± 2 K tolerances)
- number of air-gaps in the core (40 in both the design and FEA analysis)
- aluminum foil as winding material
- operating magnetic flux density as 1.1 Tesla

- prices of iron and aluminum are taken to be 2.5 and 3.7 € per kg, respectively
- optimization criteria (min PVC)

defined in [12] are considered in the design phase. This will provide comparison of the leakage inductance percentage results in two different core materials, M33 and M4.

Magnetization and loss curves for M4 steel provided by the manufacturer are given in Figure 3. B-H curve required for FEA analysis software is defined from the magnetization curve in Figure 3. Core loss calculation in the analytical design phase is performed as defined in Equation (1) which is obtained via curve fitting of the total loss characteristic given in Fig.3.

$$P_c = 0.41 B^{1.7973} \text{ W/kg} \quad (1)$$

After completing data entry, by pressing Min PVC button on the software screen given in Fig.2, the analytical calculation is started. The procedure is as follows:

Four different loops, from outer to inner, exist in the analytical calculation, such as:

- N, number of turns: minimum, maximum values and incremental rate are to be determined as input data by designer
- a, ratio of limb depth to limb width: from 1 to 2 with an incremental rate of 0.1
- t, foil thickness : from 0.05 to 1.0 with an incremental rate of 0.1
- J, current density: from 0.1 to 5 with an incremental rate of 0.1

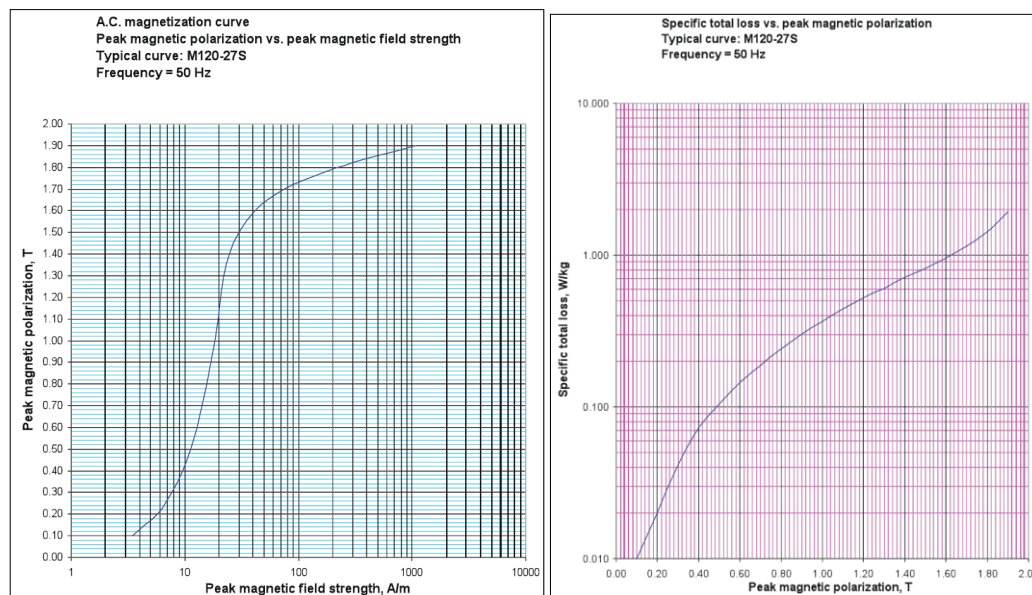


Fig.3. Magnetization and Loss Curves of M4 steel

The following design example can be considered for better understanding of the design process. A 250 kVAR 13800 V (808.66 mH, 31.38A) shunt reactor is to be designed. Input data and post-calculation results are as in Figure 2 and also summarized in Table I.

The core material for this design example is M4 steel and the winding material is Aluminum foil. Desired value of operating magnetic flux density is 1.1 Tesla, minimum and maximum temperature rise values are set at 78 and 82, respectively for 80 K temperature rise. Leakage inductance percentage is selected as 20 percent. Setting the initial number of turns (N) to 500 up to 800 with 20 incremental step, the software starts calculation with all data input given in Figure 2. After all iterations completed in almost a minute, the software will determine the optimum reactor satisfying the optimization criteria among thousands of reactors ($N \cdot a \cdot t \cdot J = 16 \cdot 10 \cdot 20 \cdot 50 = 160000$). At each iteration step firstly physical dimensions of the reactor, secondly core and winding losses, then Present Value Cost (PVC) and finally temperature rise are calculated. Since the design criteria is Min PVC, the reactor with minimum PVC among the reactors satisfying the desired temperature rise limit is determined as the optimum one and its calculated physical parameters are displayed on the screen

The results show that the optimum for defined assumptions; has 680 turns, 39200 mm² cross-sectional area (limb width 140 mm and depth 280 mm, ratio of depth to width is 2). The software rounds the value of limb width to multiples of ten to satisfy the realistic steel dimensions used in practice. Prices of iron and aluminum are taken to be 2.5 and 3.7 € per kg, respectively. It is worth to note that the proposed design software provides all the input data and design criteria to be modified according to the special design under consideration. The analytical results obtained from design software will be used as physical dimensions for modelling the reactor with Ansys/Maxwell software in determination of leakage inductance percentages of GISR.

B. Finite Element Analysis to Obtain Leakage Inductance Percentages

Firstly, physical dimensions and parameters of the reactor under design are obtained with the aid of the design software (Fig.2) and then the reactor is modelled in Maxwell 3D software to calculate inductance components by energy method as defined in [1, 12]. Table I represents target reactor specifications (design data), analytical calculation results obtained via the design software and FEA results for a set of iterations which are performed to obtain the leakage inductance percentage values for various power rating at 13.8 kV operating voltage, and at 80 K temperature rise. Each column in Table I represents an individual iteration, thus an individual reactor design phase. Since the estimated leakage inductance value changes at each new iteration, the physical parameters differs from each other.

Since the leakage inductance percentage is not known, design starts with an initial estimate. For the design of 250 kVAR 13.8 kV (808.66 mH, 31.38 A) GISR, initial estimate of leakage inductance at first is set to 5% as given in the first column of Table I. Then the analytical calculations with this value for the reactor are performed via the design software (Fig.2). After the physical parameters of the target reactor are obtained from analytical calculations, the reactor is then modelled in Ansys/Maxwell 3D software. Once FEA is performed, inductance parameters are obtained by the energy method. FEA results, shown at the end of the first column of Table I, show that the leakage inductance percentage and the inductance decline for this design iteration are 19% and 17%, respectively. However, two criteria

- leakage inductance percentage value close to the initial estimate
 - inductance decline value smaller or equal to 5%
- should be simultaneously satisfied for the iteration to be accurately completed;

TABLE I
DESIGN DATA / ANALYTICAL CALCULATION AND FEA RESULTS
FOR A SAMPLE SET OF ITERATIONS WITH 13.8 kV OPERATING
VOLTAGE AND 80 K TEMPERATURE RISE

Design Data								
Temperature Rise	80	80	80	80	80	80	80	80
# gap	40	40	40	40	40	40	40	40
kVAR	250	250	250	500	500	500	750	750
Voltage (l-l) kV	13,8	13,8	13,8	13,8	13,8	13,8	13,8	13,8
Current Amps	31,4	31,4	31,4	62,8	62,8	62,8	94,1	94,1
L (mH)	808	808	808	404	404	404	270	270
Analytical Calculation Results								
Lleak	0,05	0,1	0,2	0,05	0,1	0,15	0,1	0,15
Bm (T)	1,1	1,1	1,1	1,1	1,1	1,1	1,1	1,1
A (mm ²) (x10 ³)	45	39,2	39,2	57,8	57,8	57,8	72,2	68,6
A _{eff} (mm ²) (x10 ³)	44,1	38,4	38,4	56,6	56,6	56,6	70,8	67,2
foil thickness (mm)	0,15	0,15	0,15	0,25	0,25	0,25	0,3	0,3
J (A/mm ²)	1,4	1,3	1,4	1,1	1,2	1,2	1,1	1,1
height (mm)	729	717	705	865	837	842	971	968
N (turns)	730	730	680	550	490	500	420	410
L _{gap} (mm)	39,2	36,1	35,2	57,2	47,9	52,8	66,0	63,2
winding width (mm)	329	329	306	303	270	275	252	246
window (mm)	474	474	451	448	415	420	397	391
a (mm)	150	140	140	170	170	170	190	190
b (mm)	300	280	280	340	340	340	380	361
Temp Rise (K)	82	81,5	81,7	82	81,7	82	81,7	81,7
wind-wind clear. (mm)	25	25	25	25	25	25	25	25
yoke clearance (mm)	120	120	120	120	120	120	120	120
tube (mm)	60	60	60	60	60	60	60	60
PVC (euro)	2518	2421	2191	3912	3745	3570	4914	4720
Simulation Results								
L _{total} (mH)	944,8	892,7	789,8	470,4	434,3	418	291	273,5
L _{gap} (mH)	765,7	725	644,6	383,3	362,9	342,9	242,1	228,7
L _{iron} (mH)	1,69	1,82	1,55	0,62	0,68	0,57	0,34	0,33
L _{leakage} (mH)	177,5	165,9	143,6	86,5	70,7	74,6	48,5	44,3
Leakage %	18,8	18,6	18,2	18,4	16,3	17,8	16,7	16,2
Inductance decline %	16,85	10,39	2,34	16,33	7,40	3,38	7,96	1,42

Since the solution is not satisfying, a second iteration with the information given in the second column of Table I is performed for the same GISR. Now, initial leakage percentage is estimated to be 10% and defined as input data into the design software for analytical calculations. With the new set of physical parameters obtained for this case, the same procedure is repeated. Post-simulation results in the second column of Table I show that the inductance decline and the

leakage inductance percentage criteria after the second iteration have not been met yet. Finally, a new design iteration and FEA are performed with 20% leakage inductance percentage estimation. Criteria to finalize the design for this individual reactor are finally satisfied in this third iteration, as being 18.18% leakage inductance and 2.34% inductance decline as shown in the third column of Table I.

Table I shows only a few iteration to explain the procedure in determining the leakage inductance percentage values. The process is repeated for each set of operating voltages, temperature rise values and power rating. As a result, hundreds of design and simulation iteration are performed to obtain leakage inductance percentages as graphical curves given in Figure 5.

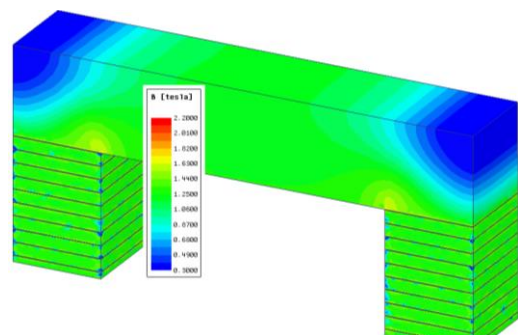
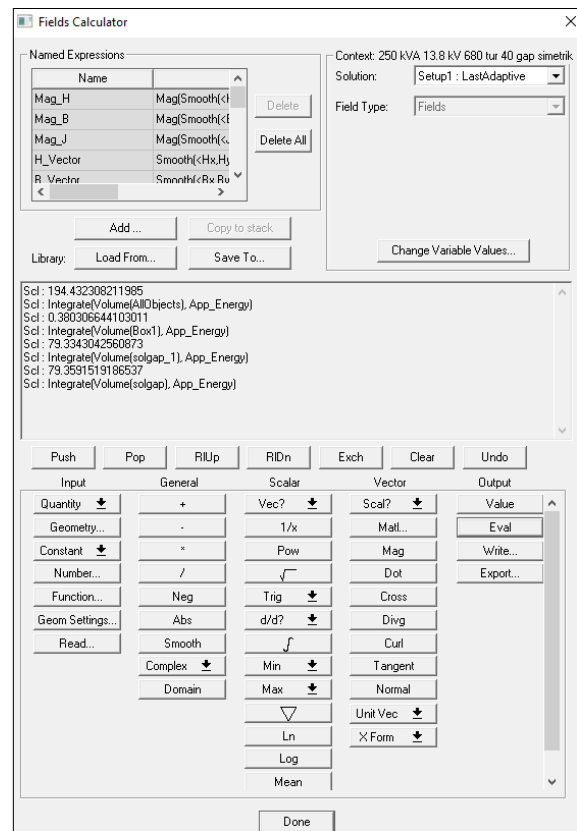


Fig.4. Calculation of energy in each volume and plot of magnetic field density for 250 kVAR 13.8 kV GISR with 20% leakage inductance and 80 K temperature rise

Calculation of inductance percentages in FEA are performed by energy method [1, 12]. The energy (co-energy may also be

used for linear operation which is the case in general) in each volume; occupied by iron, air-gaps, and the total volume surrounding the core is calculated by integrating the energy density after post-simulation via the fields calculator of the software as shown in Fig.4. After having the resultant energy value, the inductance of the related volume is obtained by (2). By this way, all the inductance components of the GISR are obtained. Calculation of energy in each volume and plot of magnetic field density for 250 kVAR 13.8 kV GISR with 20% leakage inductance and 80 K temperature rise are given in Figure 4.

$$W = \frac{1}{2} L I^2 \tag{2}$$

C. Leakage Inductance Percentages as Graphical Curves

Variations in leakage inductance percentage against reactive power in terms of operating voltage and temperature rise for M4 steel are graphically represented in Figure 5. In addition, the graphs for M33 steel are given in Figure 5 [12] for comparison.

Leakage inductance percentage of M4 steel at 0.4 kV operating voltage is almost same for all temperature rise values up to 0.1 MVAR and beyond this power range it slightly increases. However, for M33 steel, leakage inductance percentages are different for all temperature values and increases dramatically. For M4 steel, leakage inductance percentage is lower at 80 K and 100 K, however, it is higher at 60 K temperature rise.

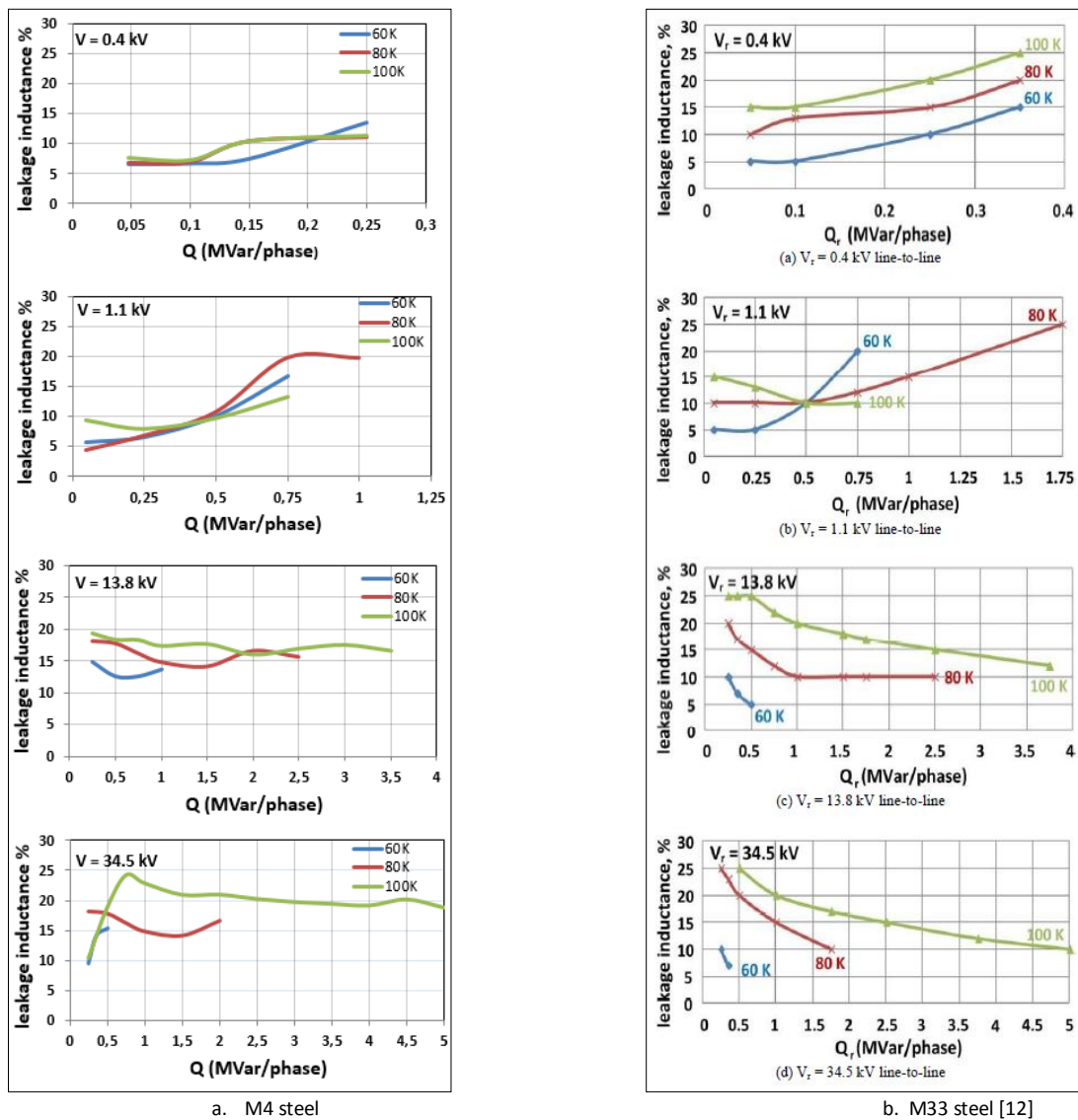


Fig.5. Leakage inductance % of M4 steel and M33 steel

At 1.1 kV, leakage inductance percentage increases as reactive power increases for all temperature rise values, beyond 250 kVAR/phase for M4 steel and 500 kVAR/phase for M33 steel.

At 500 kVAR/phase, leakage inductance percentage seems to be 10 both for M4 and M33 steel for all temperature rise values.

At 13.8 kV, leakage inductance percentage decreases slightly for all temperature rise values for both materials, whereas the decrease is more dramatical for M33 steel.

The curves at 34.5 kV show considerable difference in behavior of two materials. Although the leakage inductance percentage values of M4 steel are lower at 0.4, 1.1 and 13.8 kV operating voltage, they are getting higher at 34.5 kV level as reactive power increases.

Graphical results show that both steel has its specific leakage inductance percentage values. Any change in core material during the design phase therefore requires new set of graphical curves for reliable results.

III. CONCLUSION

Neglecting leakage inductance in the design phase of GISR results in an expensive and bulky device as shown by presented simulation results. Graphical curves representing leakage inductance percentages for M4 steel in addition to the existing curves for M33 steel, will offer variety to the literature in design of GISR. Results presented in this work will provide manufacturers and designers to implement a fast, accurate and economical design by taking the effect of leakage inductance into consideration. The proposed design software for analytical calculations is a valuable design tool and provides design of GISR in a wide range.

ACKNOWLEDGMENT

This work has been funded by the Scientific and Technological Research Council of Turkey (TUBITAK) Code 1002 Grant in the scope of project 118E687.

REFERENCES

- [1] Atilla Donuk, Mihai Rotaru, Jan K. Sykulski, "Defining and computing equivalent inductances of gapped iron core reactors." *Przegląd Elektrotechniczny (Electrical Review)*, vol.88.7b, 2012, pp.52-55.
- [2] Lee R., Stephens D., "Influence of core gap in design of current limiting transformers." *IEEE Transactions on Magnetics*, vol. 9.3, 1973, pp. 408-410.
- [3] J.P. Vora, H.C. Barnes, B.L. Johnson, "New shunt reactor principle proved-design data and factory test results for units built on insulated core principle." *IEEE Transactions on Power Apparatus and Systems*, Vol. PAS-92.3, 1973, pp. 900-906.
- [4] B. Tomczuk, K. Babczyk, "Calculation of self and mutual inductances and 3-D magnetic fields of chokes with air gaps in core." *Electrical Engineering*, vol. 83, 2001, pp. 41-46.
- [5] A. Bossi, G. Tontini, F. Coppadoro, "Influence of dimensional parameters on the design of gapped-core shunt reactors." *IEEE Transactions on Power Apparatus and Systems*, vol. 98.4, 1979, pp. 1144-1144.
- [6] A. Lotfi, M. Faridi, "Design optimization of gapped-core shunt reactors." *IEEE Transactions on Magnetics*, Vol. 48.4, 2012, pp. 1673-1676.
- [7] Y. Zhao, F. Chen, X. Ma, Z. Zhou, "Optimum design of dry-type air-gapped iron-core reactor based on dynamic programming and circular traversing algorithm." *International Conference on Electromagnetic Field Problems and Applications*, Dalian, Liaoning, China, 2012.

- [8] B. Tong, Y. Qingxin, Y. Rongge, Z. Lihua, Z. Changgeng, "Research on stress characteristics of shunt reactor considering magnetization and magnetostrictive anisotropy." *IEEE Transactions on Magnetics*, vol.54.3, 2018.
- [9] H.J.Kim, G.H. Lee, C.H. Jang, J.P. Lee, "Cost-effective design of an inverter output reactor in ASD applications." *IEEE Transactions on Industrial Electronics*, vol. 48.6, 2001, pp. 1128-1135.
- [10] W.A. Roshen, "Fringing field formulas and winding loss due to an air gap." *IEEE Transactions on Magnetics*, vol. 43.8, 2007, pp. 3387-3394.
- [11] H.D. Gersem, K. Hameyer, "A finite element model for foil winding simulation." *IEEE Transactions on Magnetics*, vol. 37.5, 2001, pp. 3427-3432.
- [12] Atilla Dönük, *Modeling and Design of Iron-Core Shunt Reactors with Discretely Distributed Air-Gaps*, Middle East Technical University, 2012.
- [13] A. Donuk, H.F. Bilgin, M. Ermis, "A practical approach to the design of power shunt-reactors with discretely distributed air-gaps." *International Review of Modelling and Simulations*, vol.6.2, 2013, pp. 567-576.

BIOGRAPHIES



ATILLA DÖNÜK was born in Malatya, Turkey, in 1977. He received BS degree in Electrical and Electronics Engineering from İnönü University in 2000, and PhD degree in Electrical and Electronics Engineering from Middle East Technical University in 2012 where he has been a

Research Assistant between September 2002 and February 2013. He worked as a guest researcher in the Research Group of Power Systems in Electrical and Computer Science at University of Southampton (UK) between November 2010 and October 2011. He has worked in Department of Electrical-Electronics Engineering at Atatürk University between February 2013 and July 2013. Since July 2013, he has been working as assistant professor in Department of Electrical and Electronics Engineering at Adnan Menderes University. His research interests include electrical machine design, power electronics and renewable energy applications.

A Hybrid Framework for Matching Printing Design Files to Product Photos

A. KAPLAN and E. AKAGUNDUZ

Abstract— We propose a real-time image matching framework, which is hybrid in the sense that it uses both hand-crafted features and deep features obtained from a well-tuned deep convolutional network. The matching problem, which we concentrate on, is specific to a certain application, that is, printing design to product photo matching. Printing designs are any kind of template image files, created using a design tool, thus are perfect image signals. For this purpose, we create an image set that includes printing design and corresponding product photo pairs with collaboration of an actual printing facility. Using this image set, we benchmark various hand-crafted (SIFT, SURF, GIST, HoG) and deep features for matching performance. Various segmentation algorithms including deep learning based segmentation methods are applied to select feature regions. Results show that SIFT features selected from deep segmented regions achieves up to 96% product photo to design file matching success in our dataset. We propose a framework in which deep learning is utilized with highest contribution, but without disabling real-time operation using an ordinary desktop computer.

Index Terms— image matching, hand-crafted features, deep features, semantic segmentation, product image processing

I. INTRODUCTION

IMAGE MATCHING is a broad title that covers or partially relates to various topics among a number of different computer vision problems, namely image-based localization, multi-view 3D reconstruction, structure-from-motion, image retrieval, tracking, just to name a few. This title may refer to finding a transformed version of an image [1], or may refer to a different version of the problem, such as finding an image with a similar semantic context [2]. Regardless of the problem definition, image matching boils down to a simple statement: finding a similarity model between (at least) two images, which would satisfy the pairings for a given image set.

The algorithms proposed under this title in recent years can

be mainly split into two principle categories. The first category consists of approaches that utilize hand-crafted representations. Among these methods, the bag-of-visual-words (BoVW) algorithm [3] proved to be very successful, irrespective of the type of the hand-crafted feature used, and is still the state-of-the-art approach due to its flexibility, compactness and speed.


However, as a part of the growing wave of interest on deep-learning-based methods, a second category of approaches recently focus on image matching using convolutional neural networks (CNN) [4]. The strength of these methods comes from the abstract features that merge at the deeper layers of CNNs [5]. The earlier approaches of this category [6-9] performs particularly good at problems like image category classification, object detection and/or localization, mainly because of their capability to convolve abstract features into image categories or object definitions. There are also attempts with promising results, which aim at transferring pre-trained and well-tuned CNNs into image retrieval frameworks [10, 11]. Nonetheless, these network structures are not well-suited to match a given image to its pair, since they use fully connected layers that lead to a classification layer (such as soft-max). This final layer is used to classify the extracted abstract features into object categories. Therefore, their structure is not designed with the purpose of finding pairs.

Very recently a new CNN structure, namely the Siamese network (SN), has been proposed specifically for the problem of image matching [12]. SNs can learn feature spaces that map similar image pairs close to each other and dissimilar image pairs with a selective distance, by using labelled pairs. This approach has also been successfully applied to similar problems that require an image-to-image matching, such as face recognition [13] or aerial-to-ground image matching [14]. SNs improve matching performance dramatically, however they come up with two main drawbacks. Firstly, they necessitate the creation of a large-scale image set, because in order to span the entire space of possible transformations from the original image to the image to be matched, a massive number of image pairs are required. Such an image set collection and annotation effort is extremely expensive and usually, industrially impracticable. Secondly, these networks make a separate full forward deep CNN run for each candidate image in the image set, thus are slow even with dedicated hardware, such as a GPU.

Deep learning is a powerful tool. In less than a decade, nearly all vision problems shifted to CNN domain.

ALPER KAPLAN, is with Cognitive Science Program, Graduate School of Social Sciences, Yeditepe University, Istanbul, Turkey (e-mail: alperkaplan@outlook.com).

ERDEM AKAGUNDUZ, is with Department of Electrical and Electronics Engineering, Çankaya University, Ankara, Turkey, (e-mail: akagunduz@cankaya.edu.tr).

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

Manuscript received January 20, 2020; accepted April 16, 2020.
DOI: [10.17694/bajece.677326](https://doi.org/10.17694/bajece.677326)

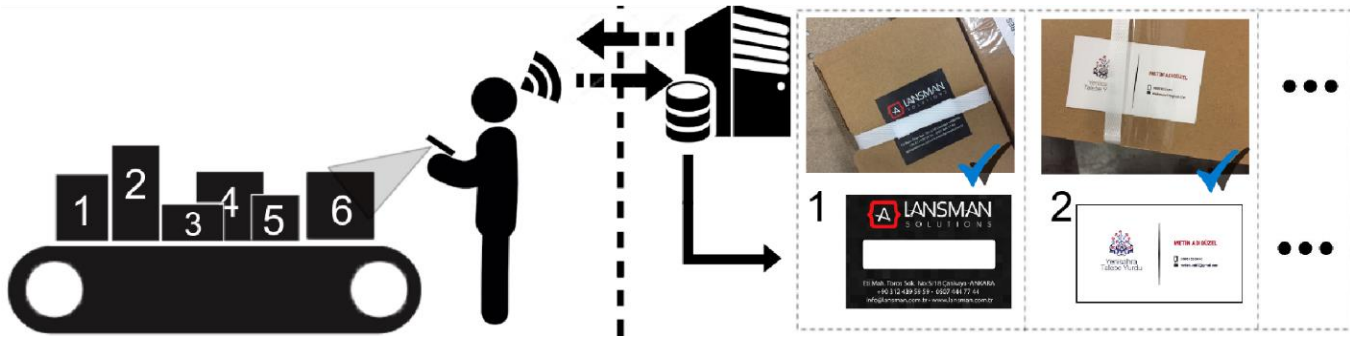


Fig.1. Pictorial representation of our operational problem definition

Nevertheless, it still comes with a price. As the layers of a CNN get deeper, the hardware requirements for real-time operation become more and more expensive. We still don't have a mobile solution, which may replace the high-cost and power-hungry GPUs that allow real-time deep learning operations. And when it comes to the problem of image matching, our best solution yet, namely the Siamese networks architecture, require massive training sets and forward-run for all possible candidate images. In conclusion we still need ingenious solutions for real-time, operation-specific and affordable image matching frameworks.

A. Problem Definition and the Proposed Solution

In this paper, we study a particular version of the image matching problem, in which we match printing design files to product photos. Printing designs are any kind of template image files, created using a design tool and used as templates for printing a flyer, banner, poster, etc. These files are computer generated, thus they possess no signal-based deficiencies like noise or optic blur. Most of them are in vector format, hence, are resolution-free.

On the other hand, photographs of a printed product suffer many unwanted effects, such as uncontrolled shooting angle, uncontrolled illumination, occlusions, printing deficiencies in colour, camera noise, optic blur, etc. Matching them to their original design files requires learning the unknown transformation that the photographing action creates. This transformation is not deterministic by nature and is affected by predominant uncontrolled factors, such as the photographer, the camera or the background.

A pictorial representation of our problem definition and system framework is depicted in Figure 1. In a sample scenario of our problem definition, an operator (or an automatic visualization system) shoots the photographs of some printed products by using a computation-limited device (such as a mobile phone, etc.). Then this device sends the product photo to a server machine, in which the photo is matched to its design file pair, in real-time.

In order to solve this problem, we propose a real-time image matching framework, which is hybrid in the sense that it uses both hand-crafted features and deep features obtained from a

well-tuned very deep CNN [6]. The hand-crafted or deep features are extracted only from a region that is designated by a fine-tuned deep CNN. In our framework, this feature region segmentation operation is the only "deep" operation that is applied on the product photo, thus we avoid running a deep CNN for each possible pair in the image set, as it is done for Siamese networks. This also prevents us from using expensive deep learning hardware (a GPU), but still permits us to provide real-time operation. By using the hand-crafted or deep features extracted from the deep segmented region, a BoW framework is utilized in order to find the correct image pair.

In following section, we provide the details of the image set that includes printing design files and corresponding product photo pairs. Section 3 explains the deep learning experiments, which aim at solving feature region segmentation problem. Section 4 represents the BoVW framework, in which different hand-crafted and/or deep feature extraction, and product segmentation methods are benchmarked for optimal performance. Section 5 presents the experimental results, whereas the final section concludes the paper and gives directions for future work.

II. DESIGN FILE - PRODUCT PHOTO PAIRS IMAGE SET

The existing image sets [15-18] prepared for matching or retrieval problems in the literature are very diverse in category. They deal with different problems such as retrieving RAW images, medical images, outdoor images, or even satellite images. Consequently, for each image set the problem definition is different. That's why, in order to provide a solution for our specific problem definition, we need to create a specific image set that includes printing design files and corresponding product photos.

As a consequence, an image set creation effort was carried out. To this end, an operator took the photographs of 2000 products at the production line. Product photos are images of a sample product (e.g. a flyer) usually stitched over a cargo box, which carries the other printed samples (Figure 2). The idea is to recognize this product (i.e. its ID) by matching the image of the sample on the cargo box with the design file at the server.

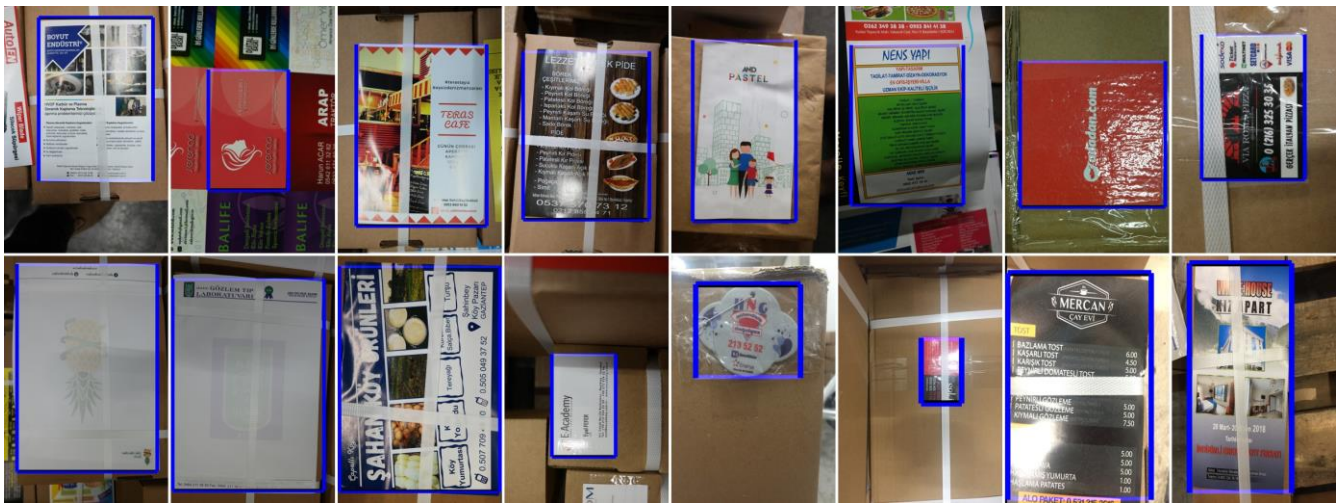


Fig.2. Sample product photos from the image set

For some of the products, there exists more than a single design file. A good example is a business card (Figure 1), which usually has information on both sides, and thus has two separate design files. In these cases, the problem definition is to match the product photo (which could be any face of the card) to one of the design files in the image set. Consequently, for the photographed 2000 product samples, 3458 design files were added to the image set. The operators were advised to shoot the product with a perpendicular angle so that the product (usually, but not necessarily rectangular in shape) would fit the image with uniform margins. However, this weak protocol was not successfully applied to all images, mainly because of human-errors, and it is difficult to say that the image set is rotation or scale controlled (please see Figures 1 and 2).

In addition to the product shooting and design file labelling efforts, an annotation effort was also carried out. For each photographed product the rectangle that encapsulates the product sample was annotated on the images (depicted as blue rectangles on Figure 2). These annotations will later be used as ground truth to our deep learning framework in the following section.

In Figure 2, several examples from the image set are provided. As it can be seen from this figure, the set includes various types of background clutter, occlusions caused by packaging (tapes, chords, etc.), unwanted flash light reflections, non-uniform illumination, folding of the sample product and such. Thus, it is important that the matching solution we propose, must be robust to these types of effects.

III. DEEP PRODUCT SEGMENTATION

As mentioned in the introduction section, we apply a deep-segmentation supported bag-of-visual-words method to match product photos to design files. This framework, with rigorous benchmarking, is provided in the next section. However, before we get into the details of our matching framework, in this chapter we present some methods for segmenting the

product region in product photos using different deep CNN (DCNN) architectures.

Finding the pixels that belong to a specific object category is known as semantic segmentation in the literature. The reader may refer to various surveys on this problem [19-26]. The literature involves hundreds of different approaches to semantic segmentation. The most common component among these approaches is undoubtedly the utilization of the abstract features of pre-trained DCNNs, by fine-tuning or transfer learning.

In this paper, in order to segment the pixels of a product photo using deep learning, we adapt three different architectures: “FCN32s”, “FCN8s” and finally the proposed “VGG-Regression-Net”, as we name it.

TABLE I
VGG-REGRESSION-NET ARCHITECTURE

Input Layer	224×224×3 RGB Image. (VGG default input image size)	
VGG Layers	Pre-trained VGG layers (1 to N)	
New Layers	(all fully connected)	
layer no.	Number of Weights	Activation Vector Size
N+1	$j \times k \times d \times 256$	$1 \times 1 \times 256$
N+2	$1 \times 1 \times 256 \times 256$	$1 \times 1 \times 256$
N+3	$1 \times 1 \times 256 \times 900$	$1 \times 1 \times 900$
Output	900x1 vector (30x30 Segmentation Result)	

FCN32s and FCN8s are well-known fully convolutional semantic segmentation networks, designed specifically for this problem [27]. They are originally trained for 21 different pixel labels. The only difference between these two architectures is that FCN8s includes skip connections that allow feature concatenation between different hierarchies within the DCNN. In order to adapt these networks to our problem, the final deconvolutional layers are set to 2 labels depth (as “product” and “background”), and while this layer is being learned from scratch, all other convolutional layers in the networks are fine-tuned during training.



Fig.3. Segmentation results for VGG-Regression-Net Layer 13, namely Conv5¹ are depicted. For four different samples, the image (left), the annotated ground truth (middle) and the DCNN output (right) are shown.

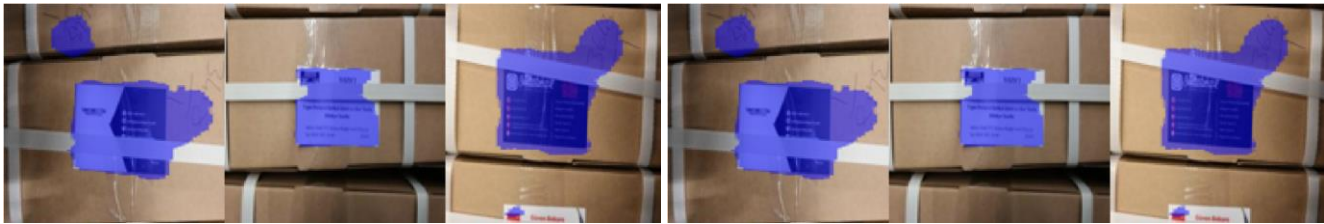


Fig.4. Segmentation results, as blue regions over the image, for FCN8s (leftmost three images) and FCN32s (rightmost three images) are depicted. FCNs may create segmentation regions with disconnected blobs, since their fully convolutional nature has no means to prevent such an output

A. VGG-Regression-Net Architecture

In addition to these fully convolutional architectures, a regression network is also proposed for the same problem. To that end, the convolution layers of a pre-trained deep network, namely VGG-VD-19L [6] are transferred to the VGG-Regression-Net (VGGRN) architecture. In VGGRN, fully connected layers are replaced and retained. The aim is to assess whether a DCNN with fully connected layers, this time trained for regression of the segmentation mask, performs better than fully convolutional architectures, such as FCN32s or FCN8s. We hypothesize that the fully connected layers can provide inference for a global composition of the image, in which the FCNs could fail to achieve.

TABLE II
VGG-REGRESSION-NET EXPERIMENTS

Layer	NCC mean	NCC std.
VGG-Regression-Net Layer 09: Conv4 ¹	0.84	±0.11
VGG-Regression-Net Layer 10: Conv4 ²	0.85	±0.12
VGG-Regression-Net Layer 11: Conv4 ³	0.87	±0.15
VGG-Regression-Net Layer 12: Conv4 ⁴	0.87	±0.13
VGG-Regression-Net Layer 13: Conv5 ¹	0.92	±0.06
VGG-Regression-Net Layer 14: Conv5 ²	0.91	±0.09
VGG-Regression-Net Layer 15: Conv5 ³	0.90	±0.14
VGG-Regression-Net Layer 16: Conv5 ⁴	0.89	±0.12
VGG-Regression-Net Layer 16: Conv5 ⁴	0.89	±0.12
FCN32s	0.87	±0.10
FCN8s	0.85	±0.09

The detailed architecture of the VGG-Regression-Net is provided in Table 1. Selected N number of pre-trained convolutional layers of VGG-VD-19L are transferred to this new architecture. The input layer of the original VGG-VD-19L receives 244×224 pixels RGB images. So any product

photo that is fed to this DCNN is first down-sampled into 244×224 pixels resolution.

In order to find the most suitable layer, multiple transfer learning experiments are run. Each separate experiment corresponds to creating a new DCNN by transferring “some” VGG-VD-19L layers and replacing new fully connected decision layers, so that we create the segmentation mask for the product in the product image. The weights of the transferred layers are frozen during training.

The ground truth of these masks are obtained by using the annotation mentioned in the previous section (please see the blue rectangles in Figure 2). For each product photo, a ground truth mask for which the pixels inside the rectangle region are 1 and the rest (i.e. the background) 0, is created and used for training. The output of the final fully-connected layer consists of 900 components, which is actually the 30×30 pixels-sized, down-sampled version of the segmentation mask.

B. Training the Deep Segmentation Architectures

Although training the FCNs is a subject of semantic segmentation, training the proposed VGG-Regression-Net architecture is a regression problem. In order to train this DCNN, L1-norm operator is implemented as a loss function. Stochastic gradient descent (SGD) with momentum is utilized and a batch size of 16 images¹ is used for batch normalization.

¹ Stochastic Gradient Descent (SGD) algorithm with momentum is employed, Initial Learning rate: 0.001, Weight Decay: 0.0004, Momentum: 0.91. MatConvNet [28] library is used for training the VGG-Regression-Net, while MATLAB Deep Learning Toolbox is utilized for training the FCNs.

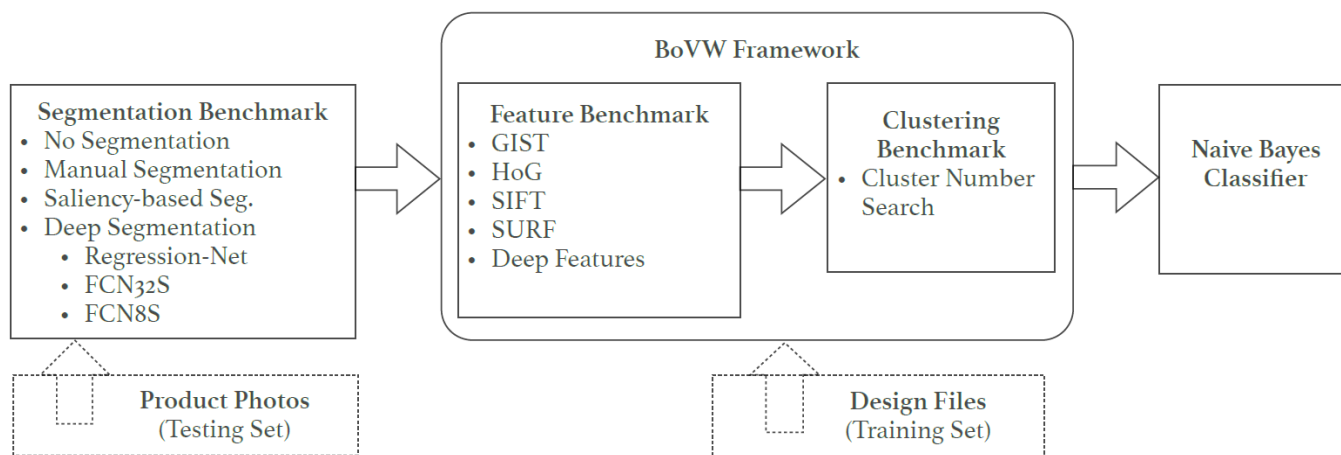


Fig.5. The overall matching framework is depicted. The framework consists of four main blocks, namely segmentation, feature extraction, clustering and classification.

For all architectures, data augmentation is applied by mirroring ($\times 2$), zooming ($\times 2$) and rotating ($\times 4$) the product photos, thus enlarging the image set by 16. For training the architectures 75% and for validation 15% of the image set are used. For this reason, each experiment is run 10 times, using a different subset for testing, which contains a separate 10% of the whole image set.

Regarding the VGG-Regression-Net architecture, in order to find the layer that provides the best abstract features for segmentation, 8 different structures are trained, by cutting the VGG-VD-19L at 8 different layers, namely conv41, conv42, conv43, conv44, conv51, conv52, conv53 and conv54, which are the 9th to 16th layers of VGG-VD-19L. Thus for 8 layers of VGG-Regression-Net, FCN32s and FCN8s, separately for 10 test sets, a total of 100 learning experiments are run.

In order to benchmark the success of different VGG-Regression-Nets and the FCNs architectures, normalized-cross correlation (NCC) of the test results with the ground truth are calculated and averaged over the entire set, respectively for each experiment. In Table 2, for each experiment the mean and the standard deviation of the segmentation accuracy (i.e. normalized-cross correlation of test results with the ground truth) are calculated. The best results are obtained using the 13th layer of VGG-VD-19L for fine-tuning experiments, with an average of 0.92 normalized cross-correlation. FCNs both perform poor.

We believe that this is mainly because of the fact that, the problem we solve here is not exactly semantic segmentation. The product photo to be segmented, a poster, a flyer et cetera is a composition of objects, not a single object. Our problem is about learning the pixels of a product inside an image, as a composition of objects. We believe that the reason why the VGG-Regression-Net architecture performed better compared to FCNs is mainly because, fully connected layers can learn the global composition of abstract features, whereas FCNs, with limited receptive fields, search for objects in local regions. Therefore, as seen in Figure 4, even disconnected blobs as segmentation results can be obtained for FCNs.

In the rest of this paper, we examine the effect of these three different segmentation methods to our matching

performance. For this purpose, each matching method is deep segmented by the two adapted FCNs and the best VGG-Regression-Net experiment, which is obtained by using the abstract features from layer 13 (namely conv5¹).

IV. IMAGE MATCHING FRAMEWORK

As previously mentioned in the introductory sections, the aim of this study is to find a real-time solution to product photo and design file matching problem. For this purpose, we propose a framework with various benchmarking experiments. The proposed framework is hybrid in the sense that it fuses a conventional pattern recognition method that uses hand-crafted features with a deep segmentation technique. And while doing this, the study presents benchmarking of different methods, in order to correctly acknowledge the best performance for the given framework.

In Figure 5, the high level depiction of our framework can be seen with the benchmarking processes we utilize. The matching is accomplished by using the BoVW method [3] together with a Naïve Bayes classifier. The main advantage of using BoVW is that it provides a fixed-length representation of the image, regardless of the number or type of features obtained from that image. Moreover, BoVW + Naïve Bayes online operation (testing) is extremely fast. It requires a relatively slower offline training phase, in which the features obtained from training set is clustered into N sets. But needless to say, this does not affect the real-time operation in our framework. Within the BoVW framework, two principle benchmarking efforts are carried out, first being the benchmarking for selection of the hand-crafted or deep features of BoVW and second being the benchmarking for the optimal number of clusters for BoVW.

BoVW relies on the features obtained from a test image, which is a product photo in our case. The product photo does not only include the “product” but considerable background as well; ergo, it may be crucial to select the features only from the product region in the photo. For this purpose, as seen in Figure 5, another benchmarking effort for finding the product region is also carried out, using different segmentation

methods including the deep segmentation techniques explained in the previous section.

In the following subsections, we explain each benchmarking effort separately, following their process order in the framework. Thus, we start with the segmentation benchmark. Then we delve into our results and carry out discussions on the optimum method.

A. Segmentation Benchmark

As mentioned above, selecting the features only from the product regions may dramatically affect the matching performance of a BoVW model. For this reason, in our experiments we utilize five product segmentation methods out of three categories and compare their results within the complete framework. We also include the case where no segmentation is carried out, so that we can clearly assess the contribution of a tested segmentation method.

We categorize our tested segmentation strategies in three titles, namely manual segmentation, unsupervised segmentation and supervised segmentation.

1) Manual Segmentation

Manual segmentation is accomplished by the operator. As seen in Figure 1, the operator shoots the products with a mobile device and at this very moment, can manually segment the product region in the photo using the same mobile device. This is an unwanted scenario because it increases the operation time, which contradicts with the general purpose of the proposed system. However, we still choose to utilize this segmentation method in our framework, so as to see the effect of “perfectly” segmenting the product in a photo to our overall matching success and we regard this method as a ground truth for the segmentation step. In our image set we already have these manual annotations (please see Section 2 and Figure 2).

2) Unsupervised Segmentation: Visual Saliency

The first automatic segmentation method we utilize is an unsupervised method to find the product region in a photo. Unsupervised segmentation had been a very hot topic [29] before deep learning overwhelmingly manipulated the field with the idea of employing large-scale data to any problem. Since, for the sake of cheap and fast operation, we try to avoid deep operation as much as we can, we select a visual saliency-based method as our unsupervised segmentation method.

For this purpose, we use Graph-based Visual Saliency (GBVS) method, which is a bottom-up visual saliency model. By creating Markov chains among image pixels, the GBVS algorithm calculates saliency values from equilibrium distributions over pixel map locations [30]. Although it is not a direct segmentation technique, visual saliency is being used as an objectness measure and is utilized for segmenting objects in an image [31]. In Figure 6, a sample result on a product photo can be seen. Firstly, the saliency heat map is calculated. Then by using a constant threshold (0.11 in our experiments), the object region is segmented. In the same

figure, the segmented object can be seen in the rightmost image.

There are various methods to segment an object from an image without any prior information, in other words in an unsupervised manner. The reason we choose to use GBVS is simply because of its speed-accuracy trade-off [31]. In an extended study, it is possible to search for the most optimum method to segment an object in an unsupervised manner, however we find this effort beyond the scope of this study.

3) Supervised Segmentation: Deep Learning

Supervision is simply utilizing domain-specific data. Thus, compared to any unsupervised method, it is more susceptible to over-fitting. However, if there is sufficient training data, supervised methods are preferable most of the time. In order to segment the product in a supervised manner, we utilize the three deep segmentation methods, as explained in Section 3.

B. Image Features Benchmark

The general idea of BoVW is very simple: “representing an image as a fixed-length set of features”. The so-called features consist of keypoints and descriptors. Keypoints denote the salient locations in the image, ideally invariant to transformations. Descriptor is the description “around” the keypoint. BoVW use both keypoints and descriptors to construct vocabularies and represent each image as a frequency histogram of features that are in the image. Similarity measures to a test image can be calculated using these frequency histograms, and thus a classification can be performed.

For a BoVW framework, the most important question is obviously “which feature/descriptor to use”. Depending on the problem definition, imaging modality, performance requirements and computational budget, different methods can be used. For a comparison of local feature detectors and descriptors for visual object categorization, the reader may refer to [32]. In this study we employ 5 popular features, namely, GIST [33], histogram of gradients (HoG) [34], SIFT [35], SURF [36], and deep features, which are obtained using a special CNN layer, namely the Spatial Pyramid Pooling (SPP) Layer [37].

Among the aforementioned five features types, SIFT [35] and SURF [36] are, by definition, local; thus they are well-suited for BoVW. For HoG [34], the locality should be pre-defined, i.e. provided by the user for local a region with a fixed area. We use blocks of 16x16 on a uniform grid and obtain HoG features individually from each block.

GIST [33], on the other hand, is a global descriptor, more than a feature. It literally catches a “gist” of the scene by using multi-scale low level features. Hence it is incompatible for a BoVW model, and accordingly it is implemented out of the BoVW framework. We calculate the GIST descriptors for each training and test data. In consequence, by calculating the Euclidean distances between the GIST descriptors, we match a product photo to an image design file.

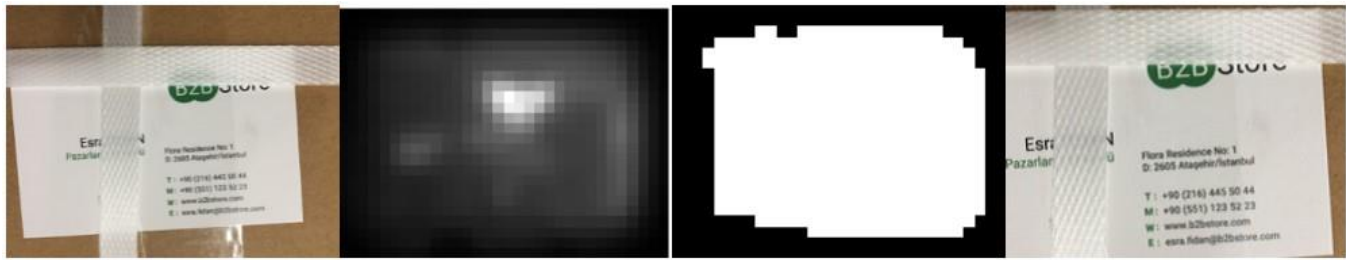


Fig.6. The objects in the product photos are segmented using the GBVS algorithm [30], which is selected as the unsupervised segmentation method for our segmentation benchmark.

Similarly, to GIST descriptor, the output of an SPP layer [37] does not need a histogramisation effort. This layer's output is already fixed-length. SPP collects activations from layers of different hierarchies, and concatenates them in a single fixed-length vector. For this purpose, we have utilized combinations of activations from different layers of the FCN32s network as an input to the SPP layer. By calculating the (weighted²) Euclidean distances between these vectors, matching is performed.

C. Hyper-Parameter Optimization

The proposed framework includes different methods with benchmarking of various intermediate steps (segmentation, feature extraction etc.). Thus there are many hyper-parameters that may affect the system performance. In this study, we optimize only the cluster number of the BoVW framework. We believe that this is the most important hyper-parameter, mainly because it is independent of the utilized segmentation, feature extraction or the classification steps. The number of clusters is the vocabulary size and thus the heart of a BoVW model. Accordingly, in the next section, we also provide results for different cluster numbers, thus showing the effect of vocabulary size on performance.

D. Classifier

The final block of the proposed framework is classification. BoVW provides a fixed length histogram representation for any image, and classification within this vector space is another step, which serves as the final decision of the system. In our framework, the final goal is to find the design file that matches the given product photo. Various classification methods can be employed for a BoVW system and the reader may refer to [38] for a detailed comparison.

BoVW concept is an adaptation of the bag of words (BoW) idea from natural language processing. In BoW, the so-called words are calculated by clustering the entire training set features into K number of subsets. Then, the number of each "word" in a document is counted, and a frequency histogram is created by using the word occurrences. We have the same concept in BoVW, but instead of words, we use image features. Image features can be anything, like salient regions in an image. We normalize frequency histograms to obtain probability distribution functions that represent the possibility of having a given visual word in an image. At this point we utilize Naïve Bayes algorithm. Occurrence of a word is

² In an SPP layer [37], the activations are weighted by the size of the pooling layer.

assumed as an independent event (which is the Naïve part) and all feature probabilities are multiplied to find the matching probability of an image to another.

We have chosen Naïve Bayes algorithm as our classifier mainly because of two reasons. Firstly, it is fast and compatible with real-time processing. And secondly training requires a small amount of samples to estimate the model parameters. This is why it has always been an optimal [39] partner for BoVW. We believe that with another, maybe mathematically more complex classifier, our results may further be improved. For the sake of computation speed we employ Naïve Bayes for all BoVW experiments in this study, and leave the benchmarking of different classifiers to a future study.

V. EXPERIMENTAL RESULTS

Before we present our comparative results with rigorous discussions in this section, the details of the experimental parameters are provided below.

A. Experimental Setup

The main objective of our experiments is to find an optimal method to product photo and design file matching. The absolute value of matching success depends on the number of design files to be compared in the image set. If there are only, for example, 10 design files to match, regardless of the benchmarked methods, the success would be relatively higher compared to a case in which thousands of possible design file candidates exist.

In our experiments, we have selected the number of product photos to match as 60³. Moreover, we have selected the number of design files as 100, so that the 60 product photos will match some of the design files in this 100 element set, whereas the rest are just fillers⁴.

³ This parameter can be optimized with further experimentation. However, this would require running thousands of experiments with 26 different methods, which we have chosen leave to a future study.

⁴ As explained in Section 2, the number design files that match a set of 60 product photos vary. A single page flyer has a single corresponding design file, whereas a two-sided business card has two design files matches for both sides. Thus the exact number of fillers (i.e. randomly selected non-pair design files) in a set of 100 design files changes according to the set of product photos.

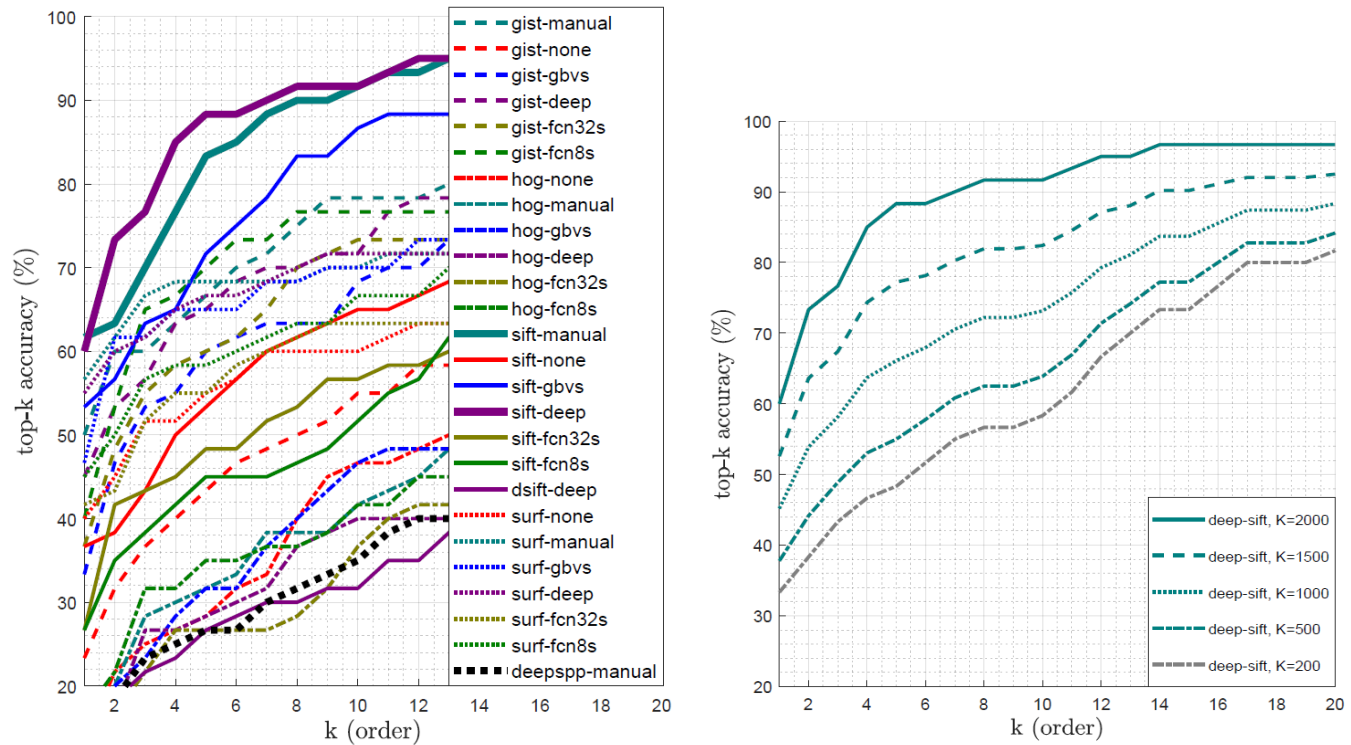


Fig.7. Top-k accuracy curves are depicted. In these curve x-axis denotes the order of match, i.e. the order of similarity of the actual pair in the training set. The y-axis denotes the average success for that order value. a) (left) Top-k accuracy curves for all method. b) (right) Top-k accuracy curves for *sift-deep* under varying number of clusters in BoVW.

Accordingly, we have created 1000 different randomly selected 60 product photo - 100 design file sets, using the total 2000 product photos and 3458 design files. For each benchmarked method, 1000 experiments are run using these 1000 different test cases. At each experiment, for each product photo, the order of match is recorded. For this purpose, when a product photo is matched to 100 design files for an experiment, the Naïve Bayes probabilities (or the Euclidean distances) are calculated and sorted. The rank of the probability of the corresponding design file is recorded as the order of that product photo's matching performance⁵.

B. Results

For benchmarking, we employ 26 different methods using a combination of 6 different segmentation methods and 5 different features as presented in Section 4. For instance, the case in which SIFT features, that are obtained from only the product region segmented using VGG-Regression-Net, is referred to as "sift-deep". Or, the case, in which GIST global features are obtained from a manual segmented object region in a product photo, is called "gist-manual".

All 26 methods are tested for the same 1000 experiment sets and for each product photo in each experiment, the order of match is recorded. Using this order measure, we also calculate the average order of being matched. For example,

⁵ For example, for a single product photo, we check the similarities to the given 100 design files in that experiment. The actual design file that matches the given product photo has the order 4, when all Naïve Bayes probabilities (or the Euclidean distances) are sorted. Then the order for this product photo in this experiment is simply 4.

for a specific method, the percentage of product photos that have smaller or equal order "k" is recorded as the top-k accuracy. For example, if the top-k accuracy for order k=10 is 95%, this shows that by only checking best 10 possible match results, it is possible to match the correct design file with 0.95 probability.

In Figure 7.a, the top-k accuracy for all methods are depicted. The numerical values are provided in Table III. The best accuracy is obtained with "sift-deep" method, for which SIFT features that are obtained from only the product region segmented via the VGG-Regression-Net, are fed to the BoVW framework. The sift-deep method also slightly outperforms sift-manual method, for which the segmentation is performed by the human operator. This indicates that human operators can make mistakes in product region annotation, whereas deep learning-based segmentation method can generalize these errors and perform much better.

After observing the success of SIFT, we have applied another version of the SIFT descriptor, namely the "Dense SIFT" (DSIFT) [40]. DSIFT is the same algorithm as the SIFT but it is run on a denser grid of locations. That is why DSIFT provide, on average, 10 times higher number of keypoints, compared to SIFT. In Figure 7.a, the performance of DSIFT, which is quite poor, can also be seen. This is, we believe, because of the fact that, increasing the number of keypoints does not support representation, but conversely creates more false alarm matches between feature clusters.



Fig.8. Sample results for the sift-deep method are seen. The samples on the top row are found with a perfect hit. The samples in the bottom row are matched with orders 8, 70 and 11 from left to right, respectively.

TABLE III
NUMERICAL RESULTS FOR EACH METHOD

Method / Order	1st	5th	10th	15th	20th
gist-manual	50.0%	66.7%	78.3%	81.7%	86.7%
gist-none	23.3%	43.3%	55.0%	63.3%	65.0%
gist-gbvs	33.3%	60.0%	68.3%	73.3%	75.0%
gist-deep	45.0%	65.0%	71.7%	80.0%	88.3%
gist-fcn32s	36.7%	60.0%	73.3%	76.7%	80.0%
gist-fcn8s	40.0%	70.0%	76.7%	76.7%	78.3%
hog-none	13.3%	28.3%	46.7%	50.0%	55.0%
hog-manual	15.0%	31.7%	41.7%	51.7%	58.3%
hog-gbvs	16.7%	31.7%	46.7%	50.0%	58.3%
hog-deep	11.7%	28.3%	40.0%	45.0%	55.0%
hog-fcn32s	8.3%	26.7%	36.7%	45.0%	53.3%
hog-fcn8s	16.7%	35.0%	41.7%	46.7%	50.0%
sift-manual	61.7%	83.3%	91.7%	95.0%	95.0%
sift-none	36.7%	53.3%	65.0%	76.7%	83.3%
sift-gbvs	53.3%	71.7%	86.7%	88.3%	90.0%
sift-deep	60.0%	88.3%	91.7%	96.7%	96.7%
sift-fcn32s	26.7%	48.3%	56.7%	65.0%	80.0%
sift-fcn8s	26.7%	45.0%	51.7%	66.7%	80.0%
dsift-deep	8.3%	26.7%	31.7%	45.0%	45.0%
surf-none	40.0%	55.0%	60.0%	66.7%	70.0%
surf-manual	56.7%	68.3%	70.0%	71.7%	71.7%
surf-gbvs	46.7%	65.0%	70.0%	76.7%	76.7%
surf-deep	55.0%	66.7%	71.7%	73.3%	76.7%
surf-fcn32s	41.7%	55.0%	63.3%	63.3%	63.3%
surf-fcn8s	45.0%	58.3%	66.7%	70.0%	73.3%
deepspp-manual	18.3%	26.7%	35.0%	40.0%	43.3%

The “deepspp” method, in which deep CNN features are fed to a SPP layer, performs poor, even with manual (perfect) segmentation of the product. Deep features carry abstract information, which may fill the semantic gap of any vision problem. Consequently, this poor performance of deep features was intriguing for us. For this reason, we have carried out extensive deep visualization experiments to uncover this issue. SPP creates activations from all (thousands even when only the deepest layer is used) neurons from the selected

layers, most of which are unfortunately noise and are usually dropped out within the deep CNN. An SPP layer does not have the ability to select deep features according to their quality. That is why this method performs unsurprisingly poor within a BoVW framework, compared to a more selective and scale-invariant descriptor method, such as the SIFT. We have utilized different combinations of activations from various layers of the FCN32s network. The best performance was obtained when only the activations from the final max-pooled convolutional layer (pool5 - $13 \times 13 \times 512$) was utilized. Only the resulting curve for this case is depicted in Figure 7.a. Our visualisation experiments clearly show that, regardless of the layer the deep features are obtained, selective activations are always overwhelmingly outnumbered by noisy, unselective and insignificant activations, which cannot lead to any semantic decision.

Consistent with our observations presented in Section 3, segmenting with FCN32s and FCN8s does not contribute the BoVW matching success positively. Semantic segmentation of product photos as if they are plain objects, is apparently not helping the feature selection operation enough.

In Figure 7.b, a parameter optimization effort for cluster numbers is depicted. As the number of clusters in BoVW increases, so as the success rates. In our tests, we tested up to 2000 clusters, which is the best case. This number can further be increased for higher success with a price of dramatically increasing our training time. The top-k accuracy for all methods that utilize BoVW in Figure 7.a are calculated using 2000 clusters, which is our optimal case.

In Figure 8, some sample results for the sift-deep method can be seen. The three samples on the top row are found with order 1, i.e. with a perfect hit. The samples in the bottom row are matched with orders 8, 70 and 11 from left to right, respectively. The mismatch cases are usually because of strong clutter or impaired design files.

The actual implementation of the system shows that the average “end-to-end” matching time for a product photo, using the sift-deep method in a regular, no-GPU desktop computer is less than 4 seconds, including communication delays⁶. This is a feasible duration for the operation considering that it is much faster than the operator manually searching for the product id, which takes about a minute for a single product photo. Still, the computation duration is open

⁶ Image upload: 0.65s) + (deep segmentation: 2.15s) + (SIFT extraction: 0.69s) + (matching 0.23s) + (downloading the results 0.025s) = (TOTAL 3.74s on average). Feature extraction for design files are performed offline.

to improvement with better hardware and further software optimization.

VI. CONCLUSIONS

The real-time image matching framework we propose in this paper is hybrid in the sense that it uses both hand-crafted features and deep features obtained from a well-tuned DCNN. We concentrate on a specific application, that is to say, printing design to product photo matching. Since photographs of a printed product suffer many unwanted effects, such as uncontrolled shooting angle, uncontrolled illumination, occlusions, printing deficiencies in color, camera noise, optic blur, et cetera, we benchmark different hand-crafted and deep features to choose an optimal performance and propose a framework, in which deep learning is utilized with highest contribution.

Our results show that a deep segmentation supported BoVW method gives satisfactory results for the proposed operational concept. What is more, hand-crafted features, when deep segmented from a region of interest may lead to better results, compared to deep features, which may include overwhelming number of noisy and unselective activations.

Like all current problems in computer vision, image matching problem is also moving to the DCNN domain. On the other hand, DCNNs require millions of data and expensive hardware. That's why we still need ingenious, practical and cheap industrial solutions until deep CNN hardware becomes standard in the following years.

In the meantime, we continue our studies on deep CNN structures, specifically on Siamese networks. We are currently building a Siamese network which can learn similarities between a product photo and design file pair. In order to train such a Siamese network, our analyses show that the current dataset must be significantly larger, compared to the dataset utilized in this study. Thus, we first focus our studies on enlarging our image set for training of such a system. Furthermore, a Siamese network will bring higher computation burden. For that matter, we are also studying embedded deep learning solutions that will utilize system-on-chip solutions for real-time operations.

ACKNOWLEDGMENT

This research was partially supported by the National Science Council of Turkey (TUBITAK - TEYDEB), with the project title "Customer-Information Matching, Invoicing and Barcoding of Post-Cut Products in Custom Printed Products using Image Processing Methods", under the 1507 program and with the project number "7170364". The authors would like to thank the owner of the project, Şans Printing Industries (bidolubaski.com) for their support and hard-work.

REFERENCES

- [1] T. Dharani, I. L. Aroquiaraj, "A survey on content based image retrieval," International Conference on Pattern Recognition, Informatics and Mobile Engineering, Tamilnadu, India, pp 485-490, 2013.
- [2] Y. Liu, D. Zhang, G. Lu, W.Y. Ma, "A survey of content-based image retrieval with high-level semantics," Pattern Recognition, vol. 40, 1, 2007, pp 262 - 282.
- [3] J. Sivic, A. Zisserman, "Video Google: a text retrieval approach to object matching in videos," International Conference on Computer Vision, 9th IEEE, Nice, France, vol. 2, pp 1470-1477, 2003.
- [4] H. Wang, Y. Cai, Y. Zhang, H. Pan, W. Lv, H. Han, "Deep learning for image retrieval: What works and what doesn't," International Conference on Data Mining Workshop, Washington, DC, US, pp 1576-1583, 2015.
- [5] J. Yosinski, J. Clune, A. Nguyen, T. Fuchs, H. Lipson, "Understanding neural networks through deep visualization," Deep Learning Workshop, International Conference on Machine Learning, Lille 2015, pp 2015.
- [6] K. Simonyan, A. Zisserman, "Very deep convolutional networks for large-scale image recognition" International Conference on Learning Representations Workshops, San Diego, CA, US, 2015.
- [7] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, "Going deeper with convolutions," IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, US, pp 1-9, June 2015.
- [8] A. Krizhevsky, I. Sutskever, G. E. Hinton, "ImageNet classification with deep convolutional neural networks," Advances in Neural Information Processing Systems (NIPS), Lake Tahoe, NV, US, pp 1097-1105, 2012.
- [9] P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, Y. LeCun, "Overfeat: Integrated recognition, localization and detection using convolutional networks," International Conference on Learning Representations, ICLR, Banff, Canada, 2014.
- [10] A. Babenko, A. Slesarev, A. Chigorin, V. Lempitsky, "Neural codes for image retrieval," European Conference in Computer Vision, Zurich, Switzerland, pp 584-599, 2014.
- [11] V. Chandrasekhar, J. Lin, O. Morère, H. Goh, A. Veillard, "A practical guide to CNNs and fisher vectors for image instance retrieval," Signal Processing, vol. 128, pp 426-439, 2016.
- [12] I. Melekhov, J. Kannala, and E. Rahtu, "Siamese network features for image matching," International Conference on Pattern Recognition (ICPR), Cancún, Mexico, pp 378-383, 2016.
- [13] Y. Taigman, M. Yang, M. Ranzato, L. Wolf, "DeepFace: Closing the gap to human-level performance in face verification," IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, US, pp 1701-1708, 2014.
- [14] T. Lin, Y. Cui, S. Belongie, J. Hays, "Learning deep representations for ground-to-aerial geolocalization," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, US, pp 5007-5015, 2015.
- [15] D. Cai, X. Gu, C. Wang, "A revisit on deep hashings for large-scale content based image retrieval," ArXiv.CoRR, vol. abs/1711.06016, pp 1-11, 2017.
- [16] R. Datta, J. Li, J. Z. Wang, "Content-based image retrieval: Approaches and trends of the new age", ACM SIGMM International Workshop on Multimedia Information Retrieval, New York, NY, USA, pp. 253-262, 2005.
- [17] P. Clough, H. Müller, T. Deselaers, M. Grubinger, T. Martin Lehmann, J. R. Jensen, W. Hersh, "The CLEF 2005 Cross-Language Image Retrieval track," International Conference of the Cross-Language Evaluation Forum for European Languages, Vienna, Austria, vol. 1171, pp. 535-557, 2005.
- [18] G. Schaefer, "UCID-RAW - a colour image database in raw format," European Congress on Computational Methods in Applied Sciences and Engineering, Porto, Portugal, pp 179-184, 2017.
- [19] T. Ahmad, P. Campr, M. Cadik, G. Bebis, "Comparison of semantic segmentation approaches for horizon/sky line detection," International Joint Conference on Neural Networks (IJCNN), Anchorage, AK, US, pp 4436-4443, 2017.
- [20] F. Jiang, A. Grigorev, S. Rho, Z. Tian, Y. Fu, W. Jifara, A. Khan, S. Liu, "Medical image semantic segmentation based on deep learning," Neural Computing and Applications, vol 29. 5, pp 1257-1265, 2018.
- [21] M. Siam, S. Elkerdawy, M. Jägersand, S. Yogamani, "Deep semantic segmentation for automated driving: Taxonomy, roadmap and challenges," IEEE International Conference on Intelligent Transportation Systems, Yokohama, Japan, pp. 1-8, 2017.
- [22] I. Ulku, E. Akagunduz, "A Survey on Deep Learning-based Architectures for Semantic Segmentation on 2D images," ArXiv.Corr, vol. abs/1912.10230, pp 1-20, 2019.
- [23] M. H. Saffar, M. Fayyaz, M. Sabokrou, M. Fathy, "Semantic video segmentation: A review on recent approaches," ArXiv.Corr, vol. abs/1806.06172, pp 1-24, 2018.
- [24] H. Yu, Z. Yang, L. Tan, Y. Wang, W. Sun, M. Sun, Y. Tang, "Methods and datasets on semantic segmentation: A review," Neurocomputing, vol. 304, pp. 82 - 103, 2018.

- [25] Y. Guo, Y. Liu, T. Georgiou, M. S. Lew, "A review of semantic segmentation using deep neural networks," *International Journal of Multimedia Information Retrieval*, vol. 7, pp. 87-93, Jun 2018.
- [26] A. Garcia-Garcia, S. Orts-Escolano, S. Oprea, V. Villena-Martinez, and J. G. Rodriguez, "A review on deep learning techniques applied to semantic segmentation," *ArXiv.Corr*, vol. abs/1704.06857, pp 1-19, 2017.
- [27] E. Shelhamer, J. Long, T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, pp. 640-651, Apr. 2017.
- [28] A. Vedaldi, K. Lenc, "MatConvNet: Convolutional neural networks for Matlab," *ACM International Conference on Multimedia, Brisbane Australia*, pp. 689-692, 2015.
- [29] H. Zhang, J. E. Fritts, and S. A. Goldman, "Image segmentation evaluation: A survey of unsupervised methods," *Computer Vision and Image Understanding*, vol. 110. 2, pp. 260 - 280, 2008.
- [30] J. Harel, C. Koch, and P. Perona, "Graph-based visual saliency," *Advances in Neural Information Processing Systems, Vancouver, Canada*, pp 545-552, 2006.
- [31] Y. Xu, J. Li, J. Chen, G. Shen, Y. Gao, "A novel approach for visual saliency detection and segmentation based on objectness and top-down attention," *International Conference on Image, Vision and Computing, Chengdu, China*, pp 361-365, 2017.
- [32] J. Lankinen, V. Kangas, J. Kamarainen, "A comparison of local feature detectors and descriptors for visual object categorization by intra-class repeatability and matching," *International Conference on Pattern Recognition, Tsukuba, Japan*, pp 780-783, 2012.
- [33] A. Oliva, A. Torralba, "Modeling the shape of the scene: A holistic representation of the spatial envelope," *International Journal of Computer Vision*, vol. 42, pp. 145-175, 2001.
- [34] N. Dalal, B. Triggs, "Histograms of oriented gradients for human detection," *IEEE Computer Vision and Pattern Recognition, San Diego, CA, US*, pp. 886-893, 2005.
- [35] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, pp. 91-110, 2004.
- [36] H. Bay, T. Tuytelaars, L. Van Gool, "SURF: Speeded up robust features," *European Conference on Computer Vision, Graz, Austria*, pp 404-417, 2006.
- [37] K. He, X. Zhang, S. Ren, J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," *Arxiv.CoRR*, vol. abs/1406.4729, pp 1-13, 2014.
- [38] C. Hentschel, H. Sack, "Does one size really fit all?: Evaluating classifiers in bag-of-visual-words classification," *International Conference on Knowledge Technologies and Data-driven Business, New York, NY, USA*, pp. 7:1-7:8, 2014.
- [39] L. I. Kuncheva, "On the optimality of naïve bayes with dependent binary features," *Pattern Recognition Letters*, vol. 27, pp. 830-837, 2006.
- [40] K. Lenc, A. Vedaldi, "Understanding image representations by measuring their equivariance and equivalence," *IEEE Conference On Computer Vision and Pattern Recognition, Boston, MA, US*, pp. 991-999, 2015.

BIOGRAPHIES



ALPER KAPLAN was graduated from Yeditepe University, Computer Engineering Department in 2016. Currently, he is a research assistant and a master's student at Cognitive Science Program of Yeditepe University. He continues his studies by focusing on his thesis about machine-generated music. In addition to his academic studies, he

works as a software developer in an e-commerce company. His interests consist of machine learning, machine-generated art/music, music theory, natural language processing, image processing.



ERDEM AKAGÜNDÜZ received his Ph.D. degree in Middle East Technical University, Electrical and Electronics Engineering Department, Ankara, Turkey, in 2011. He worked a research and teaching assistant with the METU EEE CVIS Lab. from 2001 to 2008. Between 2009-2016 he worked as an algorithm design engineer with

ASELSAN. Before starting his academic career in Turkey, as a post-doctoral research associate, we visited the University of York, UK, in 2016. He is currently an Assistant Professor in Çankaya University, Electrical and Electronics Engineering Department, Ankara, Turkey. His research interests include deep learning, computer vision and sound processing.

A Stacking-based Ensemble Learning Method for Outlier Detection

AA. ABRO, E. TAŞCI, A. UĞUR

Abstract—Outlier detection is considered as one of the crucial research areas for data mining. Many methods have been studied widely and utilized for achieving better results in outlier detection from existing literature; however, the effects of these few ways are inadequate. In this paper, a stacking-based ensemble classifier has been proposed along with four base learners (namely, Rotation Forest, Random Forest, Bagging and Boosting) and a Meta-learner (namely, Logistic Regression) to progress the outlier detection performance. The proposed mechanism is evaluated on five datasets from the ODDS library by adopting five performance criteria. The experimental outcomes demonstrate that the proposed method outperforms than the conventional ensemble approaches concerning the accuracy, AUC (Area Under Curve), precision, recall and F-measure values. This method can be used for image recognition and machine learning problems, such as binary classification.

Index Terms— Outlier detection, Ensemble learning, Machine Learning, Classification, Data Mining.

I. INTRODUCTION

OUTLIER IS defined as an observation that deviates from other observations or suspicious events that are generated by different mechanisms. Outliers are anomalous, irregular, or outlying reflections, the distortion of estimations in statistical models [1].

This is one of the best approaches of data analysis to deal with observations having numerous datasets, as automated tools are being used in it to find patterns and relationships. In recent years, outlier detection has been widely used in several industries, such as medical, to detect credit card frauds and sensors (IoT).

ABDUL AHAD ABRO, is with the Department of Computer Engineering, Ege University, Izmir, Turkey, (e-mail: abduhadabro1@gmail.com).



<https://orcid.org/0000-0002-3591-9231>

ERDAL TAŞCI, is with the Department of Computer Engineering, Ege University, Izmir, Turkey, (e-mail: arif.erdal.tasci@ege.edu.tr).



<https://orcid.org/0000-0001-6754-2187>

AYBARS UĞUR, is with the Department of Computer Engineering, Ege University, Izmir, Turkey, (e-mail: aybars.ugur@ege.edu.tr).



<https://orcid.org/0000-0003-3622-7672>

Manuscript received January 24, 2020; accepted April 14, 2020.
DOI: [10.17694/bajece.679662](https://doi.org/10.17694/bajece.679662)

Ensemble Learning is a machine learning technique that aggregates various base models to generate a single predictive model. Numerous methods are used in Ensemble Learning to reduce bias (boosting), variance (bagging), or to progress predictions (stacking) [2]. It also means that the concept provides a promissory field of future research.

While Random Forest was developed approximately two decades ago, it gives a powerful performance, simplicity in implementation and interpretability [3].

On the other hand, Rotation Forest, which is proposed by Pardo [4-5], provides favourable outcomes when compared to AdaBoost, Random Subspaces, Bagging and Iterated Bagging. The principal contribution of this paper is a) A stacking-based ensemble learning method which improves the outlier detection performance ii) A comparative analysis of four base learners and one Meta-ensemble learner on five datasets from the ODDS library in terms of five evaluation criteria; accuracy (Acc), AUC, precision, recall and F-measure.

This paper is structured with several different sections. In section II, related work presents ideas about ensemble methods. Section III discusses the proposed method in detail. Section IV, provides experimental work, detail of datasets and outcomes. Section V, is related to the evaluation of performance and results. Lastly, conclusion and future work are suggested in Section VI.

II. RELATED WORK

Outliers are mainly segregated into three main areas: Collective outliers, global outliers and contextual outliers [3]. Global Outliers consider that outliers are associated with all the available data points. Contextual Outliers consider that data separated from other data points in the context. However, Collective Outlier values are different data groups that are inaccurate according to a complete dataset. Outlier values are also known as abnormal as they examine the change to identify unexpected behaviour [4].

A static ensemble shows the base learner and the fusion rule is fixed for each single test point [5]. Generally, Bagging and Random subspace methods are employed in these processes. For instance, methods used to generate numerous diverse training subsets for base learners are combined in bagging and random subspace. Many ensemble approach are also applied over clustering algorithms. Therefore, the aggregation and structure standard of the ensemble is set for each single test point in this form of outlier detection strategy [6].

In other studies, Rotboost is a classifier of an ensemble, inferred by combining the AdaBoost and Rotation forest. There are various datasets from the UCI ML repository,

among which a classification tree that is being utilized as the base learning algorithm. It has been shown by their results that Rotboost could generate a lower prediction error in an ensemble classifier in comparison to Rotation Forest or AdaBoost [5]. The ensemble learning approaches such as bagging mainly emphasis to get an ensemble model with less variance than its components; whereas, boosting and stacking generally try to generate strong models less biased than their components even if variance can also be condensed. Random Forest (a subprocess of the Meta-ensemble method) is used as a base learner in the rotation forest. This approach has enhanced performances [7].

In [8], polarized images have been classified using Random Forest and Rotation Forest and it is concluded that Rotation Forest provides more accurate results than SVM and Random Forest; however, Random Forest provides faster results than Rotation Forest.

It is examined whether Rotation Forest is the best classifier that assists in resolving problems with continuity or not. [6].

In [9], A-Stacking and A-Bagging, the adaptive versions of ensemble learning approaches are proposed. A-Bagging method has been applied by using the same base learners over numerous subsets of data and the predictions are aggregated by using weighted majority voting.

In [10], it is shown that ML algorithms provide satisfactory performance for the prediction of the outcomes in comparison with logistic regression.

III. DETAILS OF THE PROPOSED METHOD

In this paper, we have proposed a framework of a Stacking-based ensemble learning method, including rotation forest, random forest, bagging, boosting and logistic regression. There are numerous phases of the system, such as related with datasets, base and stacking-based ensemble learners. In order to obtain the generalization performance of the system, 10-fold cross-validation is used for all learners and datasets. The ranges of the values in data pre-processing may be high when compared to non-outlier datasets. In this scenario, classification algorithms could be affected significantly or negatively by some features.

In this work, four base learners and one Meta-learner are employed with one Stacking-based Meta classifier. Rotation Forest classifier depends upon feature extraction for ensembles. Typically, it provides more authentic results than AdaBoost and Random Forest. The Random Forest classifier is based on several collections of tree classifiers and randomly selected sub-spaces of data are being used to create each classifier independently.

Ensemble Learning such as bagging and boosting assist in diminishing various influences such as classification error. Furthermore, combinations of many classifiers drop variance, particularly in the case of unstable classifiers, which may generate a more reliable classification than a single classifier. The main idea of this study is to establish and provide data comprised of detecting outliers to present new methods related to outlier detection in classification with logistic regression.

Whereas, logistic regression predicts to analyze, explain and indicate the interrelation between one nominal and a dependent binary variable, ratio-level independent or interval

variables. Weka (Waikato Environment for Knowledge Analysis)[11], could examine and test multiple outliers[12], without losing the impact of swamping and masking. We demonstrated the behaviour of our method through simulation with different percentages of outliers and sample sizes. In this process, the different datasets have been utilized referred to from the ODDS library.

IV. EXPERIMENTAL WORK

In the experimental process, five datasets have been used from the ODDS library for classifications [13].

The characteristics of datasets are analyzed concerning the attributes and the number of instances. These datasets are generally used to solve issues related to machine learning. There are no missing values in these datasets and there are various numerical attribute descriptions, which are illustrated in Table I. As it can be observed from Table I, various datasets, the number of samples, dims and outliers are presented for each dataset. Datasets are chosen according to their distinct parameters from the ODDS library source. It is determined by investigating the appropriate data or datasets which are being utilized in the findings of outliers. The proposed stacking-based ensemble learning method has been introduced for this process. This method utilized the imbalanced classification problems of binary (two-class) where the positive case such as (class 1) is taken as an outlier and negative case (class 0) is taken as normal.

TABLE I
DATASETS DESCRIPTIONS

Datasets	Samples	Dims	Outliers
Class	214	9	9 (4.2%)
Letter Recognition	1600	32	100(6.25%)
Shuttle	49097	9	3511 (7%)
Forest Cover	286048	10	2747(0.9%)
Vertebral	240	6	30 (12.5%)

In this work, four different ensemble learning approaches have been carried out along with the ensemble learning method, which is considered suitable for the detection of outliers. However, the performance metrics are calculated based on outlier detection according to binary classification problems. In this method, a technique has been used, which is known as logistic regression from the field of statistics and it is being used to solve binary classification issues. A stacking-based ensemble method, along with logistic regression and four different baseline methods have been presented in Fig. 1.

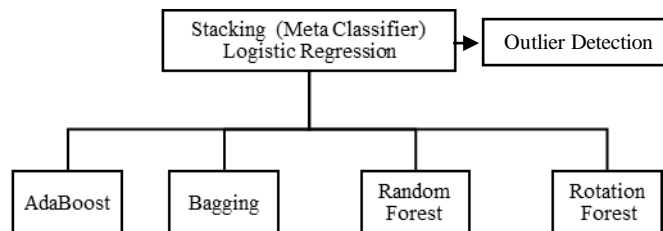


Fig.1. Ensemble Learning baseline methods

Bagging is a modest and very influential ensemble process. It is considered as the Bootstrap procedure to a high-variance ML algorithm. Simultaneously, Boosting denotes a group of

algorithms that employ weighted averages to interchange the weak learners into stronger learners. The random forest consists of multiple random decision trees [14]. Rotation forest is a tree-based ensemble that performs and transforms on subsets of attributes before constructing each tree.

V. PERFORMANCE EVALUATION

A. Evaluation Measures

This section describes the five performance evaluation measures of the proposed method, consisting of accuracy, AUC, precision, recall and F-measure.

Accuracy represents how near a measurement is to an identified or accepted figure. It is further defined in Eq.1.

$$Acc = \frac{TP+TN}{TP+FP+FN+TN} \quad (1)$$

In equation 1, TN, FN, FP and TP show the number of True Negatives, False Negatives, False Positives and True Positives.

AUC represents the Area under the ROC Curve. AUC calculates the whole two-dimensional area beneath the whole ROC curve from (0,0) to (1,1).

Precision is a positive analytical value [15]. Precision defines how reliable measurements are, although they are farther from the accepted value.

The equation of precision is shown in Eq.2.

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

The Recall is the hit rate [15]. The recall is the reverse of precision; it calculates false negatives against true positives. The equation is illustrated in Eq. 3.

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

F-measure can be defined as the weighted average [16] of precision and recall. This rating considers both false positives and false negatives. The equation is illustrated in Eq. 4.

$$F = 2x \frac{Precision+Recall}{Precision+Recall} \quad (4)$$

Tables II-VII present accuracy, AUC, precision, recall and F-measure individual values with ensemble methods for all datasets.

To sum up, Tables II-VI, have been designed according to the diverse data sets concerning the numerous approaches of ensemble learning in terms of different specifications. In Table II, logistic regression has better outcomes, which provides 99.5327% Acc in comparison to others. Likely, in Table III, rotation forest indicates 95.1875% Acc adequate consequences. Similarly, in Table IV, the random forest presents 99.9939% Acc effective results. Likewise, in Table

V, the random forest illustrates the 99.9857% Acc productive outcomes. However, in the end, logistic regression shows a 92.5% Acc result in Table VI.

TABLE II
RESULTS OF ENSEMBLE LEARNING METHODS BY UTILIZING THE GLASS DATASET

Glass					
Methods	Acc (%)	AUC	Precision	Recall	F-Measure
Bagging	96.2617	0.988	0.954	0.963	0.954
AdaBoost	99.0654	0.996	0.991	0.991	0.990
Random Forest	97.6636	0.997	0.975	0.977	0.974
Rotation Forest	97.1963	0.993	0.969	0.972	0.968
Logistic Regression	99.5327	0.999	0.996	0.995	0.995

TABLE III
RESULTS OF ENSEMBLE LEARNING METHODS BY UTILIZING THE LETTER RECOGNITION DATASET

Letter Recognition					
Methods	Acc (%)	AUC	Precision	Recall	F-Measure
Bagging	94.8750	0.944	0.949	0.949	0.932
AdaBoost	93.7500	0.744	0.938	0.938	0.968
Random Forest	95.0000	0.987	0.953	0.950	0.934
Rotation Forest	95.1875	0.930	0.952	0.952	0.938
Logistic Regression	94.0625	0.813	0.925	0.941	0.926

TABLE IV
RESULTS OF ENSEMBLE LEARNING METHODS BY UTILIZING THE SHUTTLE DATASET

Shuttle					
Methods	Acc (%)	AUC	Precision	Recall	F-Measure
Bagging	99.9919	0.999	1.000	1.000	1.000
AdaBoost	99.8330	1.000	0.998	0.998	0.998
Random Forest	99.9939	1.000	1.000	1.000	1.000
Rotation Forest	99.9817	1.000	1.000	1.000	1.000
Logistic Regression	99.6497	0.988	0.997	0.996	0.996

TABLE V
RESULTS OF ENSEMBLE LEARNING METHODS BY UTILIZING THE FOREST COVER DATASET

Forest Cover					
Methods	Acc (%)	AUC	Precision	Recall	F-Measure
Bagging	99.9790	1.000	1.000	1.000	1.000
AdaBoost	99.8133	0.999	0.998	0.998	0.998
Random Forest	99.9857	1.000	1.000	1.000	1.000
Rotation Forest	99.9773	1.000	1.000	1.000	1.000
Logistic Regression	99.8941	1.000	0.999	0.999	0.999

TABLE VI
RESULTS OF ENSEMBLE LEARNING METHODS BY UTILIZING THE
VERTEBRAL DATASET

Vertebral					
Methods	Acc (%)	AUC	Precision	Recall	F-Measure
Bagging	91.6667	0.887	0.912	0.917	0.903
AdaBoost	87.0833	0.879	0.844	0.871	0.852
Random Forest	91.6667	0.889	0.910	0.917	0.905
Rotation Forest	91.6667	0.929	0.909	0.917	0.909
Logistic Regression	92.5000	0.930	0.919	0.925	0.920

In general, bagging has more successive consequences than boosting, whereas, the random forest provides more effective outputs than rotation forest in most of the datasets. On the other hand, logistic regression has also provided satisfactory results to some extent, which is illustrated in Tables II and VI. In Table VII, a stacking-based ensemble learning method has been applied, in which the model is trained with the combined prediction preceding model. The logistic regression has been set as a Meta classifier and experienced the diverse datasets with numerous methods like rotation forest, random forest, boosting and bagging in the given order.

The letter recognition, forest cover and vertebral datasets have significant outputs concerning the accuracy, AUC, precision, recall and F-measure parameters in Table VII; however, glass and shuttle datasets show similar outcomes for Tables II and IV.

Table VII demonstrates the comparison of all datasets results, with respect to our proposed stacking-based meta-ensemble learning method. As it is clearly shown in Table VII, a Meta-ensemble classifier, stacking with four base learners (namely, Rotation Forest, Random Forest, Bagging and Boosting) and one Meta-learner (namely, Logistic Regression) provide highly accurate outcomes as compare to others.

TABLE VII
OUR STACKING-BASED ENSEMBLE LEARNING METHOD

Classifier	Proposed Stacking Meta Classifier Logistic Regression					
	Acc (%)	Impr. (%)	AUC	Precision	Recall	F-Measure
Glass	*99.5327	0.0000	0.997	*0.996	*0.995	*0.995
Letter Recognition	97.4375	2.2500	*0.987	0.973	0.974	0.974
Shuttle	*99.9939	0.0000	*1.000	*1.000	*1.000	*1.000
Forest Cover	99.9860	0.0003	*1.000	*1.000	*1.000	*1.000
Vertebral	93.3333	0.8333	0.903	0.931	0.933	0.926

- * Indicates the similar performance results concerning base learner.
- High Acc, AUC, Precision, Recall and F-measure is shown in Bold, while the greyed shows insufficient results.
- Impr. represents improvement according to best results of Tables II-VI.

Moreover, in Table VII, it is analyzed that when stacking based ensemble learning method combines with logistic regression, it provides more accurate outcomes than logistic

regression; whereas, logistic regression does not provide better outcomes when applied individually.

VI. CONCLUSION AND FUTURE WORK

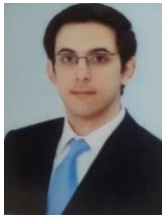
This research work has proposed an approach of ensemble classifiers in multiple datasets efficiently for outlier detection. However, these different methods such as Random forest, Rotation forest, Bagging and Boosting (base learners) and Meta-learner logistic regression under stacking classifiers occupy more space and consume more time for computations. This method enables us to provide more productive and effective outputs by using the advantages of these algorithms. We believe that this scenario is suitable for both research and commercial applications. The performance of the classifier models can be different depending on the datasets that will be chosen. In the future, other hybridization of ensemble learning methods will be employed for performance improvement.

REFERENCES

- [1] Ö. G. Alma, S. Kurt and U. Aybars, "Genetic algorithms for outlier detection in multiple regression with different information criteria," vol. 9655, 2011.
- [2] C. Pardo, J. F. Diez-Pastor, C. Garcia-Osorio and J. J. Rodríguez, "Rotation Forests for regression," *Appl. Math. Comput.*, vol. 219, no. 19, pp. 9914–9924, 2013.
- [3] L. Chen, S. Gao and X. Cao, "Research on real-time outlier detection over big data streams," *Int. J. Comput. Appl.*, vol. 7074, pp. 1–9, 2017.
- [4] N. Simidjievski, "Predicting long-term population dynamics with bagging and boosting of process-based models," vol. 42, pp. 8484–8496, 2015.
- [5] C. Zhang and J. Zhang, "RotBoost : A technique for combining Rotation Forest and AdaBoost," vol. 29, pp. 1524–1536, 2008.
- [6] A. Bagnall, M. Flynn, J. Large, J. Line, A. Bostrom and G. Cawley, "Is rotation forest the best classifier for problems with continuous features?," 2018.
- [7] E. Taşçı, "A Meta-Ensemble Classifier Approach: Random Rotation Forest," *Balk. J. Electr. Comput. Eng.*, vol. 7, no. 2, pp. 182–187, 2019.
- [8] P. Du, A. Samat, B. Waske, S. Liu and Z. Li, "Random Forest and Rotation Forest for fully polarized SAR image classification using polarimetric and spatial features," *ISPRS J. Photogramm. Remote Sens.*, vol. 105, pp. 38–53, 2015.
- [9] S. Agarwal and C. R. Chowdary, "A-Stacking and A-Bagging: Adaptive versions of ensemble learning algorithms for spoof fingerprint detection," *Expert Syst. Appl.*, vol. 146, p. 113160, 2020.
- [10] J. Zhou Feng, Y. Wang, J. Peng, M. wei Sun, J. Zeng and H. Jiang, "Comparison between logistic regression and machine learning algorithms on survival prediction of traumatic brain injuries," *J. Crit. Care*, vol. 54, pp. 110–116, 2019.
- [11] Eibe Frank, Mark A. Hall and Ian H. Witten (2016). *The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques"*, Morgan Kaufmann, Fourth Edition, 2016.
- [12] T. A. Engel, A. S. Charão, M. Kirsch-Pinheiro and L. A. Steffanel, "Performance improvement of data mining in weka through GPU acceleration," *Procedia Comput. Sci.*, vol. 32, pp. 93–100, 2014.
- [13] Shebuti Rayana (2016). *ODDS Library* [<http://odds.cs.stonybrook.edu>]. Stony Brook, NY: Stony Brook University, Department of Computer Science.
- [14] Y. Zhou and G. Qiu, "Random forest for label ranking," *Expert Syst. Appl.*, vol. 112, pp. 99–109, 2018.
- [15] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognit. Lett.*, vol. 27, no. 8, pp. 861–874, 2006.
- [16] L. A. Bull, K. Worden, R. Fuentes, G. Manson, E. J. Cross, and N. Dervilis, "Outlier ensembles: A robust method for damage detection and unsupervised feature extraction from high-dimensional data," *J. Sound Vib.*, vol. 453, pp. 126–150, 2019.



ABDUL AHAD ABRO was born in 1988. He received the BS degree from the University of Sindh, MSc degree from Mohammad Ali Jinnah University and enrolled in a PhD degree at the Ege University, Computer Engineering Department, Izmir, Turkey. From 2014 to 2015, he was an ad-hoc lecturer in Sindh Madressatul Islam University. He also serves as a reviewer for numerous indexed journals & conferences. His research interests are machine learning, data mining and ensemble learning.



ERDAL TAŞCI was born in 1989. He received the BS, MSc and PhD degrees from the Ege University, Computer Engineering Department, Izmir, Turkey, in 2011, 2013 and 2018, respectively. He has been working as a teaching assistant at the Department of Computer Engineering of Ege University, Izmir, Turkey. He has been reviewing for various journals including IEEE Transactions on Systems, Man and Cybernetics: Systems, Computer Methods and Programs in Biomedicine and IET Image Processing. His research interests are pattern recognition, machine learning, image processing and data mining.



AYBARS UĞUR is a full-time professor in the Department of Computer Engineering at Ege University, Izmir, Turkey. He received his BS, MSc and Ph.D. degrees in computer engineering from Ege University, Izmir, Turkey, in 1993, 1996, 2001, respectively. His research interests are artificial intelligence, machine learning, deep learning, swarm intelligence, computer vision, optimization, intelligent systems and computer graphics.

Nested Miller Compensation Based Op-Amp Design for Piezoelectric Actuators

M.A. ÇELİK, D. GÖKCEN*, F.E. AYDOS


Abstract—This study introduces the design of a practical three-stage operational amplifier (op-amp) using nested Miller compensation, particularly for piezoelectric actuators. Driving a piezoelectric actuator represents a challenge in amplifier design due to its large capacitive nature. A stable piezo driver needs to be free of oscillations and phase lag. Direct feedback compensation using a conventional Miller capacitor is an effective method as long as the capacitance of the load is considerably close to the value of the Miller capacitor. However, using a large capacitor causes a decrease in the slew rate and gain bandwidth. To avoid this, our design focused on the utilization of nested Miller compensation technique. A prototype of the design working at 100V peak to peak voltage (V_{pp}) is implemented using commercial off-the-shelf (COTS) components. The measurements show the successful driving capability and step-response of the op-amp design. In the design, Widlar current source is also utilized for thermal stability and short circuit protection. According to simulation results, the proposed op-amp has a slew rate of 0.5 V/ μ s, an open loop gain of 90dB with 3MHz Gain Bandwidth Product (GBP) and phase margin of 77°, and a common mode rejection ratio (CMRR) of 62dB.

Index Terms—Operational Amplifier, Stability, Piezoelectric, Actuator, Piezo Driver


I. INTRODUCTION

THE WIDESPREAD use of piezoelectric actuators in diverse applications stimulate significant research efforts


DİNÇER GÖKCEN, is with Department of Electrical and Electronics Engineering, Hacettepe University, Ankara, Turkey, (Corresponding Author e-mail: dgokcen@hacettepe.edu.tr).

 <https://orcid.org/0000-0003-1847-1356>

MEHMET AKİF ÇELİK, is with Department of Electrical and Electronics Engineering, Hacettepe University, Ankara, Turkey, (e-mail: mehmetakifcelik972@gmail.com).

 <https://orcid.org/0000-001-5119-8705>

FATİH EMRE AYDOS is with Department of Electrical and Electronics Engineering, Hacettepe University, Ankara, Turkey, (e-mail: eaydos.0@gmail.com).

 <https://orcid.org/0000-0002-6726-4567>

Manuscript received July 02, 2019; accepted March 13, 2020.
DOI: [10.17694/bajece.585798](https://doi.org/10.17694/bajece.585798)

for the development of piezo drivers with high linearity and low-cost [1-3]. Due to their unique capability to make displacement in nanoscale, piezo materials are widely used in many scientific applications including atomic force microscopy (AFM), scanning tunneling microscopy (STM), nano- and micromanipulators, and optical microscopy stages [4-6]. In addition to its expanding use in scientific applications, piezoelectric materials have already been utilized in many consumer products, such as inkjet printers, lenses of the smartphones, lasers, fuel injectors, touch sensors, ultrasonic cleaners, etc [2,7-11]. Piezoelectric transducers are made up of dielectric materials with the ability to convert mechanical stress into electrical signals and vice versa. Typically, a standard piezoelectric crystal will expand at nanometer scale per a certain voltage applied. The piezoelectric materials inherently act as a capacitive element (typically 0.01 microfarad or more) and their drivers require considerably high voltage range that may not be easily handled by integrated systems. The piezoelectric materials exemplify a non-ideal type of load for conventional op-amp designs because of their large capacitance. Additionally, in piezoelectric materials, mechanical resonances appear at the frequency range varying from kHz to MHz [11]. This implies that the instability in op-amp operation can induce oscillations at high frequencies and this phenomenon may also trigger undesired mechanical resonance.

The aim of this study is to build an op-amp to drive large capacitive loadings at a wide range of frequency and voltage. Nested Miller compensation, short circuit protection, and overtemperature operation are exclusively considered in the design. Particularly, Miller and nested Miller compensation techniques are compared in detail. To demonstrate the validity of the techniques provided, the design is implemented using commercial off-the-shelf (COTS) circuit components. The circuit built by COTS components is powered by a peak to peak supply voltage (V_{pp}) of 100V, has a common mode rejection ratio (CMRR) of 62 dB, 0.5 V/ μ s slew rate, and open loop gain value of 90 dB with 3MHz bandwidth and 77° phase margin. Well-known transistor types (MMBTA42 and its complementary MMBTA92) are preferred in the prototype to show the feasibility of the system and to reduce the complexity of the implementation for scientists and engineers from diverse fields.

A typical op-amp involves a differential amplifier, voltage gain, level shifter, and power output stages. The three-stage design provided in this study has an additional differential amplifier stage to achieve higher gain. Combination of each stage is supposed to provide efficiency and linearity in

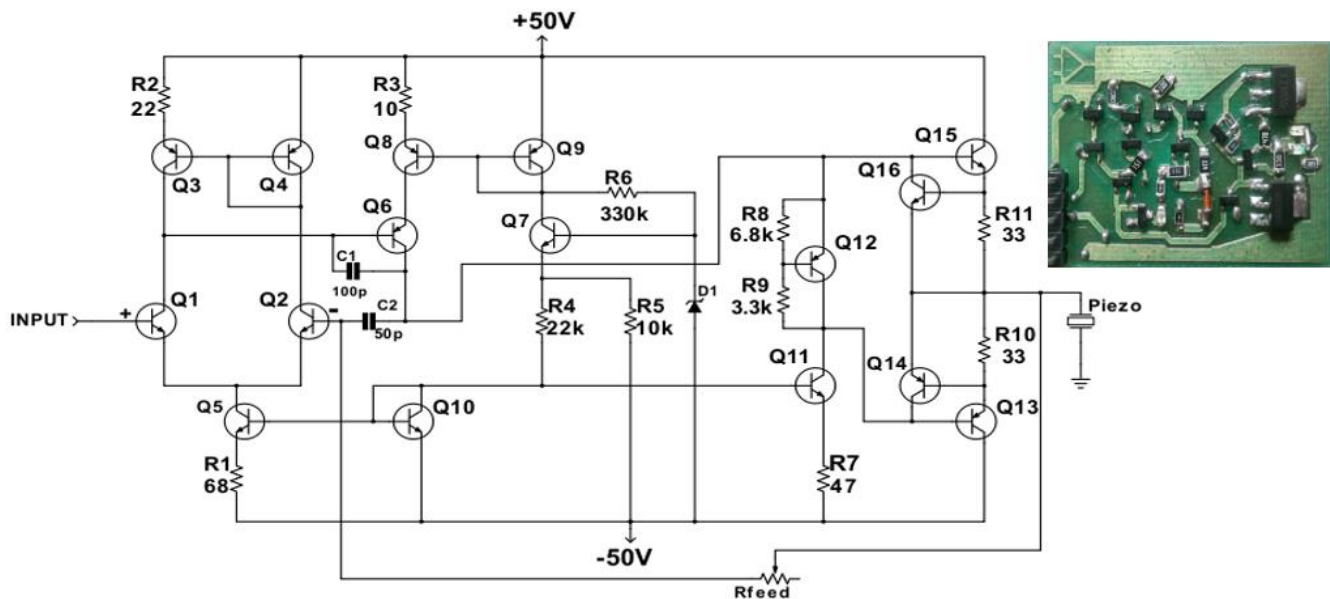


Fig. 1. Op-amp Circuit Diagram. The inset shows the picture of the op-amp module

amplification. The output stage has a Class AB amplifier configuration with two complementary transistors (FZT 855 and its complementary FZT955). The stability of the amplifier is directly associated with the compensation. Direct feedback compensation using a Miller capacitor is a well-studied concept, especially in CMOS based integrated circuits [12-14]. Miller compensation relies on the pole-splitting principle and no doubt, it efficiently solves the instability issues in many applications. Piezoelectric materials are of non-ideal loads for conventional op-amps. Their capacitance is large and subject to changes with varying voltage. Miller compensation technique using a single capacitor turns out to be not functional if the capacitance of the load is much greater than the value of the Miller capacitor. Increase in the Miller capacitor to improve the stability have adverse impacts on slew rate and gain bandwidth [15-18]. Considering the large capacitance of piezoelectric actuators as well as their varying capacitance under bias, the op-amp design given here utilizes nested Miller compensation technique. The major advantage of the nested Miller compensation is that stability can be improved with smaller capacitors and the speed of op-amp operation is considerably not affected by the compensation [15,16].

In a similar manner with many electronic devices working at high voltages, short circuit protection stands as an essential part of the op-amp design to ensure the reliability in extreme operating conditions. Widlar circuit with Class AB power amplifier configuration is utilized for short circuit protection and also ensures the overtemperature operation of the op-amp. The Widlar approach is important particularly in high-temperature applications to ensure the reliable operation of the instrument at non-standard thermal conditions [19].

II. CIRCUIT DESIGN AND PROTOTYPING

Fig. 1 shows the circuit schematic of the piezo driving op-amp with a picture of the prototype. The input stage of the op-amp circuit is a conventional n-type differential amplifier driven by a current source pair (Q3 and Q4). Inverting and noninverting inputs are respectively the bases of Q1 and Q2 with internal feedback capacitors (C1 and C2) and external feedback resistor (Rfeed). Rfeed is a zero-nulling resistor for nested Miller compensation. A current mirror consisting of Q5, Q10, and Q11 transistors yields to a bias current of 0.5mA. This part of the design is a Widlar current source, which is widely used to produce small current values using small resistors [19,20]. Widlar circuit is extensively utilized in integrated circuits because it effectively ensures the temperature stability and short circuit protection with small resistors. Q5 is the current sink for the differential pair and produces 0.25 mA bias current for both Q1 and Q2. Q8 and Q9 form another current mirror which provides 0.5 mA bias current for amplifier transistor Q6. 6.2V Zener is used for voltage regulation and also for sourcing the current. It eliminates the fluctuations at the voltage and prevents the system from abnormal changes at the current value. A stable current of 0.77mA can be observed from both Q10 and Q9 current mirror transistors. R6 sustains suitable current for Zener diode whereas R1, R4, and R5 provide 0.5 mA bias current for Q5 as sink current for the differential pair. R8 and R9 are the V_{BE} multiplier to eliminate crossover distortion for the output stage. Q13 and Q15 are the output power transistors with current limiter transistors Q14 and Q16. The output stage of the op-amp is a Class AB amplifier with temperature tolerable short circuit protection. R11- Q16 and R10 - Q14 pairs are the current limiters and constrain the output current

at about 15 mA. C1 is the Miller capacitor utilized for the direct compensation of the operational amplifier. C2 is the supporting compensation capacitor that shows derivative characteristics with the zero nulling-feedback resistor (R_{feed}). Combination of C1 and C2 capacitors presents the nested Miller compensation circuitry. Well-known BJT transistors are utilized in the prototype for simplicity. Class AB amplifier consists of an FZT 855 (Q15) and its complementary FZT955(Q13), and the rest of the circuit have MMBTA42 and its complementary MMBTA92. While choosing the components, their costs, availability in the market, low voltage drop and functionality at high voltage were assessed.

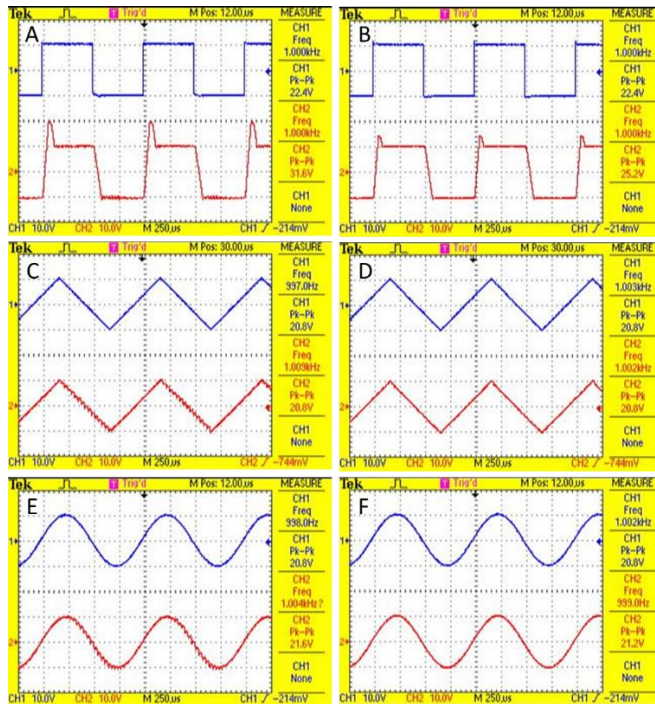


Fig. 2. Output voltage characteristics of a 47nF piezoelectric load with a) Miller compensation for 1kHz square wave input at unity gain, b) nested Miller compensation for 1kHz square wave input at unity gain, c) Miller compensation for 1kHz triangular wave input at unity gain, d) nested Miller compensation for 1kHz square wave input at unity gain, e) Miller compensation for 1kHz sinusoidal wave input at unity gain, f) nested Miller compensation for 1kHz sinusoidal wave input at unity gain. (Blue lines are inputs, whereas red lines represent outputs)

In a piezo driving amplifier, the output voltage is limited by the output impedance and instability might occur due to the large capacitance of the piezoelectric materials. The quality of the op-amp highly depends on the compensation. Direct compensation with Miller capacitor is the primary choice for many amplifier designs [12,14]. The capacitance of the load is very important while determining the value of the Miller capacitor. Principally, the capacitance of the load is not supposed to be much greater than the value of the Miller

capacitor. The step-response degrades with the increase in the Miller capacitor. When nested Miller compensation technique is used, the value of the capacitor used for the compensation significantly decreases. This phenomenon is an advantage when the system speed and die size are considered [15-18]. Especially in CMOS design, smaller capacitor is preferred since it occupies less area in the layout. When the stability becomes an issue for commercial op-amps, several external compensation techniques can be considered. One of the common external compensation techniques is using an integrator at the output by utilization of an external capacitor and a resistor. This technique is not convenient when the size, cost, and speed is a concern. To our knowledge, the proposed circuit design introduces nested Miller compensation for piezo drivers for the first time in the literature. The circuit involves nested Miller compensation with C1 and C2 capacitors to improve the stability of the output signal. As shown in Fig. 1, Miller and supporting compensation capacitors are 100pF and 50pF, respectively. The novelty of this research is the use of small capacitors for compensation of a system that has a capacitive loading of 47nF. Nested Miller compensation allows the use of small capacitance to reduce the oscillations and overshoot. The op-amp design has 0.5 V/ μ s slew rate and 1 MHz Gain Bandwidth Product (GBP). By tuning C1 and C2 capacitors as well as zero nulling resistor (R_{feed}) connected to the compensation circuitry, the op-amp can be modified to work for a variety of piezoelectric materials without significant degradation of the step-response.

Especially for large capacitive loads, control of oscillations at the output voltage becomes a crucial task to be examined for a universal piezo driver design. To determine the technical specifications and to understand the stability of the op-amp prototype, a number of measurements were conducted using Agilent 33220A Signal Generator, Tektronix TDS 2022B Oscilloscope, Agilent 34405A Multimeter, and Agilent E3648A DC Power Supply. R_{feed} value is maintained at 100k Ω during the measurements. Output voltage characteristics obtained from the prototype for various types of signals are shown in Fig. 2. Compensation at systems with multiple poles is frequently done by the domination of one of the poles [12]. Miller compensation determines a dominant pole via pole-zero cancellation [20]. When a capacitive load (47nF) is inserted into the output of the op-amp, oscillations are observed besides overshoots at the step response in the compensation configuration with a single Miller capacitor of 100pF (Fig. 2a, 2c, and 2e). Fig. 2b, 2d, and 2f show the results with nested Miller compensation. Nested Miller compensation with two capacitors brings additional zeros to the system. The output voltage characteristics show that Miller compensation using a single transistor may not be adequate to lower the oscillations. Instead of increasing the value of the Miller capacitor, we used nested Miller compensation capacitor (C2) not to affect step-response characteristics, and so the slew rate. In principle, the effect of the poles on the imaginary axis of the root locus must be eliminated by the zeros coming from the nested

compensation by the means of pole-zero cancellations. Basically, nested Miller compensation introduces complex zeros to the system. The roots causing undamped oscillation must be much greater than the roots causing underdamped oscillation. In an uncompensated system, the poles go into the right-hand plane (RHP) for higher gain and the system becomes unstable [15,20]. With the nested Miller compensation, purely imaginary poles do not affect the stability, and op-amp becomes stable at a high gain range. As also stated by Baker et al, the left-hand plane (LHP) zeros improve the speed of the op-amp [15].

III. COMPENSATION AND ZERO NULLING RESISTOR

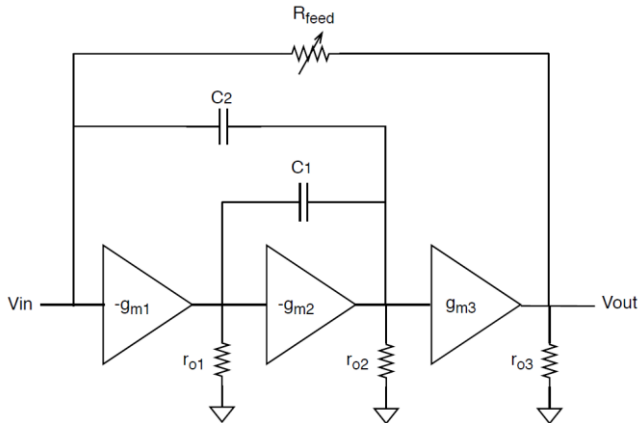


Fig. 3. Block Diagram of the proposed three-stage opamp.

Fig. 3 shows the symbolic block diagram of the proposed three-stage opamp. g_{m1} , g_{m2} , and g_{m3} refer to the transconductances of the stages, whereas corresponding output impedances are given by r_{o1} , r_{o2} , and r_{o3} . C_1 and C_2 forms the nested Miller compensation and R_{feed} is the zero nulling resistor. Small-signal analysis of the system given in Fig. 1 provides a straightforward approach to find the transfer function of the proposed three-stage opamp [16,17].

The transfer function of the opamp with nested Miller compensation configuration is given in Eq.1. where the voltage gain A_0 can be expressed by the following equation.

It is important to note that if the capacitive loading is taken into account in the transfer function, the expression in Eq. 1 will not change. Assuming two poles are widely separated, the denominator of $Av(s)$ is used to find out the poles via the equation given in Eq.3.

$$D(s) = 1 + \frac{s}{\omega_{p1}} + \frac{s^2}{\omega_{p1}\omega_{p2}} \quad (3)$$

Thus, from the transfer function, the dominant pole, ω_{p1} , is simplified to

$$\omega_{p1} = \frac{1}{g_{m2}C_1r_{o1}r_{o2} + C_1r_{o1} + C_1r_{o2} + g_{m3}C_2r_{o2}r_{o3} + C_2R_{feed} + C_2r_{o3} + C_2r_{o2}} \quad (4)$$

and in a more simplified form:

$$\omega_{p1} \cong \frac{1}{g_{m2}C_1r_{o1}r_{o2} + g_{m3}C_2r_{o2}r_{o3}} \quad (5)$$

The numerator of $Av(s)$ provides the function for the zero.

$$\omega_z = \left(C_2R_{feed} - \frac{C_1}{g_{m2}} - \frac{C_2}{g_{m3}} \right)^{-1} \quad (6)$$

RHP zero cancellation can be done via using R_{feed} resistor connected to both input and output.

$$R_{feed} = \frac{C_1}{C_2g_{m2}} + \frac{1}{g_{m3}} \quad (7)$$

When large capacitive loading is taken into account, zero becomes identical

$$\omega_z' = \left(C_L C_2 R_{feed} - \frac{C_L C_1}{g_{m2}} - \frac{C_L C_2}{g_{m3}} \right)^{-1} \quad (8)$$

Thus, R_{feed} value is independent of the loading but determined by C_1 and C_2 values as well as g_{m2} and g_{m3} . In a similar way with Mita et al., the relation between C_1 , C_2 and C_L is assumed as give below to achieve a reasonable phase margin.

$$C_2 C_L > C_1^2 \quad (9)$$

$$Av(s) = A_0 \frac{1 + \left(C_2 R_{feed} - \frac{C_1}{g_{m2}} - \frac{C_2}{g_{m3}} \right) s + \left(\frac{C_1 C_2}{g_{m2} g_{m3}} - \frac{C_1 C_2}{g_{m2}} \right) s^2}{1 + \left(g_{m2} C_1 r_{o1} r_{o2} + C_1 r_{o1} + C_1 r_{o2} + g_{m3} C_2 r_{o2} r_{o3} + C_2 R_{feed} + C_2 r_{o3} + C_2 r_{o2} \right) s + \left(g_{m2} C_1 C_2 r_{o1} r_{o2} R_{feed} + g_{m2} C_1 C_2 R_{feed} + g_{m3} C_1 C_2 r_{o1} r_{o2} r_{o3} + C_1 C_2 r_{o2} R_{feed} + C_1 C_2 r_{o2} r_{o3} + C_1 C_2 r_{o1} R_{feed} + C_1 C_2 r_{o1} r_{o3} + C_1 C_2 r_{o1} r_{o2} \right) s^2} \quad (1)$$

$$A_0 = g_{m1} g_{m2} g_{m3} r_{o1} r_{o2} r_{o3} \quad (2)$$

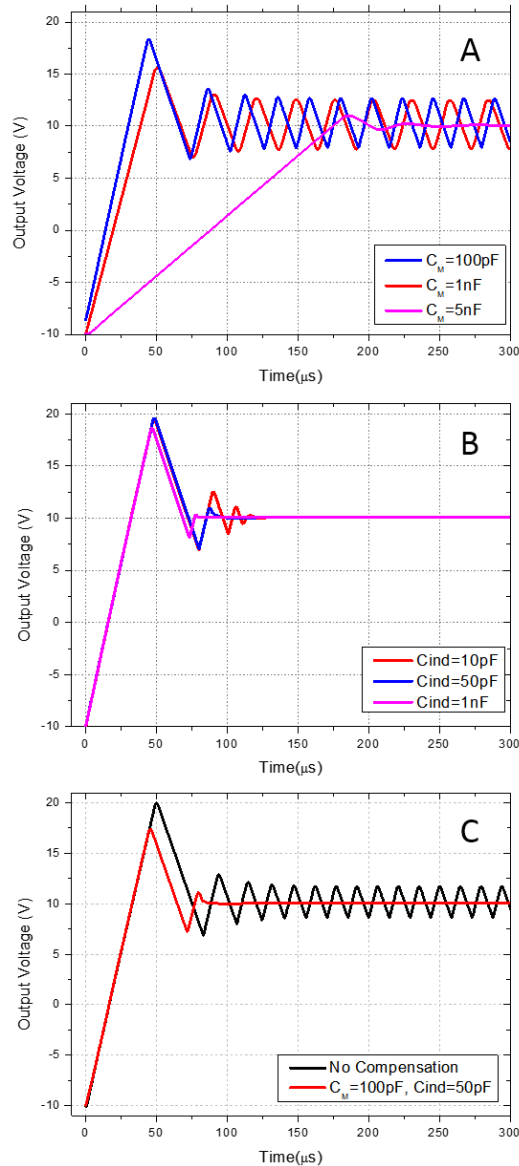


Fig. 4. Step response simulations of the op-amp with (a) only Miller capacitor (C_M) and (b) additional capacitor to form nested Miller (C_{ind}). (c) Comparison of a non-compensated op-amp with the nested Miller compensated using the same capacitance values with the prototype.

IV. SIMULATION RESULTS

The measurements collected for various scenario is confirmed via the simulations done in LTSpice using the models of the components used in the prototype (i.e. MMBTA42, MMBTA92, FZT 855, FZT955). In simulations, a 47nF capacitor is used for mimicking the piezoelectric material loading and R_{feed} value is kept at 100k Ω . Direct compensation with Miller capacitance is a very convenient method, however, the stability of the system can be ensured at capacitances very close to the loading capacitance. When the loading has large a capacitance, like piezoelectric materials, an adverse impact of using a single Miller capacitor on the

processing speed can be observed drastically, as shown in Fig.4a. Instability in the signal brings up mechanical vibrations associated with the resonance frequency.

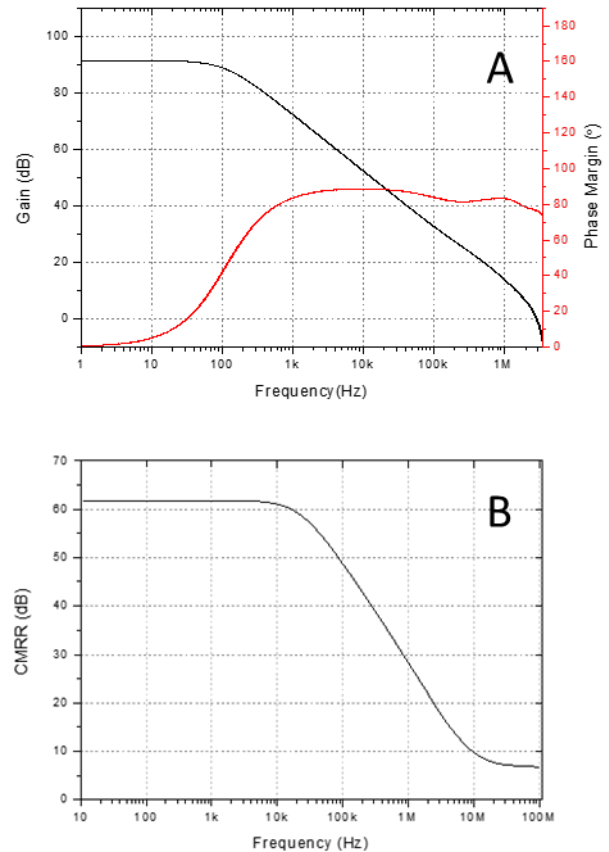


Fig. 5. a) Open loop gain vs. frequency and phase margin plots b) Common mode rejection ratio (CMRR) vs. frequency

Fig.4a illustrates step responses to exemplify the use of direct compensation with various values of Miller capacitors. With the increase at the value of Miller capacitors, overshoot settles down, however, the step response shows that the speed of the system decreases drastically and oscillations do not disappear with the use of small Miller capacitors. Fig. 4b shows the utilization of a supporting capacitor without the original Miller capacitor. Utilization of the supporting capacitor at the pF level effectively reduces the oscillations. As shown in Fig. 4b, insertion of an additional compensation capacitor eliminates a significant amount of oscillations, but increasing the value of the supporting capacitor from 50pF to 1nF does not make significant improvements in the overshoot. As compared with the Miller compensation, nested Miller compensation provides remarkable progress at the oscillation cancellation, but it does not remove the significant portion of the overshoot. Miller capacitor is very effective on the reduction of overshoot whereas nested Miller compensation is more advantageous for the step response. Consequently, combining two capacitors can be used to enhance the stability of the system, so step response and overshoot can be adequately compensated. Fig. 4c demonstrates how a non-compensated system is stabilized

using a single Miller capacitor and the configuration for the nested Miller compensation.

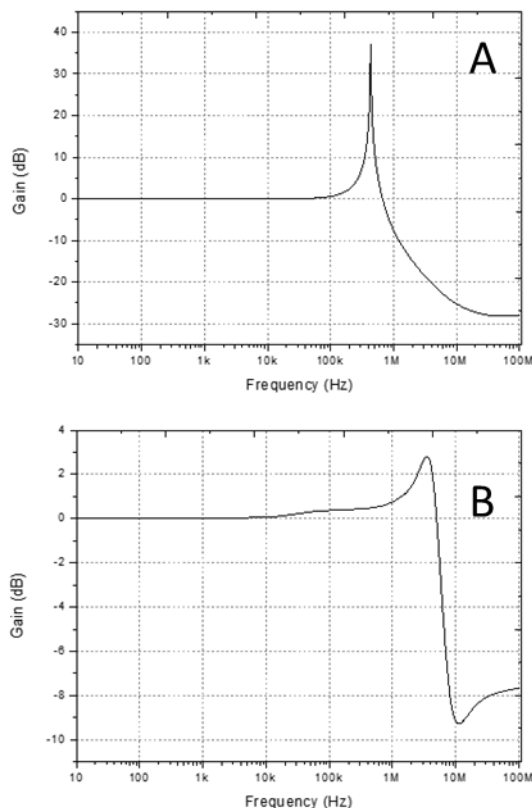


Fig. 6. Frequency responses at unity close loop gain a) only with the presence of Miller Capacitor ($C1=100\text{pF}$, No $C2$, $A_{VCL} = 1 \text{ V/V}$) b) with nested Miller compensation ($C1=100\text{pF}$, $C2=50\text{pF}$, $A_{VCL} = 1\text{V/V}$)

V. FREQUENCY RESPONSE CHARACTERISTICS

Bode plot recorded at open loop gain is shown in Fig. 5a. Accordingly, open loop gain value of 90 dB with 3MHz bandwidth and 77° phase margin is achieved in the design. The common mode rejection ratio (CMRR) is one of the parameters defining the quality of the op-amp. CMRR is calculated by the ratio of differential voltage gain to common mode voltage gain. The op-amp design given here has CMRR of 62dB, as shown in Fig. 5b. The plotted CMRR expresses the capability of differential-mode amplification whereas the common-mode voltage is excluded.

Fig. 6a shows the frequency response of the amplifier in case only Miller compensation is used ($C2$ removed from the circuit in Fig. 1). The plot suggests that the amplifier gradually becomes unstable after 100 kHz and reaches resonance frequency at around 400 kHz. At the resonance frequency, oscillations can be observed at the output voltage characteristics and this induces an excessive gain in the system. Resonance in the signal causes mechanical vibrations which result in undesired motions in the piezoelectric positioners, stages, and actuators. In case only 100pF Miller

capacitor is used, the recommended operating range of the op-amp would be limited to 100 kHz due to the resonance frequency. As the Miller capacitor is increased the stability of the system improves, but the capacitance value will be very large and it will have adverse impacts on the speed of the op-amp [15-18].

Impact of the nested compensation on resonance frequency is evaluated in detail. As shown in Fig. 6b, the resonance frequency is shifted to about 3.5 MHz when nested compensation is used. The excessive gain previously observed at 400kHz considerably decreases and becomes a tolerable value ($\sim 0.5\text{dB}$) for stable operation. As compared to the system with regular Miller compensation, bandwidth and stability are improved at the higher frequencies with the addition of a 50pF supporting capacitor.

VI. CONCLUSION

Piezoelectric positioners with high precision movement ability are of emerging technologies in metrology and instrumentation. Accurate control of the positioners in micro and nanoscale without mechanical resonance is only possible with highly stable op-amps. The three-stage op-amp design provided in this context introduces a feasible technique to build a fundamental part of the piezoelectric_positioner controllers. Without the requirement of a dedicated semiconductor manufacturing line, the op-amp can be produced using COTS components. The challenge with piezoelectric materials is their large capacitance and resonance frequency. The conventional direct (Miller) compensation may not bring out ultimate solutions to oscillations when the capacitance of the load is greater than Miller capacitance integrated into the circuit. Experimental results showed that the stability of the op-amp is improved by nested Miller compensation. In the prototype design, regular Miller capacitor reduces the overshoot whereas nested compensation is more beneficial to eliminate the oscillations while maintaining the step-response. Additionally, features for short circuit protection and temperature stability enable the use of designed opamp in extreme conditions. In conclusion, the rail-to-rail op-amp design covered in this study can be modified for the diverse use of piezoelectric materials in scientific instrumentation.

VII. ACKNOWLEDGMENT

This work is financially supported by Hacettepe University BAP under FAY-2017-14008.

VIII. REFERENCES

- [1] C. M. Dougherty, L. Xua, J. Pulskamp, S. Bedair, R. Polcawich, B. Morgan, and R. Bashirullah, "A 10V Fully-Integrated Switched-Mode Step-up Piezo Drive Stage in $0.13 \mu\text{m}$ CMOS Using Nested-Bootstrapped Switch Cells," *IEEE Journal of Solid-State Circuits*, 2016, vol. 51, pp. 1475-1486.
- [2] S. C. Doret, "Simple, low-noise piezo driver with feed-forward for broad tuning of external cavity diode lasers," *Review of Scientific Instruments*, 2018, vol. 89.

- [3] H. Ma, R. V. D. Zee, and B. Nauta, "A High-Voltage Class-D Power Amplifier With Switching Frequency Regulation for Improved High-Efficiency Output Power Range," *IEEE Journal of Solid-State Circuits*, 2015, vol. 50, pp. 1451-1462.
- [4] H. Tang and Y. Li, "Development and Active Disturbance Rejection Control of a Compliant Micro-/Nanopositioning Piezostage with Dual Mode," *IEEE Transactions on Industrial Electronics*, 2014, vol. 61, pp. 1475-1492.
- [5] S. Polit and J. Dong, "Development of a high-bandwidth XY nanopositioning stage for high-rate micro-/nanomanufacturing," *IEEE/ASME Transactions on Mechatronics*, 2011 vol. 16, pp. 724-733.
- [6] S. P. Wadikhaye, Y. K. Yong, and S. O. R. Moheimani, "A serial-kinematic nanopositioner for high-speed atomic force microscopy," *Review of Scientific Instruments*, 2014, vol. 85, pp. 105104(1)-105104(10).
- [7] L. T. Creagh and M. McDonald, "Design and Performance of Inkjet Print Heads for Non-Graphic-Arts Applications," *MRS Bulletin*, 2003, vol. 28, pp. 807-811.
- [8] A. Michael and C. Y. Kwok, "Piezoelectric micro-lens actuator," *Sensors and Actuators A-Physical*, 2015, vol. 236, pp. 116-129.
- [9] B. Oh, S. Oh, K. Lee, and M. Sunwoo, "Development of an injector driver for piezo actuated common rail injectors," in *SAE Technical Papers - 14th Asia Pacific Automotive Engineering Conference*, Hollywood, CA, 2007.
- [10] S. Baglio, G. Muscato, and N. Savalli, "Tactile measuring systems for the recognition of unknown surfaces," *IEEE Transactions on Instrumentation and Measurement*, 2002, vol. 51, pp. 522-531.
- [11] A. Robichaud, P.-V. Cicek, D. Deslandes, and F. Nabki, "Frequency Tuning Technique of Piezoelectric Ultrasonic Transducers for Ranging Applications," *Journal of Microelectromechanical Systems*, 2018, vol. 27, pp. 570-579.
- [12] B. Razavi, *Design of Analog CMOS Integrated Circuits*: McGraw-Hill, 2001.
- [13] A. S. Sedra and K. C. Smith, *Microelectronic Circuits*: Oxford University Press, 2016.
- [14] K. N. Leung and P. K. T. Mok, "Analysis of multistage amplifier-frequency compensation," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 2001, vol. 48, pp. 1041-1056.
- [15] V. Saxena and R. J. Baker, "Indirect feedback compensation of CMOS op-amps," 2006 IEEE Workshop on Microelectronics and Electron Devices, 2006 (WMED '06), 14 April 2006.
- [16] D. Marano, G. Palumbo, and S. Pennisi, "Step-Response Optimisation Techniques for Low-Power, Three-stage operational Amplifiers Driving Large Capacitive Loads" *IET Circuits, Devices and Systems*, 2010, vol. 4(2), pp. 87-98.
- [17] R. Mita, G. Palumbo, S. Pennisi, "Design Guidelines for Reversed Nested Miller Compensation in Three-Stage Amplifiers", *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 2003, vol. 50, pp. 227-233.
- [18] S. T. NGuyen and T. T. Bui, "A Design Procedure for Three-Stage Operational Amplifier Using Indirect Compensation Technique," presented at The 2014 International Conference on Advanced Technologies for Communications (ATC 2014), Hanoi, Vietnam, 2014.
- [19] S. Sengupta, K. Saurabh, and P. E. Allen, "A process, voltage, and temperature compensated CMOS constant current reference," in *Proceedings - IEEE International Symposium on Circuits and Systems-2004 IEEE International Symposium on Circuits and Systems - Proceedings*, Vancouver, 2004, pp. 1325-1328.
- [20] P. E. Allen and D. R. Holberg, *CMOS Analog Circuit Design*, 3rd ed. New York: Oxford University Press, 2012.
- [21] ON Semiconductor, "MPSA42 / MMBTA42 / PZTA42 NPN High-Voltage Amplifier," *MMBTA42 Datasheet*, Oct. 2014.

BIOGRAPHIES



M. Akif ÇELİK received his B.Sc. degree in Electrical and Electronics Engineering from Hacettepe University, Ankara, in 2018. He started his M.S. in the same department in 2019. He is currently working as a system design engineer in Aselsan. His research interests include analog design, scientific instrumentation, solid-state electronic devices and digital signal processing.



Dinçer GÖKCEN received his B.E. degree in electrical engineering from Yıldız Technical University in 2005 and his Ph.D. degree in electrical engineering from the University of Houston in 2010. He has been with the National Institute of Standards and Technology, GlobalFoundries, and Aselsan. In 2016, he joined Hacettepe University as an Assistant Professor. His research interests include nanoelectronics, nanofabrication, scientific instrumentation, quantum devices, and sensors.



F. Emre AYDOS received his B.Sc. degree in Electrical and Electronics Engineering from Hacettepe University, Ankara, in 2018. Afterward, he started his Master's Degree in the same department in 2019. He is currently working as a hardware design engineer in Aselsan. His research interests include hardware design, power electronics, and control systems.

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 2020
Volume: 8
Issue: 2

CONTENTS

N.H. Kaplan, I.Karabey Aksakalli, U. Kilic, I. Erer; A Novel Approach for Copy-move Forgery Detection using Bilateral Filtering.....	114-120
E. Belge, H. K. Kaba, A. Parlak, A. Altan, R. Hacıoğlu ; Estimation of Small Unmanned Aerial Vehicle Lateral Dynamic Model with System Identification Approaches,	121-126
M. Yılmaz, M. F. Çorapsız, M. R. Çorapsız; Voltage Control of Cuk Converter with PI and Fuzzy Logic Controller in Continuous Current Mode,.....	127-134
P. Arikan, S. Balci, F. Battal; Determination of the roll-off value in the air-gapped inductor of a DC-DC boost converter circuit with FEA parametric simulations,	135-141
O. Akgun, A. Akan and O. Yılmaz; The Spectral and Statistical Analysis of Gait Dynamics in ALS Disease,	142-146
Y. Sarica and M. Cetin; Artificial Intelligence Based Game Levelling,	147-153
S. Bilgili, A. K. Demir; Determination of the Ideal Protocol Stack for the Transmission of Health Data over 6LoWPAN IoT Networks.....	154-163
A. Dönük; Determination of Leakage Inductance Percentage for Gapped Iron-Core Shunt-Reactors with M4 Steel as Core Material,.....	164-169
A. Kaplan and E. Akagunduz; A Hybrid Framework for Matching Printing Design Files to Product Photos.....	170-180
AA. Abro, E. Taşci, A. Uğur; A Stacking-based Ensemble Learning Method for Outlier Detection,.....	181-185
M.A. Çelik, D. Gökçen, F.E. Aydos; Nested Miller Compensation Based Op-Amp Design for Piezoelectric Actuators,.....	186-192

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

