ISSN 1694-7398
journals.manas.edu.kg

June 2021
Volume 9, Issue 1

MANAS JOURNAL OF ENGINEERING

MJEN

BISHKEK 2021

KYRGYZ TURKISH MANAS UNIVERSITY
MANAS JOURNAL OF ENGINEERING (MJEN)

ISSN: 1694- 7398
Year: 2021
Volume: 9
Issue: 1
http://journals.manas.edu.kg
journals@manas.edu.kg

PUBLICATION PERIOD

Manas Journal of Engineering (MJEN) is published twice year, MJEN is a peer reviewed journal.

CORRESPONDENCE ADDRESS
Kyrgyz Turkish Manas University
Chyngyz Aitmatov Avenue 56 Bishkek, KYRGYZSTAN
URL: http://journals.manas.edu.kg
e-mail: journals@manas.edu.kg
Tel : +996 312 492763- Fax: +996 312 541935

KYRGYZ-TURKISH MANAS UNIVERSITY
MANAS JOURNAL OF ENGINEERING

CONTENT

MANAS Journal of Engineering
ISSN 1694-7398 | e-ISSN 1694-7398

Volume 9, Issue 1, (2021) Pages 1-9
https://doi.org/10.51354/mjen.802818

Improvement of Football Match Score Prediction by Selecting Effective Features for Italy Serie A League

Yavuz Selim Taspinar [1,*], Ilkay Cinar[2], Murat Koklu[3]

[1] Selcuk University, Doganhisar Vocational School, Department of Transport and Traffic Services, Konya, Turkey, ytaspinar@selcuk.edu.tr, ORCID: 0000-0002-7278-4241
[2]Selcuk University, Faculty of Technology, Department of Computer Engineering, Konya, Turkey, ilkay.cinar@selcuk.edu.tr, ORCID: 0000-0003-0611-3316
[3]Selcuk University, Faculty of Technology, Department of Computer Engineering, Konya, Turkey, mkoklu@selcuk.edu.tr, ORCID: 0000-0002-2737-2360

ABSTRACT

Football is one of the most popular sports in terms of number of fans in the world. This situation arises from the unpredictable nature of football. People are becoming more and more connected to this sport as it combines emotions such as excitement and joy that it creates in people. Match result prediction is a very challenging problem, and recently the solution to this problem has become very popular. With the result of this unpredictable game the events that occur during the match that affect this result are tried to be predicted by machine learning methods. This study demonstrates our work on finding the most effective features in match result prediction using match statistics from the Italian Serie A League's 2027 pieces match between the 2014-2015 and 2019-2020 seasons and with 54 features for each match. Feature selection testing was conducted to estimate the results of a football match and determine the most important factors. The selection of features was made using the ANOVA method and it was predicted that 28 of the 54 features would be effective in predicting match results. After this stage, fairly high rates classification success was achieved using the logistic regression method. 88.85% as a result of the prediction made with all features and 89.63% success was achieved as a result of the prediction made with 28 selected features. With these results, it is possible to say that process of feature selection increase success in match result prediction.

ARTICLE INFO

Research article

Received: 2.10.2020
Accepted: 19.04.2021

Keywords:
Soccer,
Football,
Sport Result Prediction,
Feature Selection,
Team Sport

**Corresponding author*

1 Introduction

In recent years, football continues to attract the attention of people from various age groups whose social and cultural status is different. In addition, it remains one of the sports with the largest number of spectators and fans worldwide. In football, outcome prediction is seen as a rather difficult problem because of the large number of factors that cannot be predicted and can affect results. In football, outcome prediction is seen as a rather difficult problem because of the large number of factors that cannot be predicted and can affect results. There are too many football teams at different levels in regional and national leagues in all countries of the world [1]. The success that can be achieved in a single match in football does not mean that it will be successful in other matches in the league. Also, a good team can sometimes be defeated against teams weaker than itself. Possession of the ball in the match, shots thrown to the goal, fouls, corner kick and many other factors that occur during the match affect the outcome of the match [2]. While it is difficult to predict results due to these emerging situations, people in academia and industry have sought to achieve positive results by conducting research on Football match prediction [3]. Various machine learning techniques and statistical methods are used to estimate the results and analyse the factors affecting the outcome of the match [4, 5]. Considering the history of the sport of football, thousands of matches have been played in the period up to the present day. A large number of statistical data about these matches can be accessed and used through sports sites on the internet.

A naive Bayesian method was used in a study to predict the results of Tottenham Hotspur football clubs for the period 1995-1997. They had noted that the Bayesian network outperforms other machine learning techniques, such as

nearest neighbors, decision trees. They had achieved an error rate of 40.79% by focusing on a specific team and a specific time period [6].

The study, which used Bayesian networks to predict the results of Spanish football team Barcelona's football matches, took into account weather conditions, the psychological state of the players and whether any of the main players had been injured. They had achieved a 92% accuracy rate in their study over 20 matches for a single season and a single team [7]. In another study, which used a regular probit regression model to predict the outcome of a football match, they proved that interesting factors such as the distance the away team travelled to the match had an effect on the outcome of the match. However, the study also had analysed the economic gains and price efficiency of the fixed-win betting market rather than match prediction [8, 9].

New approaches are being implemented to solve the secrets in football using artificial neural networks. In some studies, artificial neural networks derived from machine learning methods can even be said to be the best predictive model [10]. However, over the years, this claim has been refuted by developments in artificial neural networks. The outcome prediction was made with the help of images taken during match with convolutional neural networks from deep learning methods [11]. Deep neural networks have been used to predict football match outcomes in another study [12]. By using different match results in different leagues, match results were also estimated with the help of various machine learning methods. Match data of teams in the English Premier League, multiple linear regression, Artificial Neural Network (ANN), discriminant function analysis [13], Bayesian network, expert Bayesian network, decision tree, k-nearest neighbor [6] and ANN [14] methods were used in match result prediction.

Football match result estimation is a multi-class classification problem and in most studies the number of classes was taken as 2 or 3 classes. 2-class classifications home team won-away team won, 3-class classifications draw-home team won-away team won in the form classes were created. In a study with the collection of match data from various leagues, Long Short Term Memory Neural Network (LSTM NN) classification, LSTM NN regression methods were used and different results were obtained using different class numbers. They achieved 70.2% results in tests conducted using 2 classes and 52.5% results in tests conducted with 3 classes [15]. In the classifications for which 2-class match results were estimated, it was observed that they obtained higher results than the methods with 3-classification [16-17]. In the English Premier League, 69.5% success was achieved with a 2-class classification using 4 features with logistic regression method on 2280 pieces match data. When these studies in the literature are examined in detail, the number of match data in the leagues used, the number of features used and the method used are 3 factors that affect prediction success. When the numbers of data from matches in leagues are taken into account, there are also studies that use more than 200 thousand match data [18-20], as well as studies that use less than 100 match data [16, 6, 21]. The most important factor in predicting match results is the events that occur during the match. The features obtained from these unpredictable events are used as features in classification problems. Some of these features may be meaningless in match result predictions. These meaningless features are extracted from all features, meaningful data is evaluated and given as an introduction to classification algorithms. More than 100 features in literature studies [22], [15] and classifications made using only 4 features [23, 19] are available. This situation can vary for the data in each league and does not have a standard [1]. Trial and error or data simplification methods can be used to find the most effective features. Similar results were obtained in studies with different methods, feature numbers and 3 classes on the same data. 52.4% by XGBoost regression method using 66 features [18], 51.5% by Hybrid Bayesian Network method using 4 features [19], 51.9% by K nearest neighbour method using 8 features [20] achieved classification success. Based on this information, the feature selection phase of the data from the matches comes across to us as an extremely important factor for match result prediction.

In a study conducted between 1997 and 2003, tests were carried out using the multiple logistic regression method using Match records belonging to the Australian Football League. They achieved 66.7% accuracy in their work and stated that the key variables were the team's offensive strength, home advantage, distance travelled and ground recognition [24].

In another study, they used the Bayesian hierarchical model to predict the results of matches played in the Serie A League between 1991-1992. They showed that the most effective features when making predictions are home advantage, team attack and team defense variables [25]. In a similar study, the most effective features were shown to be attack and defender [26].

In another study, which used ANN and logistic regression, 95% success was achieved in match result prediction with match data from the English Premier League 2014-2015 season. They had shown that the most effective features in the classification were home and away teams, goals, shorts, corner, odds, attack strength, players' performance index, managers' performance index, managers' win, and teams' win streak [27]. In the literature, entropy, probability distributions and feature selection with different algorithms have been made [28].

This study, the Serie A League which is at the top level of the Italian football leagues and consists of 20 teams, uses football match data played between the 2014-2015 and 2019-2020 seasons to answer the question of which are the most effective features for predicting match results. A study was carried out

to determine the features affecting the match result using logistic regression method with datasets containing a total of 55 features including 2027 pieces football matches and one match result feature for each match. The dataset was created by scraping method on the web. With this dataset, tests were carried out by selecting the most effective features in the result estimation among 54 features that affect the outcome of the match using ANOVA.

The article is edited as follows. In the second part of the article provides information about the dataset used, the features of the dataset, and the methods we use. In Chapter 3, tests were carried out to determine effective features and estimate results. In section 4, performance analysis is given about the results obtained.

2 Material and Methods

Information about the data to be used in the study is given in this section. The data in the dataset is undergone various processes to be used in the study. These operations consist of putting together match data from the entire season and checking for missing data. After these operations, the selection of the effective features of each match was made. After the feature selection process, classification process was made for the prediction of the match result. The classification process is done separately for selected features and all features, and the test results was shown in Chapter 3. The introduction of the logistic regression method used in feature selection and classification is made in this section. The operations to be carried out in Figure 1 are shown with a flow diagram.

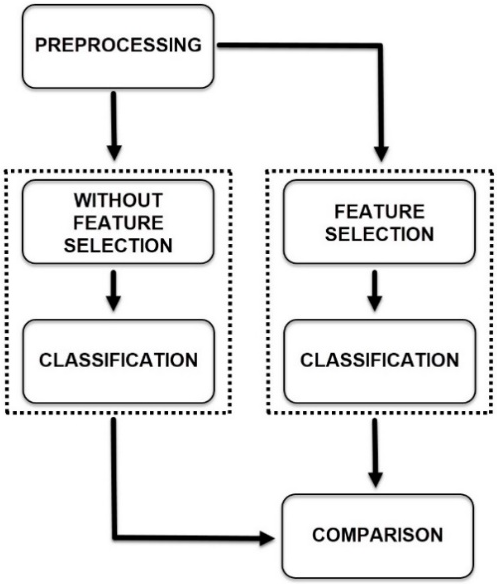


Figure 1. *Match score prediction flow diagram*

In this study, machine analysis and classification processes were carried out with the MATLAB program. A computer with Intel i5 10200H CPU, 8 GB Ram and GTX1650Ti graphics card was used.

2.1 Dataset

The dataset includes data from matches played in the Italian Serie A League at 40 weeks in the 2014-2015 season, 38 weeks in the 2015-2016 season, 41 weeks in the 2016-2017 season, 40 weeks in the 2017-2018 season, 41 weeks in the 2018-2019 season, and 23 weeks in the 2019-2020 season. The data were collected by scraping method on the web. Web Scraping method is a useful method for collecting data for use in researches [29]. Match data scoreboard.com from the website [30] was taken by this method. In the data set includes a total of 2027 pieces football matches, 54 features for each match, and a match result class with together these features. Table 1 contains list of features in the datasets.

Table 1. *Features of matches included in the datasets*

Features From Matches		
Possession Home	Crosses Home (%)	Headed Clearances Home
Possession Away	Crosses Away (%)	Headed Clearances Away
Total Shots Home	Dribbles Home (%)	Aerial Duels Home (%)
Total Shots Away	Dribbles Away (%)	Aerial Duels Away (%)
On Target Home	Corners Home	Blocked Crosses Home
On Target Away	Corners Away	Blocked Crosses Away
Off Target Home	Offsides Home	Saves Home
Off Target Away	Offsides Away	Saves Away
Blocked Home	Recoveries Home	Catches Home (%)

Blocked Away	Recoveries Away	Catches Away (%)
Passing Home (%)	Tackles Home (%)	Fouls Committed Home
Passing Away (%)	Tackles Away (%)	Fouls Committed Away
Attacking 3rd Home (%)	Interceptions Home	Fouls Won Home
Attacking 3rd Away (%)	Interceptions Away	Fouls Won Away
Key Passes Home	Blocks Home	Yellow Cards Home
Key Passes Away	Blocks Away	Yellow Cards Away
Clear Cut Chances Home	Clearances Home	Red Cards Home
Clear Cut Chances Away	Clearances Away	Red Cards Away
		Match Result

Besides the 54 features in the dataset match results are also an important factor in the outcome prediction. The ratios of Match Results for 2027 matches included in the dataset are shown in Figure 2. Draw: The end of the match with an equal score, Home Win: The result of the match is that the number of goals scored by the home team is greater, Away Win: The result of the match is that the number of goals scored by the away team is greater.

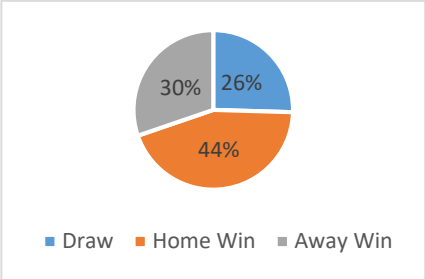


Figure 2. *Match score distribution*

2.2 Feature Selection

Feature selection is a totality of operations to select relevant features for solving a problem, to discard unnecessary ones and to increase the success of classification. A large number of data is being studied in order to increase classification accuracy. This is a big problem and it is quite difficult for algorithms to work with large data sets. Therefore, irrelevant features in the data are discarded and pre-processing steps are applied to reduce the number of features and the number of data. Thanks to the correct selection of features, learning speed can mostly be increased, as well as improvement in classification success according to the amount of data can be provided in a positive way [28].

ANOVA method of variance analysis was used for feature selection in the study. ANOVA is used to analyse how independent variables interact among themselves and the effects of these interactions on the dependent variable [31]. The dependent variables here are the features obtained from match statistics, while the independent variable is the match result.

2.3 Logistic Regression

Logistic regression is one of the statistical models used frequently in studies. In logistical regression, the dependent variable is estimated from one or more variables. Logistic regression clarifies the relationship between dependent variables and independent variables. In logistical regression, variables do not need to show normal distribution [32]. The values predicted in logistic regression are limited to 0 and 1 as they are probabilities. This is because logistical regression predicts the probability of outcomes, not the results itself [33].

2.4 ANOVA (Analysis of Variance)

ANOVA is a statistical analysis method used to study equalities over the values of more than two groups of features found in the datasets. It is used to compare the average values in small clusters formed by dividing the dataset into clusters by assigning variable labels to the values in the dataset [34].

2.5 Confusion Matrix

The measure of classification performance is measured by a confusion matrix, which records true and false recognized instances for each class [35]. The example confusion matrix for a two-class classifier is given in Table 2 [36].

Table 2. *Confusion matrix for binary classification*

		Predicted Class	
		P	**N**
Actual Class	**P**	TP	FN
	N	FP	TN

Elements of the Confusion matrix are expressed as TP: True

Positive, FN: False Negative, FP: False Positive, TN: True Negative counts. The 3-class confusion matrix used in the study is shown in Table 3.

Table 3. *3-Class confusion matrix (C: Class, T: True, F: False)*

		Predicted Class		
		C_1	C_2	C_3
Actual Class	C_1	T_1	F_{12}	F_{13}
	C_2	F_{21}	T_2	F_{23}
	C_3	F_{31}	F_{32}	T_3

Performance criteria for classification methods are provided with their formulas in Table 4 [37-39]. These performance criteria give percentile accuracy rates of the classification.

Table 4. *Performance criteria, formula, and evaluation condition table*

Performance Criteria	Formula
ACCURACY	$\frac{tp+tn}{tp+fp+tn+fn}*100$
PRECISION	$\frac{tp}{tp+fp}*100$
RECALL	$\frac{tp}{tp+fn}*100$
SPECIFICITY	$\frac{tn}{tn+fp}*100$
F1-SCORE	$2*\frac{precision*recall}{precision+recall}$

3 Experimental Results

The main aim in the experiments is to examine the contribution of these features to classification success by selecting the most effective features derived from match statistics, primarily for result prediction. In the study, logistic regression method was used for classification operations. The classification result obtained using all the features in the dataset was compared with the classification result obtained with the selected features. In this way, it is envisaged that faster classifications can be made in large data sets by looking at these effective features.

First, data pre-processing operations were performed to ensure that the data can be processed smoothly. After this process, the selection of features was realized with ANOVA. However, it was observed that only the features belonging to the home team or only the away team were selected among the selected features. To ensure data integrity, the features coloured in Table 4 are considered as selected features and added to the list. This situation is true for the datasets used and can result in different situations in different datasets. This is why every feature in the table is included for both teams. The features selected with ANOVA are shown in Table 5.

Table 5. *Selected features with ANOVA*

1	possession_home	15	clearances_home
2	possession_away	16	clearances_away
3	total_shots_home	17	headed_clearances_home
4	total_shots_away	18	headed_clearances_away
5	on_target_home	19	aerial_duels_home
6	on_target_away	20	aerial_duels_away
7	attacking_3rd_home	21	blocked_crosses_home
8	attacking_3rd_away	22	blocked_crosses_away
9	key_passes_home	23	saves_home
10	key_passes_away	24	saves_away
11	clear-cut_chances_home	25	yellow_cards_home
12	clear-cut_chances_away	26	yellow_cards_away
13	recoveries_home	27	red_cards_home
14	recoveries_away	28	red_cards_away

Among the selected features, graphs showing the differentiation of the 10 most effective features according to the result state are shown in Figure 3.

.

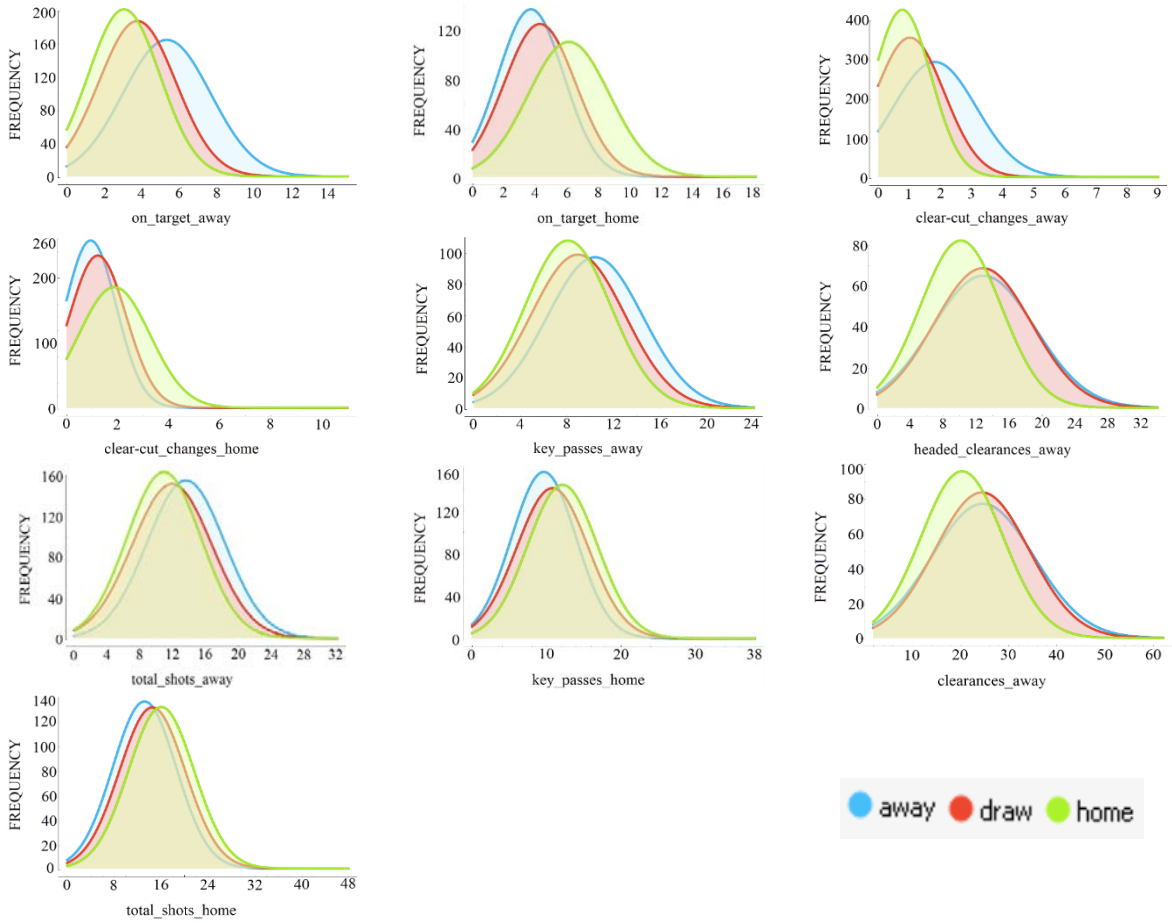


Figure 3. *Frequency distributions of selected top 10 features*

The repetition of values in selected features by considering the result state is an important factor for the classification problem. For example, the frequency of values in a feature is expected to differ with the frequency of result values, because the repetition of values in a feature at the same rate as the numbers of draws, home team and away team wins makes classification difficult. Non-effective features are eliminated with these situations in mind and can not be used in the classification phase.

A total of 2027 football matches were classified by logistic regression using data from the data set, which includes 54 features per match and the Match Result class with these features. The results obtained from this classification are shown in the confusion matrix in Table 6.

Table 6. *Confusion matrix for classification done using all features*

		Predicted Class		
		Draw	Home Win	Away Win
Actual Class	Draw	402	58	51
	Home Win	60	838	1
	Away Win	54	2	561

As a result of the calculations made with the data in Table 6, it was found that 77.90% of the draw status, 93.31% of the home team win, 91.51% of the away team win were correctly classified. Overall average classification success was found as 88.85%. The reason why draw state classification success is lower than classification success in the case of the home team or away team being the winner is that the draw state classification is difficult. The reason for this situation in the

event of a tie occurs because the statistics during the match are similar to other situations, and that 25.46% of 2027 pieces football match result in a draw.

Data simplification, in some cases can increase the success of classification and in some cases can decrease it. The high success of classification after the selection of important features in the dataset is related to the contribution of effective features and non-effective features to classification. However, it should also be forgotten that non-effective features may never affect the outcome of the match. Because the outcome of a football match and the events during the match is an unpredictable process. Features that are not effective apply to the dataset in our study.

Some of the effective features selected are not common features for either team. To ensure data integrity, these missing features were added within both sets and the number of selected features was increased to 28. These selected features were given as input to the logistic regression method and the outcome of the match, which was the output data, was tried to be estimated. The confusion matrix of the results obtained is shown in Table 7.

Table 7. *Confusion matrix for classification made using selected 28 features*

		Predicted Class		
		Draw	Home Win	Away Win
Actual Class	Draw	408	53	44
	Home Win	54	842	2
	Away Win	54	3	567

The success achieved as a result of classification with selected features is 89.63%. Based on this result, it can be said that non-effective features adversely affect classification success. Overall classification success showed a 0.78% increase in classification with selected features. Classification success was achieved by 79% for the result of the match resulting in a draw, 94% for the win of the home team and 92% for the win of the away team. The biggest increase in classification success was observed when the match ended in a draw. The reason for the increase in general classification success can be said to be the high percentage of draw status classification success. Because the draw situation is undesirable in match result prediction and negatively affects classification results. The success of classification by number of features is shown in Figure 4.

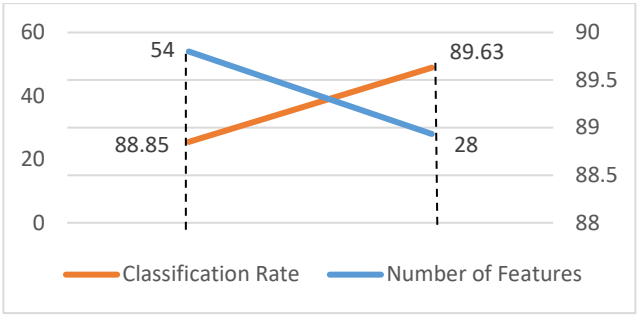


Figure 4. *Classification success by number of features*

4 Conclusion

In this study, the selection of effective features in the dataset was provided using data simplification methods. Imbalances can occur when certain features that do not contribute to classification are removed. For example, the yellow card feature of the home team is an effective feature in classification, while the yellow card feature of the away team may not be an effective feature. However, 3 features that are not effective in terms of data integrity (yellow card away, recoveries home, and block crosses home) were added and used as input data in the classification. As a result of the tests, the success rate was 88.85% in the classification performed using 54 features and the success rate was 89.63% in the classification performed with 28 features after the selection of features. As can be seen in other studies in the literature, in order to increase the success of classification, the result of the match can be treated as a 2-class. However, the handling of situations where only the home and away team are victorious does not coincide with real life. Of the 54 features we use in the classification, 26 are removed because they are non-contributing to the classification. It can be said that with the new features selected in the study carried out, it has been successful in the feature selection in terms of match result prediction success.

REFERENCES

[1] Bunker, R. and T. Susnjak, The Application of Machine Learning Techniques for Predicting Results in Team Sport: A Review. arXiv preprint arXiv:1912.11762, 2019.

[2] Zaveri, N., et al., Prediction of Football Match Score and Decision Making Process. International Journal on Recent and Innovation Trends in Computing and Communication, 2018. 6(2): p. 162-165.

[3] Chalikias, M., E. Kossieri, and P. Lalou, Football matches: Decision making in betting. Teaching Statistics, 2020.

[4] Samba, S., Football Result Prediction by Deep Learning Algorithms. 2019, Tilburg University.

[5] Barron, D., et al., Identifying playing talent in professional football using artificial neural networks. Journal of Sports Sciences, 2020: p. 1-10.

[6] Joseph, A., N.E. Fenton, and M. Neil, Predicting football results using Bayesian nets and other machine learning techniques. Knowledge-Based Systems, 2006. 19(7): p. 544-553.

[7] Owramipur, F., P. Eskandarian, and F.S. Mozneb, Football result prediction with Bayesian network in Spanish League-Barcelona team. International Journal of Computer Theory and Engineering, 2013. 5(5): p. 812.

[8] Goddard, J., Regression models for forecasting goals and match results in association football. International Journal of forecasting, 2005. 21(2): p. 331-340.

[9] Baboota, R. and H. Kaur, Predictive analysis and modelling football results using machine learning approach for English Premier League. International Journal of Forecasting, 2019. 35(2): p. 741-755.

[10] Schumaker, R.P., O.K. Solieman, and H. Chen, Sports data mining. Vol. 26. 2010: Springer Science & Business Media.

[11] Chen, M.-Y., T.-H. Chen, and S.-H. Lin, Using Convolutional Neural Networks to Forecast Sporting Event Results, in Deep Learning: Concepts and Architectures. 2020, Springer. p. 269-285.

[12] Rudrapal, D., et al., A Deep Learning Approach to Predict Football Match Result, in Computational Intelligence in Data Mining. 2020, Springer. p. 93-99.

[13] Reed, D. and P. O'Donoghue, Development and application of computer-based prediction methods. International Journal of Performance Analysis in Sport, 2005. 5(3): p. 12-28.

[14] McCabe, A. and J. Trevathan. Artificial intelligence in sports prediction. in Fifth International Conference on Information Technology: New Generations (itng 2008). 2008. IEEE.

[15] Danisik, N., P. Lacko, and M. Farkas. Football match prediction using players attributes. in 2018 World Symposium on Digital Intelligence for Systems and Machines (DISA). 2018. IEEE.

[16] Huang, K.-Y. and W.-L. Chang. A neural network method for prediction of 2006 world cup football game. in The 2010 international joint conference on neural networks (IJCNN). 2010. IEEE.

[17] Odachowski, K. and J. Grekow. Predicting the Final Result of Sporting Events Based on Changes in Bookmaker Odds. in KES. 2012.

[18] Hubáček, O., G. Šourek, and F. Železný, Learning to predict soccer results from relational data with gradient boosted trees. Machine Learning, 2019. 108(1): p. 29-47.

[19] Constantinou, A.C., Dolores: A model that predicts football match outcomes from all over the world. Machine Learning, 2019. 108(1): p. 49-75.

[20] Berrar, D., P. Lopes, and W. Dubitzky, Incorporating domain knowledge in machine learning for soccer outcome prediction. Machine Learning, 2019. 108(1): p. 97-126.

[21] Hucaljuk, J. and A. Rakipović. Predicting football scores using machine learning techniques. in 2011 Proceedings of the 34th International Convention MIPRO. 2011. IEEE.

[22] Odachowski, K. and J. Grekow. Using bookmaker odds to predict the final result of football matches. in International Conference on Knowledge-Based and Intelligent Information and Engineering Systems. 2012. Springer.

[23] Prasetio, D. Predicting football match results with logistic regression. in 2016 International Conference On Advanced Informatics: Concepts, Theory And Application (ICAICTA). 2016. IEEE.

[24] Bailey, M.J., Predicting sporting outcomes: A statistical approach. 2005, Faculty of Life and Social Sciences, Swinburne University of Technology.

[25] Baio, G. and M. Blangiardo, Bayesian hierarchical model for the prediction of football results. Journal of Applied Statistics, 2010. 37(2): p. 253-264.

[26] Min, B., et al., A compound framework for sports results prediction: A football case study. Knowledge-Based Systems, 2008. 21(7): p. 551-562.

[27] Igiri, C.P. and E.O. Nwachukwu, An improved prediction system for football a match result. IOSR Journal of Engineering (IOSRJEN), 2014. 4(12): p. 12-20.

[28] Bolón-Canedo, V., N. Sánchez-Maroño, and A. Alonso-Betanzos, Feature selection for high-dimensional data. 2015: Springer.

[29] Mitchell, R., Web scraping with Python: Collecting more data from the modern web. 2018: " O'Reilly Media, Inc.".

[30] scoreboard.com. 2020 [cited 2020 17 March]; Available from: www.scoreboard.com.

[31] Makris, S. and C. Urgesi, Neural underpinnings of superior action prediction abilities in soccer players. Social cognitive and affective neuroscience, 2015. 10(3): p. 342-351.

[32] Cruyff, M.J., et al., A review of regression procedures for randomized response data, including univariate and multivariate logistic regression, the proportional odds model and item response model, and self-protective responses, in Handbook of Statistics. 2016, Elsevier. p. 287-315.

[33] Kalantar, B., et al., Assessment of the effects of training data selection on the landslide susceptibility mapping: a comparison between support vector

machine (SVM), logistic regression (LR) and artificial neural networks (ANN). Geomatics, Natural Hazards and Risk, 2018. 9(1): p. 49-69.

[34] Pietraszek, J., et al. The fuzzy approach to assessment of ANOVA results. in International Conference on Computational Collective Intelligence. 2016. Springer.

[35] Sokolova, M., N. Japkowicz, and S. Szpakowicz. Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation. in Australasian joint conference on artificial intelligence. 2006. Springer.

[36] Luque, A., et al., The impact of class imbalance in classification performance metrics based on the binary confusion matrix. Pattern Recognition, 2019. 91: p. 216-231.

[37] Ozkan, I.A. and M. Koklu, Skin Lesion Classification using Machine Learning Algorithms. International Journal of Intelligent Systems and Applications in Engineering, 2017. 5(4): p. 285-289.

[38] Çinarer, G., Emiroğlu, B. G., & Yurttakal, A. H. (2020). Prediction of Glioma Grades Using Deep Learning with Wavelet Radiomic Features. Applied Sciences, 10(18), 6296.

[39] Karaşahin, A. T., & Tümer, A. E. Real time traffic signal timing approach based on artificial neural network. MANAS Journal of Engineering, 8(1), 49-54.

MANAS Journal of Engineering
ISSN 1694-7398 | e-ISSN 1694-7398

Volume 9, Issue 1, (2021) Pages 10-14
https://doi.org/10.51354/mjen.803396

Cohomology of semi-invariant submanifolds of cosymplectic manifolds

Ramazan Sarı

Amasya University, Gümüşhaciköy Hasan Duman Vocational School, Turkey, ramazan.sari@amasya.edu.tr, ORCID: 0000-0002-4618-8243

ABSTRACT

In this paper, we study de Rham cohomology class for semi-invariant submanifolds of a cosymplectic manifold. We show that there are de Rham cohomolgy class on semi-invariant submanifold of a cosymplectic manifold. Firstly, we define semi-invariant submanifolds of a cosyplectic manifold. We present an example for semi-invariant submanifold of a cosymplectic manifold.Later, We obtain characterizations, investigate the geometry of distributions which arise from the definition of semi-invariant submanifold. We obtain that invariant distribtion is always integrable and minimal. Moreover, necessary and sufficient conditions investigate for the anti-invariant distribution to be integrable and minimal. Finally, we prove that semi-invariant submanifold of a cosymplectic manifold has nontrivial de Rham cohomology class. Further, the theoretical methodology of mathematics are used to obtain results.

ARTICLE INFO

Research article

Received: 1.10.2020
Accepted: 28.11.2020

Keywords:
Cohomology class, cosymplectic manifold, semi-invariant submanifold

1 Introduction

Cohomology groups have an important studying area for a topological manifold. If a topological space M is a manifold, we may define the dual of the cohomology groups out of differential forms defined on M. The dual groups are called the de Rham cohomology groups. Besides physicists' familiarity with differential forms, cohomology groups have several advantages over homology groups [12].

Contact geometry has a very important place in physical and other mathematical structure. Really, this structures studied thermodynamics, geometric optics and in Hamiltonian dynamics. In these days, contact structures have obtain low dimensional topology. Contact structures first appeared on partial differential equations. Gray defined an almost contact manifold by the condition that the structural group of the tangent bundle is reducible to $U(n) \times 1$. After Sasaki studied an almost contact manifold with tensor fields and Riemannian metric [14]. Later many author studied different contact structures [9, 19]. Goldberg and Yano defined and studied cosymplectic manifolds [8]. A cosymplectic manifold can be considered as an odd-dimensional analogue of a Kaehler manifold.

Bejancu defines and study CR-submanifold which generalized invariant manifold and anti invariant manifold [1]. Later, this submanifolds have been developed different type structure [10, 13, 15]. Tripathi investigated semi-invariant submanifolds of LP-cosymplectic manifold [18]. In [5], Dirik studied contact CR- submanifolds of cosymplectic manifold.

Tanno investigated topology of contact Riemannian manifold [17]. He studied the basic topological properties of contact manifolds. Fernandez and Ibanez studied de Rham cohomologies on almost contact manifolds [6]. They investigated the relation of the coeffective cohomology of some classes of almost contact manifolds with the topology of the manifold. Chinea et al. introduced topology of cosymlectic manifold [3]. Montano et al. introduced topology of 3-cosymplectic manifolds [11]. They showed that there is an action of the Lie algebra on the basic cohomology spaces of a compact 3-cosymplectic manifold with respect to the Reeb foliation.

Chen introduced cohomology of CR-submanifold [2]. He proved that there are de Rham cohomolgy class on CR-submanifold of a Kaehler manifold. Moreover, he show that this class nontrivial such that invariant distribution and anti-invariant distribution are integrable and minimal, respectively. Later, Deshmukh and Ghazal studied cohomolgy of CR-submanifold nearly Kaehler and quasi Kaehler, respectively [4,7]. In [16], Şahin obtained cohomolgy of hemi-slant submanifoldof a Kaehler manifold.
In this paper, we study de Rham cohomology of semi-invariant submanifold of cosymlectic manifold. We obtain

that there are de Rham cohomolgy class on a semi-invariat submanifold under certain conditions.

2 Semi-ınvariant submanfolds of cosymplectic manifold

Let M be an $n-$dimensional real differentiable manifolds of differentiability class C^∞ endowed with a C^∞ vector valued linear function φ, a C^∞ vector field ξ, 1-form η and Riemannian metric g , which satisfies

$$\varphi^2 = -I + \eta \otimes \xi \quad \text{and } \eta(\xi) = 1 \tag{1}$$
$$g(\varphi X, \varphi Y) = g(X,Y) - \eta(X)\eta(Y) \tag{2}$$

for all $X, Y \in \Gamma(TM)$. Then, M said to be contact manifold. Also in contact manifold the following relations hold:

$$\varphi\xi = 0, \quad \eta o \varphi = 0, \quad rank(\varphi) = n - 1$$

and

$$g(\varphi X, Y) = -g(X, \varphi Y).$$

A contact manifold M is called cosymplectic manifold if

$$(\nabla_X \varphi)Y = 0 \tag{3}$$

for all $X, Y \in \Gamma(TM)$.

Definition 2.1. An $(2m+1)-$dimensional Riemannian submanifold B of a cosymplectic manifold M is called a semi-invariant submanifold there exists on B two differentiable orthogonal distributions D_T and $D^\perp$ satisfying:

1. $TB = D_T \oplus D^\perp \oplus sp\{\xi\}$;

2. The distribution D_T is invariant under φ, such that $\varphi D_{T(x)} = D_{T(x)}$ for all $x \in B$;

3. The distribution $D^\perp$ is anti-invariant under φ, such that $\varphi D_x^\perp \subseteq T_x^\perp M$ for any $x \in B$, where $T_x B$ and $T_x B^\perp$ are the tangent space of B at x.

Example 2.2. In what follows, $(\mathbb{R}^{2m+1}, \varphi, \eta, \xi, g)$ will denote the manifold $\mathbb{R}^{2m+1}$ with its usual cosymplectic structure given by

$$\eta = dz, \qquad \xi = \frac{\partial}{\partial z}$$

$$\varphi\left(\sum_{i=1}^{n} \left(X_i \frac{\partial}{\partial x_i} + Y_i \frac{\partial}{\partial y_i}\right) + Z\frac{\partial}{\partial z}\right) = \sum_{i=1}^{n} \left(Y_i \frac{\partial}{\partial x_i} - X_i \frac{\partial}{\partial y_i}\right) + Y_i y_i \frac{\partial}{\partial z}$$

$$g = \left(\sum_{i=1}^{n} dx_i \otimes dx_i + dy_i \otimes dy_i\right) - \eta \otimes \eta$$

$(x_1, \ldots, x_n, y_1, \ldots, y_n, z)$ representing the cartesian coordinates on $\mathbb{R}^{2m+1}$. We consider a submanifold of $\mathbb{R}^7$ defined by

$$M = X(k, f, l, w, t) = (k, 0, l, f, w, 0, t).$$

Therefore a basis of TM

$$e_1 = \frac{\partial}{\partial x_1}, \qquad e_2 = \frac{\partial}{\partial y_1}, \qquad e_3 = \frac{\partial}{\partial x_3},$$
$$e_4 = \frac{\partial}{\partial y_2}, \qquad e_5 = \frac{\partial}{\partial z} = \xi$$

Moreover,

$$e_1^* = \frac{\partial}{\partial x_2}, \qquad e_2^* = \frac{\partial}{\partial y_3}$$

from a basis of $T^\perp M$.
We determine

$$D_1 = sp\{e_1, e_2\}$$

and

$$D_2 = sp\{e_3, e_4\},$$

then D_1 is invariant distribution and D_2 is anti-invariant distribution. Then

$$TM = D_1 \oplus D_2 \oplus sp\{\xi\}$$

is a semi-invariant submanifold of $\mathbb{R}^7$.
On the other hand, let $\mathcal{V}$ be a differentiable distribution on a Riemannian manifold M with Levi civita connection ∇. We determine, for all $X, Y \in \Gamma(\mathcal{V})$,

$$\sigma(X,Y) = (\nabla_X^M Y)^\perp$$

where $(\nabla_X Y)^\perp$ denotes the component of $\nabla_X Y$ in the orthogonal complementary distribution $\mathcal{V}$ in M. Let $\{E_1, E_2, \ldots, E_p\}$ be an orthonormal frame of $\mathcal{V}$. We determine

$$H = \frac{1}{p}\sum_{j=1}^{p} \sigma(E_j, E_j).$$

Therefore H is well defined vector field on M. If $H = 0$ identically on M, we said to be $\mathcal{V}$ as minimal distribution.

3 Some basic result

Let B be a submanifold of a contact manifold M. Let the induced metric on M also be denoted by g. Then Gauss and Weingarten formulate are given respectively by

$$\nabla_X^M Y = \nabla_X^B Y + h(X,Y) \quad (4)$$
$$\nabla_X^M V = \nabla_X^{B\perp} V - A_V X \quad (5)$$

for all $X, Y \in \Gamma(TM)$ and $V \in \Gamma(TM^\perp)$, where A_V is the Weingarten endomorphism associated with V, $\nabla^{B\perp}$ is the connection in the normal bundle and h is the second fundamental from of M.
The shape operator A and the second fundamental form h related by

$$g(A_V X, Y) = g(h(X,Y), V). \quad (6)$$

Let B be a submanifold of a contact manifold M with contact structure (φ, η, ξ, g). For $X \in \Gamma(TB)$ we put

$$\varphi X = TX + NX \quad (7)$$

where TX and NX denote the tangential and normal components of φX respectively.
For $V \in \Gamma(TB^\perp)$ we put

$$\varphi V = tV + nV \quad (8)$$

where tV and nV denote the tangential and normal components of φV respectively.

Proposition 3.1. For a submanifold B of a contact manifold and $X \in \Gamma(TB), V, K \in \Gamma(TB^\perp)$, we have

$$g(X, TY) = -g(TX, Y), g(X, tV) = -g(tX, V)$$

and

$$g(K, nV) = -g(nK, V).$$

Proposition 3.2. For a submanifold B of a contact manifold and $\xi \in \Gamma(TB)$, we have

$$T\xi = 0 = N\xi, \eta \circ T = 0 = \eta \circ N$$

$$T^2 + tN = I + \eta \otimes \xi, \qquad NT + nN = 0$$

$$n^2 + Nt = I, tf + Tt = 0.$$

4 Cohomology class of semi-ınvariant submanifolds

In this section, we introduce de Rham cohomology class on semi-invariant submanifold of cosymlectic manifold. Firstly, we prove the following useful lemmas.

Lemma 4.1. Let B be a semi-invariant submanifold of cosymlectic manifold M. Therefore the distribution D_T is always integrable.

Proof. For all $U, V \in \Gamma(D_T)$ and $K \in \Gamma(D^\perp)$, using (1) and (2) we have

$$g([U,V], K) = g(\nabla_U^M \varphi V, \varphi K) - g(\nabla_V^M \varphi U, \varphi K).$$

Then, by virtue of (4), we arrive,

$$g([U,V], K) = g(h(U, \varphi V) - h(V, \varphi U), \varphi K)$$

which gives our assertion.

Lemma 4.2. Let B be a semi-invariant submanifold of cosymlectic manifold M. Therefore the distribution D_T is minimal.

Proof. Firstly, for all $U \in \Gamma(D_T)$ and $K \in \Gamma(D^\perp)$, we get

$$g(U, K) = 0.$$

Then for any, $W \in \Gamma(D_T)$, we arrive,

$$g(\nabla_W^M U, K) = g(\nabla_W^M V, K). \quad (9)$$

Therefore, using (2), (3) and (9) ,we have,

$$g(\nabla_U^M U, K) = -g(\nabla_U^M \varphi K, \varphi U).$$

Hence, from (5), we get,

$$g(\nabla_U^B U, K) = g(A_{\varphi K} U, \varphi U). \quad (10)$$

Moreover, using (2) and (3), we arrive,

$$g(\nabla_{\varphi U}^M \varphi U, K) = g(\nabla_{\varphi U}^M \varphi^2 U, \varphi K).$$

Then by virtue of (1), (9) and (5) we have,

$$g(\nabla_{\varphi U}^B \varphi U, K) = -g(A_{\varphi K} U, \varphi U). \quad (11)$$

(10) and (11) we arrive,

$$g(\nabla_U^B U + \nabla_{\varphi U}^B \varphi U, K) = 0. \quad (12)$$

Let $\{E_1, \ldots, E_q, \varphi E_1, \ldots, \varphi E_q\}$ be a ortonormal base of D_T. Then,

$$H = \frac{1}{2q}\sum_{j=1}^{q}\{\sigma(E_j,E_j)+\sigma(\varphi E_j,\varphi E_j)\}$$
$$= \frac{1}{2q}\sum_{j=1}^{q}\{(\nabla^B_{E_j}E_j)^{\perp}+(\nabla^B_{\varphi E_j}\varphi E_j)^{\perp}\}$$

By virtue of $g((\nabla^B_U W)^{\perp},K)=g(\nabla^B_U W,K)$, using (12) we have,

$$g(H,K)=0$$

which completed that proof.

Lemma 4.3. Let B be a semi-invariant submanifold of cosymlectic manifold M. Therefore the distribution $D^{\perp}$ is integrable if and only if

$$TA_{\varphi L}K = TA_{\varphi K}L$$

for all $K,L\in\Gamma(D^{\perp})$.

Proof. For all $K,L\in\Gamma(D^{\perp})$ and $U\in\Gamma(D_T)$, using (1) and (2) we have,

$$g([K,L],U)=g(\nabla^M_K\varphi L,\varphi U)-g(\nabla^M_L\varphi K,\varphi U).$$

From (5), we get,

$$g([K,L],U)=g(-A_{\varphi L}K,\varphi U)-g(-A_{\varphi K}L,\varphi U).$$

Finally, by virtue of (7) and (8), we arrive,

$$g([K,L],U)=g(TA_{\varphi L}K-TA_{\varphi K}L,U)$$

which completes proof.

Lemma 4.4. Let B be a semi-invariant submanifold of cosymlectic manifold M. Therefore the distribution $D^{\perp}$ is minimal if and only if

$$g(h(K,TU),NK)=g(\nabla^{\perp}_K\varphi K,NU)$$

for all $K\in\Gamma(D^{\perp})$ and $U\in\Gamma(D_T)$.

Proof. For all $K\in\Gamma(D^{\perp})$ and $U\in\Gamma(D_T)$, from (9), (1), (2) and (7), we have,

$$g(\nabla_K K,U)=g(\nabla_K\varphi K,TU)+g(\nabla_K\varphi K,NU).$$

By virtue of (5), we get,

$$g(\nabla_K K,U)=-g(A_{\varphi K}K,TU)+g(\nabla^{\perp}_K\varphi K,NU).$$

Then, using (6) and (7), we arrive,

$$g(\nabla_K K,U)=-g(h(K,TU),NK)+g(\nabla^{\perp}_K\varphi K,NU)$$

which gives our assertion.

Now, we denote an orthonomal frame $\{E_1,\dots,E_q,\varphi E_1,\dots,\varphi E_q\}$ of the distribution D_T. Let $\{w^1,\dots,w^q,w^{q+1},\dots,w^{2q}\}$ be the 1-forms on B satisfiying

$$w^i(K)=0,\qquad i\in\{1,\dots,2q\},$$
$$w^i(E_j)=\delta_{ij},i,j\in\{1,\dots,q\},\qquad (13)$$
$$w^k(\varphi E_j)=\delta_{kj},\qquad k\in\{q+1,\dots,2q\},\qquad ,j\in\{1,\dots,q\}$$

for all $K\in\Gamma(D^{\perp})$. Therefore, we arrive

$$w=w^1\wedge\dots\wedge w^{2q}.\qquad (14)$$

Hence w defines a $2q$-form on submanifold B.

Theorem 4.4. Let B be a closed semi-invariant submanifold of a cosymlectic manifold M . Therefore the $2q-$form w defines a canonical de Rham cohomology class given by

$$c(B)=[w]\in H^{2q}(B,\mathbb{R}),dimD_T=2q.$$

Moreover $c(B)$ is non-trivial if D_T is integrable and $D^{\perp}$ is minimal.

Proof. Firstly, using (14), we arrive,

$$dw=\sum_{k=1}^{2q}(-1)^k w^1\wedge\dots\wedge w^{2q}.$$

By virtue of (13), for all $U_1,\dots,U_{2q}\in\Gamma(D_T)$ and $K,L\in\Gamma(D^{\perp})$, we show that $dw=0$ if and only if

$$dw=(K,U_1,\dots,U_{2q})=0\qquad (15)$$

and

$$dw=(K,L,U_1,\dots,U_{2q})=0.\qquad (16)$$

Hence, $D^{\perp}$ must be integrable for (15) equality to occur and D_T must be minimal for (16) equality to occur. But two conditions always exist for semi-invariant submanifolds of cosymplectic manifold . Accordingly, w is closed form on M. Therefore, w defines a de Rham cohomolgy class $c(B)$ such that

$$c(B)=[w]\in H^{2q}(B,\mathbb{R}).$$

On the other hand, we denote $\{E_{2q+1},\dots,E_{2q+p}\}$ and $\{w^{2q+1},\dots,w^{2q+p}\}$ an orthonormal frame and dual frame of

D_T, respectively. Let $w^* = w^{2q+1} \wedge \ldots \wedge w^{2q+p}$ be p-form on M. Therefore similarly way for w, we can say that, $D^{\perp}$ is minimal and D_T is integrable, then w^* is closed, hence $2q$ -form w is coclosed. We know that, B is closed submanifold, then w is harmonic. Since w is nontrivial, the cohomology class $[w]$ characterize by w is nontrivial in $H^{2q}(B, \mathbb{R})$.

5 Discussion and conclusion

Contact geometry has an important application for many sciences such as physics, geometric optics, technology, thermodynamics, classical mechanics, medical sciences and classical mechanics. Researchers have increased studies on this field from different areas in recent years. The improvement of the contact geometry depends on the differential geometry of the manifolds with structures. Another subclass of contact geometry is the cosymplectic manifolds. Topology of cosymplectic manifolds is less explored, and there is a shortlist of papers in the mathematical literature on this topic. The works on this subject will be useful tools for the applications for topological of the cosymplectic manifolds. Hence, we studied de Rham cohomology class for semi-invariant submanifolds of cosymplectic manifolds. Consequently, the results obtained in this article provide contribution to investigate topological properties of different submanifolds in cosymplectic manifolds.

References

[1] Bejancu A., “Geometry of CR-submanifolds”, Dordrecht, the Netherlands, D. Reidel; 1986.

[2] Chen B.Y., “Cohomology of CR-submanifolds”, Annales Faculte des Sciences Toulouse, 3, (1981), 167-172.

[3] Chinea D., De Leon M. Marrero J.C., “Topology of cosymplectic manifolds”, J.Math. Pures Appl. 72, (1993), 567-591.

[4] Deshmukh S., “Cohomology of CR-submanifolds of a nearly Kaehler manifold”, Math. Chronicle, 16, (1982), 47-57.

[5] Dirik S., “On Contact CR-Submanifolds of a Cosymplectic Manifold”, Filomat, 32(13), (2018), 4787-4801.

[6] Fernandez M. Ibanez R., “Coeffective and de Rham cohomologies on almost contact manifolds”, Dif. Geo. and its App., 8, (1998), 285-303.

[7] Ghazal T., “Cohomology of CR-submanifolds of quasi-Kaehler manifolds”, Int. J. Pure Appl. Math., 52(5), (2009), 711-715.

[8] Goldberg S. I. Yano K., “Integrability of almost cosymplectic structure”, Pacific J. Math., 31, (1969), 373-382.

[9] Kenmotsu K., “A class of almost contact Riemannian manifolds”, Tohoku Math. J. 24 (1972), 93-103.

[10] Ludden G., “Submanifolds of cosymplectic manifold”, J.Differential Geometry, 4, (1970), 23.

[11] Montano B.C., Nicola A.D. Yudin I., “Topology of 3-cosymplectic manifolds”, Quarterly Journal of Mathematics, 64(1), (2013), 59-82.

[12] Nakahara M., “Geometry, topology and physics”, IOP Publishing, Bristol 1990.

[13] Sari R., Ünal İ. Aksoy Sari E., “Skew Semi Invariant Submanifolds of Para Kenmotsu Manifold”, Gümüşhane University Journal of Science and Technology Institute, 8, (2018), 112-118.

[14] Sasaki S., “On differentiable manifolds with certain structures which are closely related to almost contact structure”, I, TShoku Math. J., 12, (1960), 459-476.

[15] Shoeb M., Shahid M. H. Sharfuddin A., “On Submanifolds of a cosymplectic Manifold”, Soochow Journal of Math., 27(2), (2001),161-174.

[16] Şahin F., “Cohomology of hemi-slant submanifolds of a Kaehler manifold”, Journal of Advanced Studies in Topology, 5(2), (2014), 27-31.

[17] Tanno S., “The topology of contact Riemannian manifolds”, Illinois J. Math., 12(4), (1968), 700-717.

[18] Tripathi M.M., “Semi-invariant submanifold of LP-cosymlectic manifolds”, Bull. Malaysian Math. Sc. Soc., 24, (2001), 69-79.

[19] Vanli A., Ünal İ. Avcu K., “On complex Sasakian manifold”, African Mathematika, 32, (2021), 507-516.

MANAS Journal of Engineering
ISSN 1694-7398 | e-ISSN 1694-7398

Volume 9, Issue 1, (2021) Pages 15-21
https://doi.org/10.51354/mjen.822630

A TensorFlow implementation of Local Binary Patterns Transform

Devrim Akgun

Sakarya University, Software Engineering Department, 54187, Sakarya, Turkey, dakgun@sakarya.edu.tr, ORCID: 0000-0002-0770-599X

ABSTRACT

Feature extraction layers like Local Binary Patterns (LBP) transform can be very useful for improving the accuracy of machine learning and deep learning models depending on the problem type. Direct implementations of such layers in Python may result in long running times, and training a computer vision model may be delayed significantly. For this purpose, TensorFlow framework enables developing accelerated custom operations based on the existing operations which already have support for accelerated hardware such as multicore CPU and GPU. In this study, LBP transform which is used for feature extraction in various applications, was implemented based on TensorFlow operations. The evaluations were done using both standard Python operations and TensorFlow library for performance comparisons. The experiments were realized using images in various dimensions and various batch sizes. Numerical results show that algorithm based on TensorFlow operations provides good acceleration rates over Python runs. The implementation of LBP can be used for the accelerated computing for various feature extraction purposes including machine learning as well as in deep learning applications.

ARTICLE INFO

Research article

Received: 6.11.2020
Accepted: 28.06.2021

Keywords:
tensorflow,
local binary patterns,
deep learning,
feature extraction

1. Introduction

Intense computations in Python may take considerably longer time when compared with the programming languages such as C/C++. This is because Python is an interpreter based language where the input script is interpreted line by line. On the other hand Python provides a high level abstraction which makes most of the scientific computations and especially machine learning and deep learning applications simple to implement. Although Python realization of these algorithms usually take long running times, C/C++ compiled Python functions accelerates the computations significantly. For this purpose TensorFlow provides an open source framework for machine learning applications [1], [2]. TensorFlow also enables programmers to run their codes on GPU (Graphics Processing Unit) and TPU (Tensor processing Unit) as well as on CPU. Hence utilization of GPU resources and hardware accelerators in addition to CPU provides significant accelerations [3]. TensorFlow library includes most of the functions and layers for developing and training deep learning models. In addition, it has the flexibility of custom layers which enable programmers to design their own layers [4]. Users may describe custom functions based on the existing operations as well as writing the functions from scratch.

In deep learning applications, determining the layers of deep neural networks have importance in developing a successful model. Deep learning layers have the ability to extract features from the dataset automatically, and different layers may provide the potential to extract better features. Feature extraction layers such as convolutions and pooling in deep learning enable to automatic extraction of desired features. Although TensorFlow and Keras cover most of the frequently used layers, additional layers may increase the accuracy depending on the problem's nature. Various authors use custom layers as a combination of existing layers or their own developed layers for specific purposes such as new activation function definitions for medical diagnostic [5], wavelet-based pooling [6], solution of inverse partial differential equation [7], radial basis functions for adaptive routing problems [8].

One of the feature extraction methods is LBP transform which is widely used in machine learning and deep learning applications in addition to image processing [9], [10]. In literature there are numerous utilizations of LBP in various computer vision applications such as handwritten text recognition [11], facial expression recognition for smart applications [12], ear recognition for identity verification [13], Retrieval of histopathological image retrieval [14],

gender recognition [15], edge detection for noisy images [16], breast tumor diagnosis [17], texture image retrieval [18], face similarity comparison [19], color texture recognition [20]. Most deep learning or machine learning models for computer vision applications like image LBP demand intense computational power. In this study, a general-purpose LBP operation is written using basic TensorFlow operations. Experimental evaluations were realized using image batches ranging from 1 to 1024 and images in various dimensions ranging from 28×28 to 448×448. In order to show the acceleration of the Tensorflow based algorithm, the results were also obtained in Python. Proposed design can be used as a layer of a deep learning model as well as general purpose image processing applications. The rest of the paper was organized as follows; a brief information about LBP was given in Section 2, and the proposed design with TensorFlow was explained in Section 3. Comparative evaluations with the TensorFlow model were done in Section 4. Conclusions about the evaluations were given in the final section.

2. Local binary patterns

Deep learning models are intended to automatically extract the features required to make correct estimations. In computer vision, deep learning models, 2D convolution layers, and pooling layers are the key operators for automatic feature extraction. There are also preprocessing layers such as normalization, noise reduction, and histogram equalizations for the elaboration of the training dataset to provide better accuracy. One of the efficient feature extraction approaches is LBP transform [21], which extracts the texture features efficiently in pattern recognition studies. It can be used in deep learning applications for preprocessing or a non-trainable layer for increasing the model accuracy depending on the problem type.

$$LBP_{K,R}(i,j) = \sum_{k=0}^{K-1} f(p_k, p_c)2^k \qquad f(p_c, p_n) = \begin{cases} 1 & p_c < p_n \\ 0 & otherwise \end{cases} \qquad (1)$$

The idea of LBP transform is shown by Eq. 1, where K is the number of neighbor pixels, R is the radius, p_k is one of the selected neighbor pixels, and p_c is the center pixel of the selected window. An example application of this equation for K=9 and R=3, which corresponds to a 3×3 window for the computation of each pixel, is given in Figure 1. The first pixel that is used for comparison is selected as the Least Significant Bit (LSB), and it forms the first digit of the binary number, which corresponds to 2^0. In this example, it is selected as p_0=187, and when compared with the pixel at the center p_c=191, $p_0 \geq p_c$ produces "0". Similar operations are repeated clockwise till all comparisons are made for the remaining pixels. If all comparisons are written as a binary digit, it is obtained as; (01011000)2=88, the corresponding LBP transform value for the selected pixel. When the described operations are repeated for all pixels, the LBP transform is obtained in the form of a 3×3 matrix. Note that the input dimensions can be maintained by using padding to the input matrix. An example application of LBP for a complete image is shown in Figure 2.

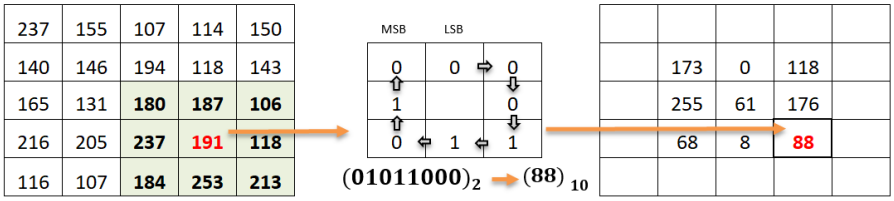


Figure 1. *An example image input image on the left and its LBP transform output on the right*

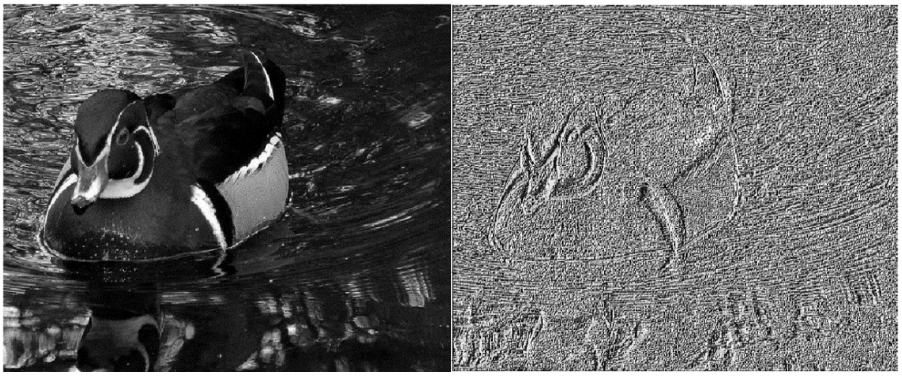


Figure 2. *An example image input image on the left and its LBP transform output on the right*

3. TensorFlow implementation

TensorFlow is an open-source numerical computation library for machine learning and deep learning applications. It has support for various high-level programming languages such as Python, C++, and Java. TensorFlow also has a low-level API to communicate with various hardware, as shown by the hierarchical block diagram given in Figure 3. There are defined implementations such as layers, losses, metrics, and various TensorFlow operations on the top of low-level API. TensorFlow provides various operations for the execution of algorithms on multicore CPU and GPU resources and makes it practical for general-purpose computations as well as training deep learning models. Data and variables in TensorFlow are defined by tensors which are N-dimensional arrays in Python. In TensorFlow 1, a session is started to compute library operations for given tensor data and variables. Recently it has been made more practical with the introduction of TensorFlow 2, removing the need for a session by running Eager execution by default.

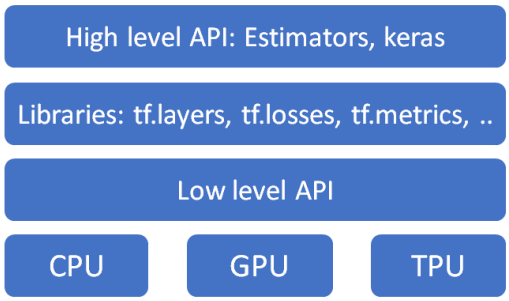


Figure 3. *Hierarchical structure of TensorFlow framework*

```
L=np.zeros((rows,cols,1),np.float32)
```

```
for i in range(1,rows-1):
    for j in range(1,cols-1):
        L[i,j]=\
        ( I[i-1,j]  >= I[i,j] )*1+\
        ( I[i-1,j+1]>= I[i,j] )*2+\
        ( I[i,j+1]  >= I[i,j] )*4+\
        ( I[i+1,j+1]>= I[i,j] )*8+\
        ( I[i+1,j]  >= I[i,j] )*16+\
        ( I[i+1,j-1]>= I[i,j] )*32+\
        ( I[i,j-1]  >= I[i,j] )*64+\
        ( I[i-1,j-1]>= I[i,j] )*128;
```

Code snippet 1. Python implementation of LBP transform

The TensorFlow framework provides various operations that can operate on tensors such as *add*(), *matmul*(), *mean*(), and *greate*r(), and combinations of these can be used to write new operations. In the LBP algorithm that is straightforward to implement, various comparison, multiplication, and adding operations are used as shown by Code snippet 1. According to the algorithm, all pixels are computed independently, and these can be defined with tensor operations. A matrix based implementation of this algorithm is given in Figure 4. Since TensorFlow operations are mainly defined for vector-matrix operations, most of the for-loops that can be defined in parallel are eliminated. In this implementation, the input variables; *P*0, *P*1, …, *P*7 define the neighbors in selected 3×3 mask in the form of matrices. The pixels in the selected masks are compared with the pixels at the center of masks one by one, as previously described in Eq. 1. After a comparison is made, false and true conditions are defined in the form of a matrix and then the comparison is done. This is repeated for every eight different pixels in a mask to form the LBP transform of the image.

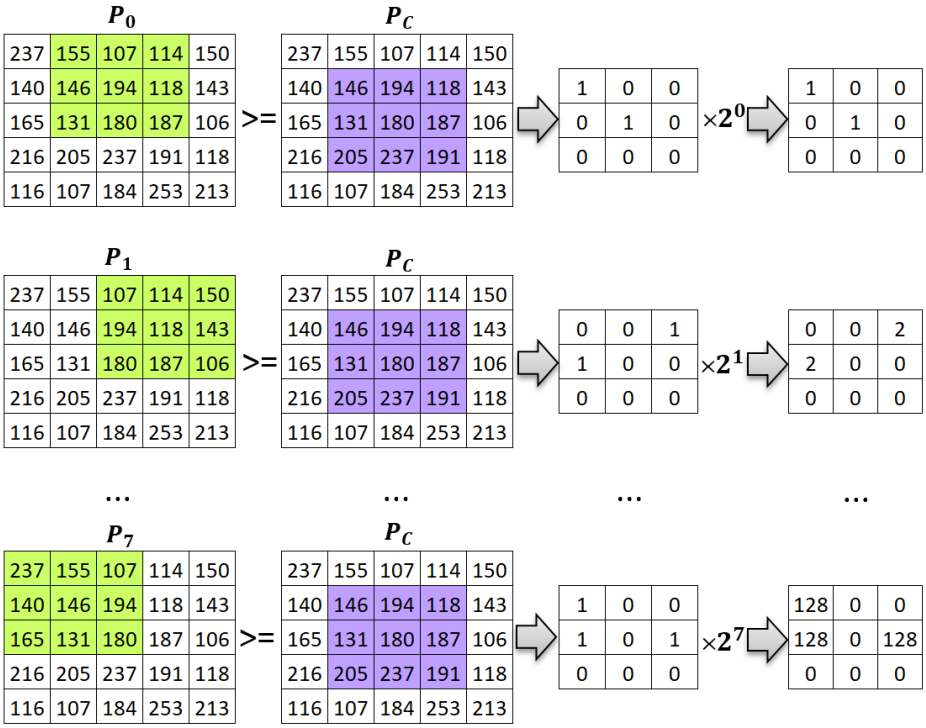


Figure 4. *TensorFlow based implementation approach*

```
g=tf.greater_equal(y01,y11)
z=tf.multiply(tf.cast(g,dtype='uint8'), tf.constant(1,dtype='uint8') )
g=tf.greater_equal(y02,y11)
tmp=tf.multiply(tf.cast(g,dtype='uint8'), tf.constant(2,dtype='uint8') )
z =tf.add(z,tmp)
g=tf.greater_equal(y12,y11)
tmp=tf.multiply(tf.cast(g,dtype='uint8'), tf.constant(4,dtype='uint8') )
z =tf.add(z,tmp)
g=tf.greater_equal(y22,y11)
tmp=tf.multiply(tf.cast(g,dtype='uint8'), tf.constant(8,dtype='uint8') )
z =tf.add(z,tmp)
g=tf.greater_equal(y21,y11)
tmp=tf.multiply(tf.cast(g,dtype='uint8'), tf.constant(16,dtype='uint8') )
z =tf.add(z,tmp)
g=tf.greater_equal(y20,y11)
tmp=tf.multiply(tf.cast(g,dtype='uint8'), tf.constant(32,dtype='uint8') )
z =tf.add(z,tmp)
g=tf.greater_equal(y10,y11)
tmp=tf.multiply(tf.cast(g,dtype='uint8'), tf.constant(64,dtype='uint8') )
z =tf.add(z,tmp)
g=tf.greater_equal(y00,y11)
tmp=tf.multiply(tf.cast(g,dtype='uint8'), tf.constant(128,dtype='uint8') )
z =tf.add(z,tmp)
```

Code snippet 2. TensorFlow implementation of LBP transform

A TensorFlow implementation of this algorithm is given by Code snippet 2. Since TensorFlow operations are mainly defined for vector-matrix operations, most of the for-loops defined in parallel are eliminated. The pixels in the selected masks are compared with the pixels at the center of masks one by one, as previously described in Eq. 1. After a comparison is made, false and true conditions are defined in the form of a matrix, and then the comparison is made. This is repeated for every eight different pixels in a mask to form the LBP transform of the image.

4. Experimental results

Experimental evaluations were realized using Python 3.7.9 and TensorFlow 2.3.1 on Ubuntu 18.04 operating system. Test hardware has an AMD FX2700 eight-core CPU and GTX1080 GPU which has 8GB of memory and 2560 CUDA cores. Time measurements for the test runs were realized with the time library of the python as given by the example code shown by Code snippet 3. All the measurements were repeated 30 times to and the average time is used to form experimental results. Image batches were selected from the ImageNet data set [22], and the sizes of the images were resized to test cases which range from 28×28 to 448×448 for all evaluations. However, it should be noted that the contents of the images usually affect the extracted features but have ignorable effects on the execution durations that are hard to measure.

```
# get current time
start_time = time.time()

#Compute LBP using TensorFlow for given batch of images
result = tf_lbp(batch_of_images).numpy()

# get current time and compute the elapsed time
elapsed_time  = time.time() - start_time
```

Code snippet 3. TensorFlow implementation of LBP transform

Table 1. *Python implementation running times (seconds) for LBP algorithm*

Batch size	Image size				
	28×28	**56×56**	**112×112**	**224×224**	**448×448**
1	0.0187	0.0774	0.3172	1.2609	4.9940
2	0.0365	0.1533	0.6269	2.5125	10.107
4	0.0713	0.3009	1.2458	5.0312	20.010
8	0.1425	0.5987	2.4868	10.000	40.354

Table 1 shows the results for the Python implementation for comparison with the TensorFlow results. Numerical results show that CPU running time varies somewhat in proportion to the image size, which is a usual case. As the image size is increased, running time is increased in the same way. For example, while a 224×224 image is processed in 1.26 seconds, a 448×448 image is processed in 4.99 seconds. Because the number of pixels is increased four times, the running time is also increased approximately four times. Similar behavior is observed when the number of images in the batch is increased, as shown by Figure 5, where speed-up evaluations are given. This is not the case for GPU running times when the numerical results given in Table 2 given for TensorFlow are investigated. This is due to the cost of initializing the GPU devices and the management overhead of the CUDA cores.

Table 2. *TensorFlow running times (seconds) for LBP algorithm*

Batch size	Image size				
	28×28	**56×56**	**112×112**	**224×224**	**448×448**
1	0.0017	0.0018	0.0019	0.0020	0.0023
2	0.0017	0.0018	0.0019	0.0021	0.0029
4	0.0017	0.0018	0.0020	0.0023	0.0039
8	0.0018	0.0018	0.0020	0.0027	0.0064
16	0.0018	0.0019	0.0022	0.0038	0.0116
32	0.0019	0.0020	0.0025	0.0063	0.0266
64	0.0019	0.0022	0.0036	0.0109	0.0787
128	0.0020	0.0027	0.0064	0.0211	0.1703
256	0.0020	0.0032	0.0082	0.0566	0.2341
512	0.0021	0.0033	0.0082	0.0563	0.2349
1024	0.0021	0.0033	0.0086	0.0577	0.2385

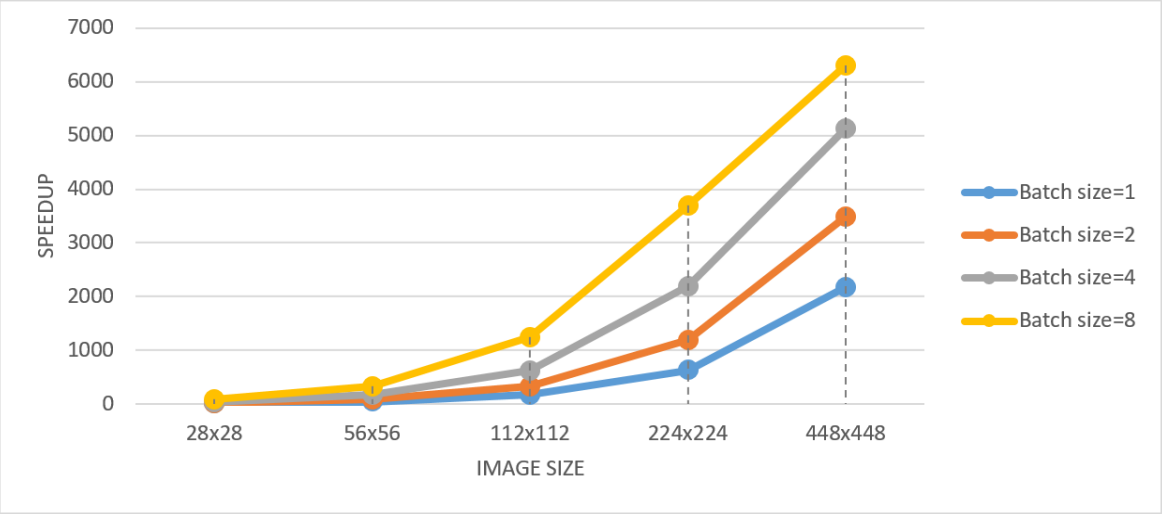


Figure 5. *TensorFlow speed-up over Python implementation*

5. Conclusions

In the presented study, a method for accelerating the LBP transform computations using TensorFlow operators has been proposed. The LBP algorithm has been defined in terms of matrix operations so that TensorFlow operators can be efficiently used. The acceleration provided by the Tensorflow method has been illustrated by comparing it with baseline Python implementation. The results show that TensorFlow running times for the LBP algorithm are far better than the direct Python runs. The significant difference between the running times of Python and TensorFlow algorithms is mainly due to the two factors. First, python codes are interpreted line by line, and therefore, results are computed slower than compiled codes. The other is GPU acceleration provided by TensorFlow library in addition to using compiled functions. Speed-up obtained by TensorFlow increases considerably as the image size is increased. This is because initializations and GPU device communications for small operations are usually costly. As the image size increases, the overhead of managing CPU and GPU device gets smaller, resulting in more efficiency. Designed LBP algorithm can be used to accelerate computer vision applications that involve LBP transform since TensorFlow allows general-purpose computations.

References

[1] Abadi M. *et al.*, "TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems," *arXiv Prepr. arXiv1603.04467*, (2016).

[2] Abadi M.*et al.*, "TensorFlow: A system for large-scale machine learning," in *Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation, OSDI 2016*, (2016), 265–283.

[3] Chien S.W.D., Markidis S., Olshevsky V., Bulatov Y., Laure E., Vetter J., "TensorFlow Doing HPC," in *2019 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*, (2019), 509–518.

[4] Agrawal A.*et al.*, "TensorFlow Eager: A Multi-Stage, Python-Embedded DSL for Machine Learning," *arXiv Prepr. arXiv1903.01855*, (2019).

[5] Parisi L., Neagu D., Ma R., Campean F., "QReLU and m-QReLU: Two novel quantum activation functions to aid medical diagnostics," *arXiv Prepr. arXiv2010.08031*, (2020).

[6] Williams T., Li R., "Wavelet pooling for convolutional neural networks," in *6th International Conference on Learning Representations, ICLR 2018 - Conference Track Proceedings*, (2018).

[7] Pakravan S., Mistani P.A., Aragon-Calvo M.A., Gibou F., "Solving inverse-PDE problems with physics-aware neural networks," *arXiv Prepr. arXiv2001.03608*, (2020).

[8] Perepelkin D., Ivanchikova M., "Research of Neural Network Architectures for Solving Adaptive Routing Problems in Multiprovider Networks of Distributed Data Centers," in *2020 9th Mediterranean Conference on Embedded Computing, MECO 2020*, (2020), 1–5.

[9] Pietikäinen M., Hadid A., Zhao G., Ahonen T., *Computer Vision Using Local Binary Patterns*, vol. 40. Springer Science & Business Media, (2011).

[10] Pietikäinen M., "Local Binary Patterns," *Scholarpedia*, 5, 3, (2010), 9775.

[11] Al-Shatnawi A., Al-Saqqar F., Alhusban S., "A holistic model for recognition of handwritten arabic text based on the local binary pattern technique," *Int. J. Interact. Mob. Technol.*, 14, 16, (2020), 20–34.

[12] Nigam S., Singh R., Misra A. K., "Local Binary Patterns Based Facial Expression Recognition for Efficient Smart Applications," in *Security in Smart Cities: Models, Applications, and Challenges*, Springer, (2019), 297–322.

[13] Hassaballah M., Alshazly H.A., Ali A.A., "Ear recognition using local binary patterns: A comparative experimental study," *Expert Syst. Appl.*, vol. 118, (2019), 182–200.

[14] Erfankhah H., Yazdi M., Babaie M., and Tizhoosh H.R., "Heterogeneity-Aware Local Binary Patterns for Retrieval of Histopathology Images," *IEEE Access*, vol. 7, (2019), 18354–18367.

[15] El-Alfy E.S.M., Binsaadoon A.G., "Automated gait-based gender identification using fuzzy local binary patterns with tuned parameters," *J. Ambient Intell. Humaniz. Comput.*, 10, 7, (2019), 2495–2504.

[16] Shen T., Huang F., Jin L., "An improved edge detection algorithm for noisy images," *ACM Int. Conf. Proceeding Ser.*, 36, 3, (2019), 84–88.

[17] Touahri R., Azizi N., Hammami N.E., Aldwairi M., Benaida F., "Automated breast tumor diagnosis using local binary patterns (LBP) based on deep learning classification," in *2019 International Conference on Computer and Information Sciences, ICCIS 2019*, (2019), 1–5.

[18] Yang W., Krishnan S., "Combining Temporal Features by Local Binary Pattern for Acoustic Scene Classification," *IEEE/ACM Trans. Audio Speech Lang. Process.*, 25, 6, (2017), 1315–1321.

[19] Gundogdu B., Bianco M. J., "Collaborative similarity metric learning for face recognition in the wild," *IET Image Process.*, 14, 9, (2020), 1733–1739.

[20] Hosny K.M., Magdy T., Lashin N.A., "Improved color texture recognition using multi-channel orthogonal moments and local binary pattern," *Multimed. Tools Appl.*, (2021), 1–16.

[21] Ojala T., Pietikäinen M., Mäenpää T., "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, 24, 7, (2002), 971–987.

[22] Deng J., Dong W., Socher R., Li L.-J., Kai Li, Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," (2010), 248–255.

MANAS Journal of Engineering
ISSN 1694-7398 | e-ISSN 1694-7398

Volume 9, Issue 1, (2021) Pages 22-29
https://doi.org/10.51354/mjen.848282

Anticarcinogenic properties of malic acid on glioblastoma cell line through necrotic cell death mechanism

Mehmet Enes Arslan

Department of Molecular Biology and Genetics, Faculty of Science, Erzurum Technical University, Erzurum, Turkey, enes.aslan@erzurum.edu.tr, ORCID: 0000-0002-1600-2305

ABSTRACT

This study aimed to investigate the anticarcinogenic and genotoxic damage potentials of malic acid on human fibroblast cells (HDFa) and glioblastoma (U87-MG) cell lines. MTT cell viability and LDH release assays were performed to understand cytotoxic features of malic acid on different cell lines. Also, Hoechst 33258 fluorescent staining was used to monitor nuclear abnormalities including micronucleus, lobbed, and notched structures. Furthermore, cellular death mechanisms behind the malic acid application were investigated via the use of flow cytometry analysis. According to cell viability analysis, malic acid showed a greater effect on U87-MG compared to HDFa cell line in terms of cytotoxicity. Similarly, chromosomal integrity assay put forth a higher number of nuclear abnormalities in U87-MG cells when compared to HDFa cell lines, and aberrations were analyzed to amplify when malic acid concentration increased. Finally, flow cytometry analysis demonstrated higher necrotic cell death in U87-MG cells than HDFa cell line. On the other hand, apoptotic cell death was the main cytotoxic mechanism against malic acid exposure in the HDFa cell line. In light of these results, it can be concluded that in higher concentrations, malic acid has an anticarcinogenic effect on glioblastoma cells via the necrotic pathway, and it also shows apoptotic properties on the fibroblast cell line. When mutagenic properties are compared, it could be understood that malic acid had a greater impact on glioblastoma cells.

ARTICLE INFO

Research article

Received: 28.12.2020
Accepted: 9.03.2021

Keywords:
malic acid,
glioblastoma,
in vitro,
necrosis,
DNA mutations

1. Introduction

Cancer is a destructive disease characterized by abnormalities in uncontrolled and rapidly dividing cells that cause the death of thousands of people worldwide every year. This is of great concern to people, as one in four people has a high risk of cancer. In the World Cancer Report, it was reported that cancer statistics could increase by 50% until 2020 and that 15 million people will be examined for several types of cancer [1–3]. Although glioblastoma is one of the most aggressive brain tumors, there are multiple treatment methods available such as surgery, radiation, and cytotoxic chemotherapy. On the other hand, recent clinic researches have shown an average survival of only 14–16 months with a 26–33% 2-year survival rate [4,5].

A number of natural compounds particularly abundant in plants have recently attracted much interest for their many biological properties including antioxidant, antitumor, antibacterial, antimutagen, and several others [6–9]. Malic acid, a type of alpha-hydroxy acid, is an organic acid that is used in many areas from food to cosmetics. It is a known fact that more than fifty cosmetic products contain malic acid [10,11]. Recent studies have shown that keratin-malate supplementation can be used as a performance-enhancing agent and increases the physical performance of athletes [12]. It has been investigated that a diet high in malic acid, which is the precursor of citrate, increases kidney stone formation [13]. The malic acid spray has been used orally to treat mouth dryness. The analysis showed that the symptoms were found to be improved in people with mouth dryness after two weeks of malic acid usage[14]. In addition to this, its availability as carrier systems for the release of poly (malic acid) targeted drugs has also been reported in many studies [15,16]. However, conditions such as redness, swelling, burning, phototoxicity caused by alpha hydroxy acids lead to safety concerns [17,18].

Previous studies showed that tumor-specific targeting is important for increasing the effectiveness of anti-tumor drugs. According to the analysis, two different covalently conjugated antibodies to the poly (β-L-malic acid) nanocarrier increased

tumor targeting efficiency [19]. Simultaneous delivery to two different AON tumor cells was shown using Polycefin, a new nanoscale PMLA-based drug, that simultaneously blocks the in vivo expression of two laminin-8 chains. In another study, simultaneous delivery to two different AON tumor cells was shown using Polycefin, a new nanoscale PMLA-based drug, that simultaneously blocks the in vivo expression of two laminin-8 chains. It has been shown to inhibit laminin-8 synthesis in U87MG and T98G in cultured human glioma cells [20]. It was also stated that the immunoconjugates of poly L-malic acid prepared with CTLA-4 and PD-1 can get past the blood-brain barrier and trigger local immune responses for glioma cancer treatment [21].

In this study, MTT and LDH tests were performed to examine the effects of malic acid on cell viability on the U87-MG glioblastoma cell line. Hoechst 33258 staining was used to investigate the nuclear status of the cell cultures. Furthermore, flow cytometry analysis was performed to understand the underlying mechanisms behind the main cellular deaths. Meanwhile, the cytotoxicity and mutagenicity effects of the malic acid, one of the alpha-hydroxy acids, were analyzed on the HDFa human fibroblast cell line to create a safety report.

2. Materials and methods

2.1. Cell cultures

Human dermal fibroblast (HDFa) and glioblastoma (U87-MG) cell cultures were supplied from American Type Culture Collection (ATCC). Solid L-malic acid (MA, CAS 97-67-6, Sigma Aldrich) was commercially available. The human fibroblast cells and glioblastoma cells were seeded in 48 well plates containing DMEM (%10 FBS and 100 U/ml penicillin-streptomycin) about 10^5 cells per well. Culture plates incubated at humidified 5% CO_2 at 37°C for 24 hours [22].

2.2. MTT analysis

MTT (3-(4,5-Dimethylthiazol-2-yl)- 2,5-diphenyltetrazolium bromide) assay, one of the most widely used viability tests, is a colorimetric analysis that measures metabolic activity using the reduction of a yellow tetrazolium salt. Culture plates were treated with L-malic acid at different concentrations (3.12 µg/mL, 6.25 µg/mL, 12.5 µg/mL, 50 µg/mL, 100 µg/mL, 200 µg/mL, 400 µg/mL). In addition, while the cells in the positive control group were treated with 1% Triton X-100, no treatment was applied to the wells designated as negative control groups. After 24 hours of incubation, an MTT reagent was added to each well. Plates were incubated for 3 hours at humidified 5% CO_2 incubator at 37°C and culture medium was discarded. Generated formazan crystals were dissolved by using dimethyl sulfoxide. The plates were analyzed at 570 nm wavelength using a microplate reader [23].

2.3. LDH release test

LDH cytotoxicity assay kit (Cayman Chemical Company®, Ann Arbor, MI, USA) was applied as per the manufacturer's instructions for the LDH assay application. The cells were seeded to 48-well plates and wide spectrum concentrations of malic acid (3.12 µg/mL, 6.25 µg/mL, 12.5 µg/mL, 50 µg/mL, 100 µg/mL, 200 µg/mL, 400 µg/mL) were administered to cell culture as triplicates for 24 hours at humidified 5% CO2 incubator at 37°C.. In addition, while the cells in the positive control group were treated with 1% Triton X-100, no treatment was applied to the wells designated as negative control groups. After that, 100 µL supernatant was transferred to a fresh 48-well plate and 100 µL of the reaction mixture was added to the samples and incubated for 30 minutes at room temperature. Finally, a microplate reader was used to analyze the absorbance of the cultures at 490 nm [24].

2.4. Hoechst staining

For the Hoechst Staining, firstly the medium was removed from the prepared culture plates by treating with different concentrations of L-malic acid. The wells were washed with PBS twice. Then cells were fixed 4% paraformaldehyde at 4°C for 30 minutes. Cells were washed with PBS. Finally, wells were incubated with 1mM Hoechst 33258 fluorescent dye at room temperature for 5 minutes. Nuclear anomalies were calculated under a fluorescence microscope. A total of 1000 cells were counted and mutations were determined for both glioblastoma and fibroblast cell lines [25].

2.5. Flow cytometry analysis

Flow cytometry analysis was performed to examine the effectiveness of apoptosis, one of the cell death pathways. Annexin V was used to analyze apoptotic cells. Briefly, the phosphatidylserine structures, normally found in the inner surface of the cell membrane, rose to the surface of the cell membrane as the indicator of apoptotic cell death. To investigate the necrotic cell ratios, propidium iodide staining was used to analyze the cell nucleus. For this purpose, $5x10^4$ cells were collected through centrifugation and the cells were resuspended in 500 µl of 1X binding buffer. 5 µl of Annexin V-FITC and 5 µl of propidium iodide (50 µg/ml) were added to the cultures and incubated in the dark for 5 minutes. The cells were then fixed with 4% paraformaldehyde in phosphate-buffered saline at 4°C for 30 minutes. Finally, cultures were examined with flow cytometry (The CyFlow® Cube 6, Germany) [26].

2.6. Statistical analysis

Statistical analyzes of the numerical data collected from the studies were performed using GraphPad Prism 7. Anova: Dunnett's and Tukey multiple comparison tests were performed following the Two Way Anova analysis for statistical evaluation, and the statistical significance level was accepted as $p < 0.05$.

3. Findings

Both MTT and LDH cytotoxicity assays showed that malic acid application for 24 hours significantly decreased cell viability in fibroblast cell culture at 100, 200, and 400 µg/ml concentrations, however, there was no significant cell viability difference compared to the negative control in lower concentrations (50 -3.12 µg/ml) (Figure 1). On the other hand, malic acid exposure inhibited cell growth in glioblastoma cell lines significantly, even at smaller concentration levels as 12,5 to 400 µg/ml compared to fibroblast cell lines (Figure 2).

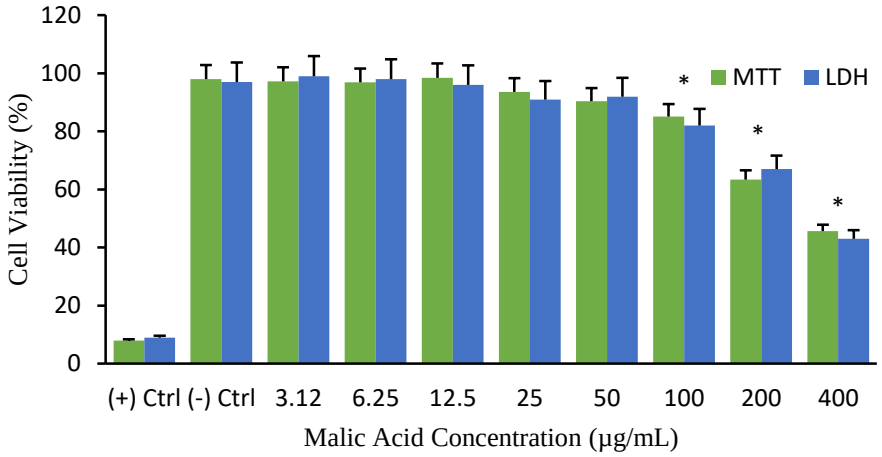


Figure 1: *Cell viability tests (MTT and LDH assays) of the fibroblast cell line (HDFa) against the malic acid application (3,12 to 400 µg/ml) for 24 hours. Symbol (*) represents a statistically significant difference (P<0.05) compared to the negative control. (GraphPad Prism 7, Anova: Tukey's post hoc test was used to calculate significantly different values)*

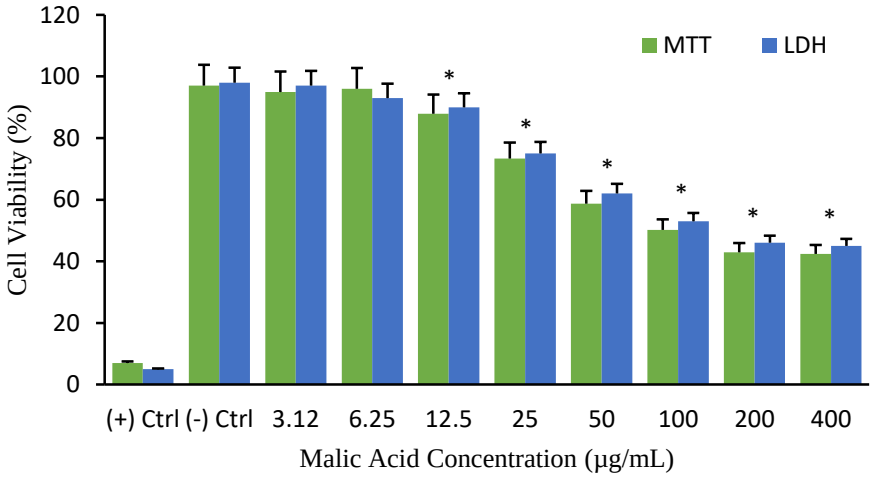


Figure 2: *Cell viability tests (MTT and LDH assays) on glioblastoma cells (U87-MG) against the malic acid application (3,12 to 400 µg/ml) for 24 hours. Symbol (*) represents a statistically significant difference (P<0.05) compared to the negative control. (GraphPad Prism 7, Anova: Tukey's post hoc test was used to calculate significantly different values)*

Secondly, Hoechst 33258 fluorescent staining analysis put forth that malic acid in higher concentrations affects nuclear structures negatively in both fibroblast and glioblastoma cell lines. Observable nucleus mutations such as micronucleus,

lobbed and notch structures were highly increased in respect to malic acid exposure (Figure 3 and 4). When two cell lines were compared to each other in terms of nuclear mutations, it was found that malic acid had more severe effects on glioblastoma cell lines with a greater number of mutations specifically in higher concentrations in contrast to fibroblast cells (Table 1).

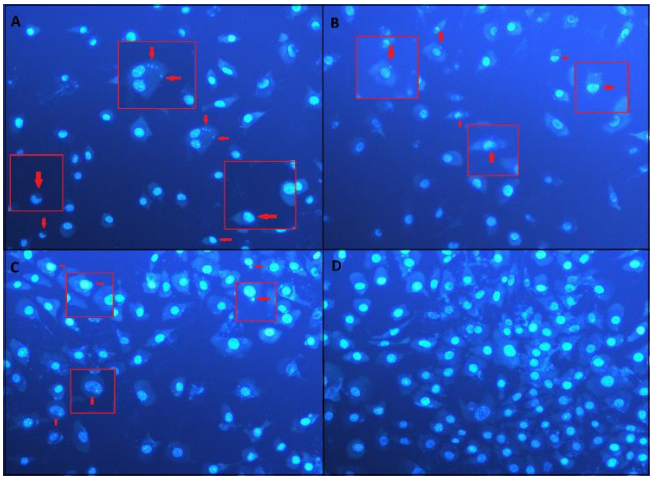


Figure 3: *Hoechst 33258 fluorescent staining of Fibroblast cell line for nuclear abnormalities (20X magnifications).* ***A)*** *400 µg/ml malic acid application for 24 hours,* ***B)*** *200 µg/ml malic acid application for 24 hours,* ***C)*** *100 µg/ml malic acid application for 24 hours,* ***D)*** *Negative control fibroblast cell line culture without malic acid application (arrows indicate nuclear abnormalities)*

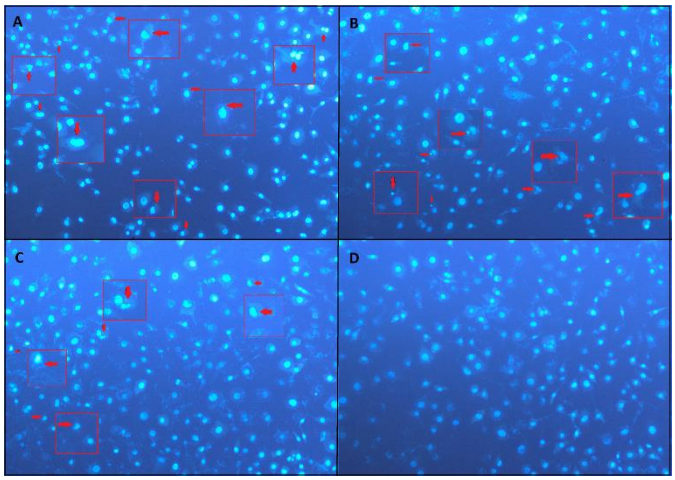


Figure 4: *Hoechst 33258 fluorescent staining of glioblastoma cells (U87-MG) for nuclear abnormalities (20X magnifications).* ***A)*** *400 µg/ml malic acid application for 24 hours,* ***B)*** *200 µg/ml malic acid application for 24 hours,* ***C)*** *100 µg/ml malic acid application for 24 hours,* ***D)*** *Negative control (U87-MG) cell culture without malic acid application (arrows indicate nuclear abnormalities)*

Table 1: *Nuclear abnormalities (NA) in Fibroblast (HDFa) and Glioblastoma (U87-MG) cell lines against malic acid applications in various concentrations. Values are mean ± SEM (standard error of the mean) from 3 samples for every 1000 cells. Different letters in the same column represent statistical differences from each other. Calculations were performed after 24 hours of application. (GraphPad Prism 7, Anova: Dunnett's multiple comparison test was used to calculate the significantly different samples)*

Treatment	**Nuclear abnormalities (NA)**			
HDFa	**Total MN**	**Total lobbed**	**Total notched**	**Mean NA/1000 cells ± SD**
400 µg/ ml	45	15	14	0,074±0,005[a]
200 µg/ ml	20	12	11	0,043±0,003[b]
100 µg/ ml	9	11	9	0,029±0,002[c]
(-) Ctrl	5	3	2	0,010±0,002[d]
U87-MG	**Total MN**	**Total lobbed**	**Total notched**	**Mean NA/1000 cells ± SD**
400 µg/ ml	54	20	19	0,093±0,007[e]
200 µg/ ml	29	16	17	0,062±0,005[f]
100 µg/ ml	15	8	10	0,033±0,003[c]
(-) Ctrl	7	8	8	0,023±0,002[g]

To examine the cytotoxicity mechanism behind the malic acid applications, Flow Cytometry analysis was performed on both fibroblast and glioblastoma cell lines for different concentrations of malic acid exposures (100, 200, and 400 µg/ml). In these examinations, the viable cell ratio was found to be inversely proportional to malic acid concentrations. Flow Cytometry results showed that cell viabilities were analyzed as 42.24%, 56.25%, and 72.39% for 400 µg/ml, 200 µg/ml, and 100 µg/ml concentration of malic acid respectively (Figure 5). In addition, it was observed that fibroblast cells decrease predominantly by apoptotic cell death mechanism. Furthermore, Flow Cytometry examinations on glioblastoma cell lines produced very opposite results. Necrotic cell death was observed after 24 hours of malic acid application to the glioblastoma cell culture and live-cell ratios were found to be 46.68%, 56.82%, and 70.73% for 400 µg/ml, 200 µg/ml, and 100 µg/ml malic acid concentrations (Figure 6).

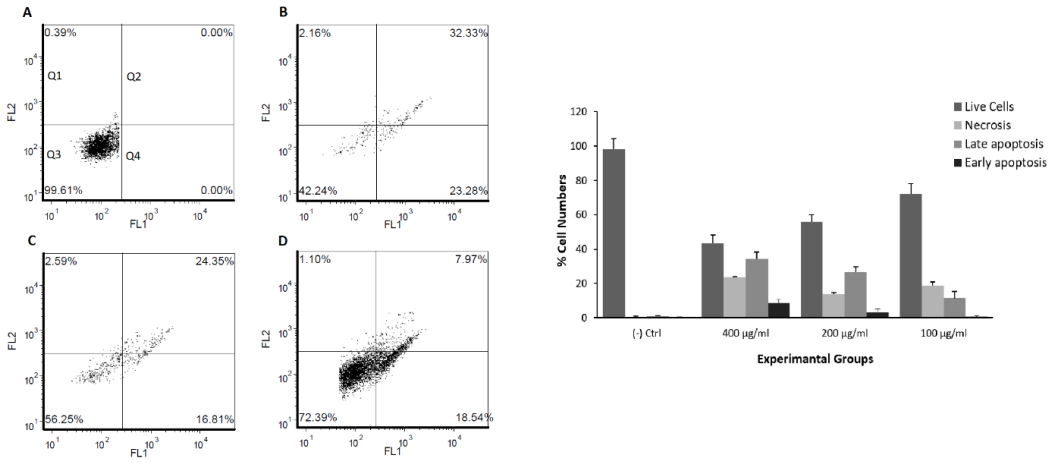


Figure 5: *Flow Cytometry analysis (Annexin V- (FL1-H) and PI- (FL2-H) labeled cells) of Fibroblast cell line against malic acid cytotoxicity.* ***A)*** *Negative control for 24 hours,* ***B)*** *400 µg/ml malic acid application for 24 hours,* ***C)*** *200 µg/ml malic acid application for 24 hours,* ***D)*** *100 µg/ml malic acid application for 24 hours. (Q1: Necrotic cells, Q2: Late apoptotic cells, Q3: living cells, and Q4: early apoptotic cells)*

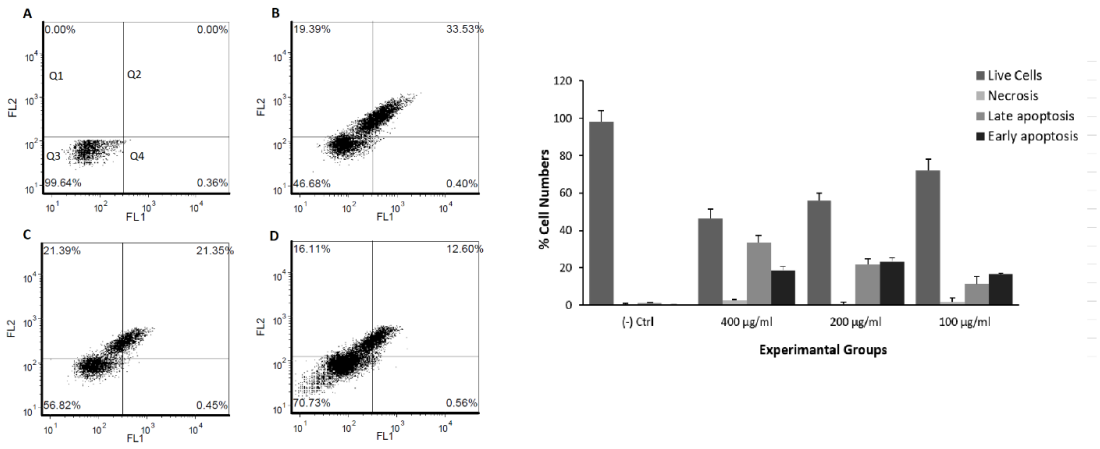


Figure 6: *Flow Cytometry analysis (Annexin V- (FL1-H) and PI- (FL2-H) labeled cells) of U87-MG cells against malic acid cytotoxicity.* ***A)*** *Negative control for 24 hours,* ***B)*** *400 µg/ml malic acid application for 24 hours,* ***C)*** *200 µg/ml malic acid application for 24 hours,* ***D)*** *100 µg/ml malic acid application for 24 hours. (Q1: Necrotic cells, Q2: Late apoptotic cells, Q3: living cells, and Q4: early apoptotic cells)*

4. Discussions

Malic acid is one of the most important chemicals used as a nutritional enhancer in the food industry. Malic acid's specific features such as hygroscopicity and increasing storage capability make it a commercially favorable molecule for several industries [27,28]. Besides, malic acid applications in the medical industry have a wide range of potential for skin problems such as warts and acne with its favorable features such as easiness to apply, being cheap, and having no serious adverse effects [29]. In our study, it was shown that lower concentrations of malic acid (50 to 3.12 µg/ml) had no significant toxic effects on the fibroblast cell lines. This result could indicate that under specific concentrations, malic acid applications would not show any crucial side effects on human health. Contrarily, cell viability assays on glioblastoma cell lines expressed the anticarcinogenic activity of malic acid even for lower concentrations (up to 12.5 µg/ml). It could be proposed that malic acid derivatives with enhanced activity could be used as anticancer agents for specific tumor treatments under certain doses without damaging healthy cells. Previous studies claimed that 5-fluorouracils conjugated poly (α-malic acid) caused no toxicity however, the molecule showed increased antitumor property on P-388 lymphocytic leukemia mice models [30]. Moreover, different conjugates of malic acid such as poly (α-malic acid)/ADR(Adriamycin) were analyzed to examine anticarcinogenic activity on human hepatoma cells in vitro [31]. In a study, Poly (β-malic acid) (PMLA) conjugated chitosan (LMC) nanoparticles were used as a co-delivery of the antitumor drug system for paclitaxel (PTX) delivery. Applications of the conjugates on hepatoma-bearing mice models were analyzed to examine prolonged survival periods and enhanced antitumor efficacy.

In this study, Hoechst 33258 fluorescent staining was used to analyze the genotoxic potential of malic acid on both fibroblast and glioblastoma cell lines. Malic acid was shown to have higher genotoxic activity against glioblastoma cell lines. Although genotoxic activity was shown on the glioblastoma cell line rather than fibroblast cells at higher concentrations in our study, several projects claimed the non-genotoxic potential of malic acid [32–34]. Besides, it was revealed previously that malic acid had genotoxic potential when reacted with trichloroacetaldehyde in an aqueous solution [11]. Ameliorating effects of malic acid against heavy metals such as cadmium (Cd) were examined in a previous research study. This study revealed that malic acid could reduce mutagenic ROS production resulted from Cd accumulation by enhancing cellular glutathione level [35]. Furthermore, several studies examined the protective effect of malic acid against aluminum-induced toxicity *in vitro* and *in vivo* [36–38].

In addition to these examinations, Flow Cytometry analysis indicated interesting results related to cellular death mechanisms in healthy and cancer cell lines against malic acid applications. Flow Cytometry analysis on fibroblast cells revealed that malic acid exposure to the healthy cell type resulted in apoptotic cell death. This finding correlated with the literature that malic acid application to the immortalized human keratinocyte line (HaCaT) was shown to activate mitochondria-dependent and endoplasmic reticulum stress-related apoptosis [18]. On the other hand, this study examined for the first time that low concentration of malic acid application to the glioblastoma cell line induced necrotic cell death. Flow Cytometry analysis showed that the increased concentrations of malic acid lead glioblastoma cells to late apoptotic death rather than necrosis.

5. Conclusion

In light of these findings, it can be understood that malic acid has anticarcinogenic activity against glioblastoma cell lines even at lower concentrations. According to genotoxicity and cytotoxicity analysis on the fibroblast cell lines, higher concentrations of malic acid can result in a mutagenic outcome however, under certain concentrations, no adverse reactions are observed. Interestingly, the cytotoxicity mechanism of malic acid on fibroblast cell lines is apoptosis at high concentrations. On the other hand, glioblastoma cells have a more necrotic tendency towards malic acid applications. From these examinations, it is concluded that malic acid applications at specific concentrations can be used as an anticancer agent without causing any detrimental and mutagenic properties.

References

[1] Pal SK, Mittal B. "Improving cancer care in India: Prospects and challenges", Asian Pacific J Cancer Prev 5, (2004), 226–228.

[2] Caparica R, Júlio A, Araújo MEM, Baby AR, Fonte P, Costa JG, et al. "Anticancer Activity of Rutin and Its Combination with Ionic Liquids on Renal Cells", Biomolecules, 10, (2020), 233.

[3] Pitucha M, Korga-Plewko A, Kozyra P, Iwan M, Kaczor AA. "2,4-Dichlorophenoxyacetic Thiosemicarbazides as a New Class of Compounds against Stomach Cancer Potentially Intercalating with DNA", Biomolecules, 10, (2020), 296.

[4] Lambert J-C, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al. "Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease", Nat Genet, 45, (2013), 1452–1458.

[5] Sato A, Sunayama J, Okada M, Watanabe E, Seino S, Shibuya K, et al. "Glioma-Initiating Cell Elimination by Metformin Activation of FOXO3 via AMPK. Stem Cells", Transl Med, 1, (2012), 811–824

[6] Marinelli L, Fornasari E, Eusepi P, Ciulla M, Genovese S, Epifano F, et al. "Carvacrol prodrugs as novel antimicrobial agents", Eur J Med Chem, 178, (2019), 515–529.

[7] Turkez H, Togar B, Polat E. "Olive leaf extract modulates permethrin induced genetic and oxidative damage in rats", Cytotechnology, 64, (2012), 459–464.

[8] Turkez H, Geyikoglu F, Yousef MI. "Beneficial effect of astaxanthin on 2,3,7,8-tetrachlorodibenzo- p -dioxin-induced liver injury in rats", Toxicol Ind Health, 29, (2013), 591–599.

[9] Ploskonos M v., Syatkin S p., Neborak E v., Hilal A, Sungrapova K y., Sokuyev R i., et al. "Polyamine Analogues of Propanediamine Series Inhibit Prostate Tumor Cell Growth and Activate the Polyamine Catabolic Pathway", Anticancer Res, 40, (2020), 1437–1441.

[10] Al Bawab A, Friberg SE, Bergamaschi MM, Santos ODH. "Some non-equilibrium phenomena in the malic acid/water/Polysorbate 81 system", Int J Pharm, 332, (2007), 140–146.

[11] Fiume MZ. "Final Report on the Safety Assessment of Malic Acid and Sodium Malate", Int J Toxicol, 20, (2001), 47–55.

[12] Tyka AK, Chwastowski M, Cison T, Palka T, Tyka A, Szygula Z, et al. "Effect of creatine malate supplementation on physical performance, body composition and selected hormone levels in spinters and long-distance runners", Acta Physiol Hung, 102, (2015), 114–122.

[13] Manfredini R, De Giorgi A, Storari A, Fabbian F. "Pears and renal stones: Possible weapon for prevention? A comprehensive narrative review", Eur Rev Med Pharmacol Sci, 20, (2016), 414-25

[14] Gõmez-Moreno G, Aguilar-Salvatierra A, Guardia J, Uribe-Marioni A, Cabrera-Ayala M, Delgado-Ruiz RA, et al. "The efficacy of a topical sialogogue spray containing 1% malic acid in patients with antidepressant-induced dry mouth: A double-blind, randomized clinical trial", Depress Anxiety, 30, (2013), 137–142.

[15] Ljubimova JY, Fujita M, Khazenzon NM, Lee B-S, Wachsmann-Hogiu S, Farkas DL, et al. "Nanoconjugate based on polymalic acid for tumor targeting", Chem Biol Interact, 171, (2008), 195–203.

[16] Ljubimova JY, Fujita M, Ljubimov A V., Torchilin VP, Black KL, Holler E. "Poly(malic acid) nanoconjugates containing various antibodies and oligonucleotides for multitargeting drug delivery", Nanomedicine, 3, (2008), 247–265.

[17] Tang S-C, Yang J-H. Dual "Effects of Alpha-Hydroxy Acids on the Skin", Molecules, 23, (2018), 863.

[18] Hsiao Y-P, Lai W-W, Wu S-B, Tsai C-H, Tang S-C, Chung J-G, et al. "Triggering Apoptotic Death of Human Epidermal Keratinocytes by Malic Acid: Involvement of Endoplasmic Reticulum Stress- and Mitochondria-Dependent Signaling Pathways", Toxins (Basel), 7, (2015), 81–96.

[19] Fujita M, Lee B-S, Khazenzon NM, Penichet ML, Wawrowsky KA, Patil R, et al. "Brain tumor tandem targeting using a combination of monoclonal antibodies attached to biopoly(β-l-malic acid)", J Control Release, 122, (2007), 356–363.

[20] Fujita M, Khazenzon NM, Ljubimov A V., Lee B-S, Virtanen I, Holler E, et al. "Inhibition of laminin-8 in vivo using a novel poly(malic acid)-based carrier reduces glioma angiogenesis", Angiogenesis, 9, (2006), 183–191

[21] Galstyan A, Markman JL, Shatalova ES, Chiechi A, Korman AJ, Patil R, et al. "Blood–brain barrier permeable nano immunoconjugates induce local immune responses for glioma therapy", Nat Commun, 10, (2019), 3850.

[22] Emsen B, Aslan A, Turkez H, Joughi A, Kaya A. "The anti-cancer efficacies of diffractaic, lobaric, and usnic acid: In vitro inhibition of glioma", J Cancer Res Ther, 14, (2018), 941.

[23] Türkez H, Arslan ME, Özdemir Ö, Chikha O. "Ameliorative effect of boric acid against nicotine-induced cytotoxicity on cultured human primary alveolar epithelial cells", BORON, 1, (2016), 104–109.

[24] Türkez H, Arslan ME, Sönmez E, Açikyildiz M, Tatar A, Geyikoğlu F. "Synthesis, characterization and cytotoxicity of boron nitride nanoparticles: emphasis on toxicogenomics", Cytotechnology, 71, (2019), 351–361.

[25] Turkez H, Arslan ME, Ozdemir O. "Genotoxicity testing: progress and prospects for the next decade", Expert Opin Drug Metab Toxicol, 13, (2017), 1–10.

[26] Marinelli L, Fornasari E, Di Stefano A, Turkez H, Arslan ME, Eusepi P, et al. "(R)-α-Lipoyl-Gly-L-Pro-L-Glu dimethyl ester as dual acting agent for the treatment of Alzheimer's disease", Neuropeptides, 66, (2017), 52–58.

[27] Iyyappan J, Bharathiraja B, Baskar G, Kamalanaban E. "Process optimization and kinetic analysis of malic acid production from crude glycerol using Aspergillus niger", Bioresour Technol, 281, (2019), 18–25.

[28] Dai Z, Zhou H, Zhang S, Gu H, Yang Q, Zhang W, et al. "Current advance in biological production of malic acid using wild type and metabolic engineered strains", Bioresour Technol, 258, (2018), 345–353.

[29] Chiriac A, Brzezinski P. "Topical malic acid in

combination with citric acid: an option to treat recalcitrant warts", Dermatol Ther, 28, (2015), 336–338.

[30] Ouchi T, Fujino A, Tanaka K, Banba T. "Synthesis and antitumor activity of conjugates of poly(α-malic acid) and 5-fluorouracils bound via ester, amide or carbamoyl bonds", J Control Release, 12, (1990), 143–153.

[31] Ouchi T, Kobayashi H, Hirai K, Ohya Y. "Design of Poly(α-malic acid)—Antitumor Drug—Saccharide Conjugate Exhibiting Cell-Specific Antitumor Activity", Pol Del Sys, 520, (1993), 382–94.

[32] Al-Ani FY, Al-Lami SK. "Absence of mutagenic activity of acidity regulators in the Ames Salmonella/microsome test", Mutat Res Toxicol, 206, (1988), 467–470.

[33] Kuroda M, Yoshida D, Mizusaki S. "Mutagenicity of pyrolyzates of natural substances toward Salmonella typhimurium TA97", Agric Biol Chem, 49, (1985), 1893–1895.

[34] Ishidate M, Sofuni T, Yoshikawa K, Hayashi M, Nohmi T, Sawada M, et al. "Primary mutagenicity screening of food additives currently used in Japan", Food Chem Toxicol, 22, (1984), 623–636.

[35] Guo H, Chen H, Hong C, Jiang D, Zheng B. "Exogenous malic acid alleviates cadmium toxicity in Miscanthus sacchariflorus through enhancing photosynthetic capacity and restraining ROS accumulation", Ecotoxicol Environ Saf, 141, (2017), 119–128.

[36] Domingo JL, Gómez M, Llobet JM, Corbella J. "Citric, malic and succinic acids as possible alternatives to deferoxamine in aluminum toxicity", J Toxicol Clin Toxicol, 26, (1988), 67–79.

[37] Majida A.J. al-Qayim and Sawsan Mashi. "Renal effects of propolis and malic acid in Aluminium Exposed Male Rats", Appl Sci Reports, 1, (2014).

[38] Al-qayim MAJ, Saadoon D. "Assessment of the Ameliorative Role of Proplis and Malic Acid in Intestinal and Liver Functions of Aluminum", Int J Sci Nat, (2013).

MANAS Journal of Engineering
ISSN 1694-7398 | e-ISSN 1694-7398

Volume 9, Issue 1, (2021) Pages 30-36
https://doi.org/10.51354/mjen.853971

Deep neural networks based wrist print region segmentation and classification

Hasan Erdinç Koçer[1],Kerim Kürşat Çevik[2,*]

[1] Department of Electrical and Electronics Engineering, Faculty of Technology, Selcuk University, Konya, Turkey, ekocer@selcuk.edu.tr, ORCID: 0000-0002-0799-2140

[2] Department of Management Information Systems, Faculty of Applied Sciences, Akdeniz University, Antalya, Turkey, kcevik@akdeniz.edu.tr, ORCID: 0000-0002-2921-506X

ABSTRACT

In recent years, biometric recognition based systems have become widespread. One of these is wrist-based recognition systems. In this study, wrist print based recognition system was developed by using near infrared (NIR) camera. Totally 220 NIR camera images taken from 10 for each both hands of 11 people. The obtained data set is allocated 70% (154 images) for training and 30% (66 images) for testing. The wrist regions are labeled on the training set images. Data sets were created with two different labeling methods. In the first data set, only the wrist regions were labeled and it was aimed to segment the wrist region from the image. In the second data set, the wrist images were labeled according to 22 classes and these classes were tried to be predicted. The labeled data was trained with YOLOV2 architecture supported by ResNet50 one of the deep neural network models. The trained model was tested with the remaining 30% of the data set. In the test process, the wrist region was determined in the NIR images with the trained model. As a results of the study, it was seen that the wrist regions were correctly detected in all first data set test images and the mean value of obtained similarity rates was 95.26%. In the test results of the second dataset, 92.43% classification success was obtained. Therefore, it can be said that the deep learning architectures ResNet and YOLO are effective in the segmentation of the wrist region.

ARTICLE INFO

Research article

Received: 5.01.2021
Accepted: 30.04.2021

Keywords:
wrist print recognition, deep neural networks, near-infrared camera, YOLO

**Corresponding author*

1. Introduction

Biometric recognition based systems are known as most secure systems. What makes biometric systems important in this way is that the security parameter they use is personal and there is no risk of theft and cannot be copied. Although there are many biometric recognition methods, fingerprint recognition, iris recognition, vein recognition and face recognition are mostly used today.

Vein recognition is a high-security biometric recognition approach based on human vascular structure. In this biometric identification approach, finger veins, hand veins, palm veins and wrist veins are used as biometric parameters. The identification systems which use wrist vein structure as biometric input are called wrist print recognition. In wrist print recognition system, the human wrist image is captured by using near-infrared (NIR) camera and illumination and then the wrist region is segmented for determining the wrist vein information. Then, the wrist print is used for identification. The wrist image acquisition unit used in this study is given in Figure 1.

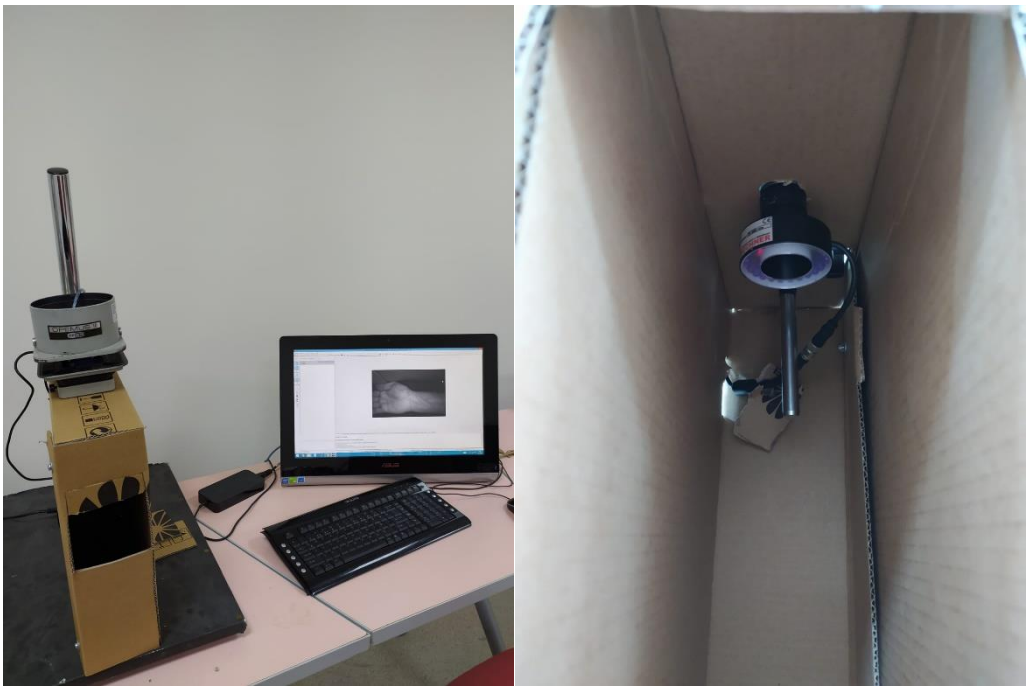


Figure 1. *A wrist image acquisition unit*

Although many studies have not been done on wrist vein recognition, we come across studies using different approaches in the literature. We can evaluate previous studies on the subject in 2 groups. The first one is hand, palm and finger vein recognition studies (will not be mentioned in this article), and the other is the wrist vein recognition works. Hartung et al. (2011) performed spectral minutiae extraction in the identification of wrist vessels [1]. In preprocessing stage, binarization and Skeletonization were implemented after noise reduction and edge enhancing. Convolution approach and Location-Based Spectral Minutiae Representation were used for extracting the feature. In this study, Hausdorff distance, Modified Hausdorff, Similarity-based Mix-matching, SML correlation and SML fast rotate methods were used for feature matching [2]. Das et al. (2014), performed the feature matching process with Support Vector Machines. Dense Local Binary Pattern approach was applied for extracting the feature [3]. In 2018, Nikisins et al. studied wrist vein recognition with cross-correlation-based comparison [4]. Adaptive histogram equalization [5] and Discrete Meyer Wavelet techniques [1] were used in preprocessing phase. Then Hessian matrix was implemented and the wrist vein features were extracted in this study. In 2020, Garcia-Martin et al. were developed portable contactless system for wrist vascular biometric identification. a new contactless database has been collected with the software algorithm TGS-CVBR®. In this work, Preprocessing and Identification Software for Contactless Vascular Biometric Recognition (PIS-CVBR®), is used for recognition the wrist images [6]. In this study three different wrist image datasets (UC3M-CV1, UC3M and PUT databases) were used and three different feature extraction methods (Scale-Invariant Feature Transform – SIFT, Speeded Up Robust Features – SURF and Oriented FAST and Rotated BRIEF (ORB) were applied to the wrist images. These wrist recognition researches are given in Table 1 including authors, year, methods, database information and best error rates in percentages.

Table 1. *Literature overview for wrist print recognition*

Researcher(s)	Year	Methods used in classification	Dataset (NIR Imaging)		Best Error Rate(s)
			Subjects	Images	
Hartung et al. [1]	2011	SML Correlation	29	348	6.13 %
		SML fast rotate			5.90 %
Das et al. [3]	2014	Support Vector Machine	50	1200	0.79 %
Nikisins et al. [4]	2018	Cross Correlation with FFT	29	348	3.58 %
Garcia-Martin et al. [7]	2020	Brute Force Matcher	50	1200	8.59%
		Fast Library for Approximate	50	1200	0.08%
		Nearest Neighbors	121	605	15.91%

As it is seen from this table, various methods were used in classification process. We have implemented Deep Neural Networks for recognition the wrist image, so a new approach is applied for wrist print recognition. In this study, the wrist vein region was detected and marked from the hand and wrist image taken using a near infrared camera. In this context, a software based on deep learning has been developed including ResNet50 and YOLO algorithms. The performance of the system was evaluated comparatively by testing the segmentation and classification process on 220 wrist images taken from 22 human wrists.

2. Deep neural networks (DNN)

Deep neural networks (DNN) can recognize objects without being affected by different properties such as different positions, directions and camera angles and environmental factors such as lighting. A deep learning algorithms are trained on tagged images. An architecture of DNN is shown in Figure 2.

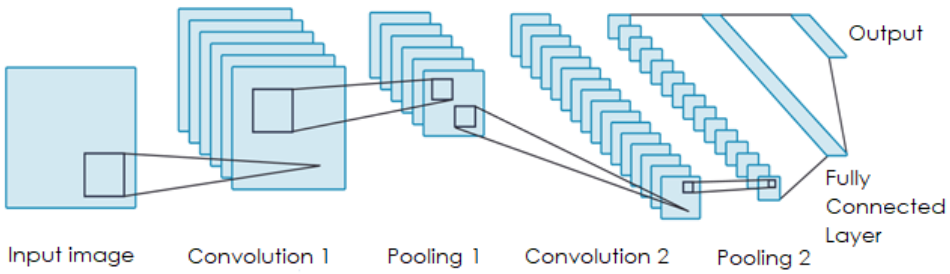


Figure 2. *A Deep neural networks architecture.*

The input usually consists of images or signals. In the convolution layers, the filtering process is applied to the 3-dimensional matrices in the previous layer. The number of filters used constitutes the depth (the size of the 3rd dimension) of the convolution layer. In the pooling layer, size reduction is applied. In the section called fully connected layer, classical artificial neural network operations are performed. The output can be defined as a vector with the length of the defined class. The number of layers in the architecture and the filter size and number can be changed by the user to suit the application. In addition, the performance may vary depending on the number of images and iterations to be used for training. If the number of training iterations is too high, both the training time increases and the model moves towards memorization. As the rate of misclassification in data that the model has not seen before will increase as a result of memorization, accuracy decreases. In the study, it was tried to determine the wrist region from the images taken by using the ResNet50 and YOLO architectures.

2.1. ResNet50

The Residual Network (also known as ResNet) uses redundant blocks with multiple layers to reduce training error. Typical ResNet models are implemented with double- or triple- layer skips that contain nonlinearities (ReLU) and batch normalization in between. During training, the weights adapt to mute the upstream layer, and amplify the previously-skipped layer. In the context of residual neural networks, a non-residual network may be described as a plain network [8, 9].

The standard ResNet architecture consists of 152 layers and has a higher classification success than other DNN architectures. The ResNet50 architecture used in this work, on the other hand, consists of 34 layers. Both layer blocks in the network are replaced by a three-layered residual block. The residual block characterization can be given as:[10, 11]

$$a^{[l+2]} = g(a^{[l]} + z^{[l+2]} \quad (1)$$

Here, a is the activations (outputs) of neurons in layer l, g is the activation function for layer l. In this equation, l is used for clarity.

2.2. You Only Look Once (YOLO)

YOLO is an algorithm for object detection using convolutional neural networks (CNN). It stands for "You Only Look Once", which means "Just Look Once". The reason for choosing this name is that the algorithm is fast enough to detect objects in one go. The most important feature that distinguishes YOLO from other algorithms is its ability to detect objects in real time. YOLO uses Darknet library and has a very fast running architecture. The reason it is so fast is that it can guess objects and coordinates by passing any picture through the neural network in one go. In other words, YOLO takes an approach that solves a single regression problem while performing the estimation process [12, 13].

When we look at YOLO's algorithm, we first see that it divides the input picture into grids of certain sizes (3x3, 5x5, 17x17 etc.). It is then checked whether the searched object is within the grid. The object detected in the similarity check is scored and marked. Then non maximum suppression is applied, which evaluates the score of each marked object, and those below a certain confidence score are separated [14, 15].

3. The proposed method

There are 3 main elements in the proposed approach. These can be listed as collecting image data set, labeling the wrist area and performing deep learning based training / testing processes. While performing the deep neural network for detection and classification of wrist images, Yolo and ResNet50 architectures were implemented. The flow chart of the system designed within the scope of this study is shown in Figure 3.

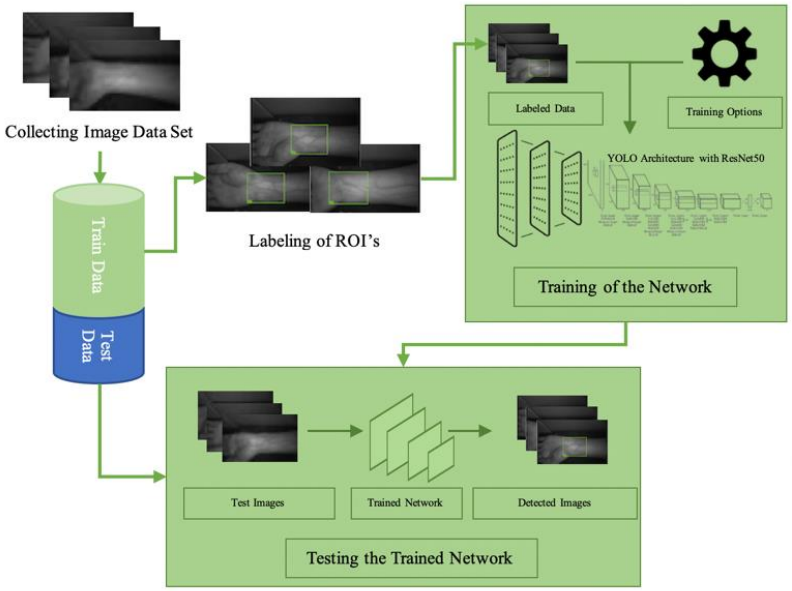


Figure 3. *Flow chart of the designed system*.

The study is primarily provided within the scope of the data collection. Two wrist images of 11 people were taken with a near infrared (NIR) camera. In the study, a NIR camera with a resolution of 752x480 pixels, a monochrome (gray level) and an infrared light source operating at 880nm wavelength was used to capture the images. By taking 10 images for both wrists of each person, a total of 220 images were provided in the data set. The images obtained are gray level with a resolution of 752x480 pixels in RGB format. Images consist of 8 bits as color depth. The sample images used in the designed system are shown in Figure 4. In the second stage, the data set is divided into training and test data. As used frequently in the literature, the obtained images were determined as homogeneous 70% (154 images) for training and the remaining 30% for testing.

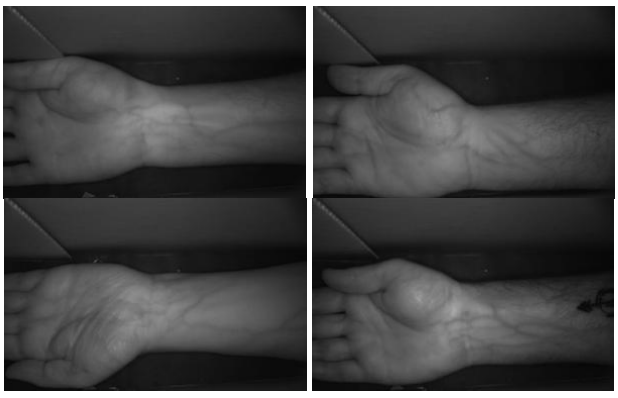


Figure 4. *Image set examples.*

In the next stage, Region of Interest (ROI) areas are labeled on the images (70%) allocated for training. In order to define the wrist print region, ROI is labeled as 224x224 dimensioned boxes where the veins on the wrist are clearly visible. Image Labeler plugin of Matlab software was used for this labeling process. An interface of Image Labeler is shown in Figure 5. The remaining test data were transferred to the test stage without labeling.

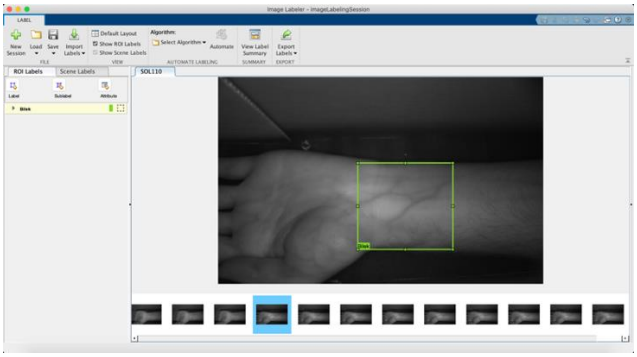


Figure 5. *Image labeling process.*

Data sets were created with two different labeling methods. In the first data set, only the wrist regions were labeled and it was aimed to segment the wrist region from the image (One class=' Wrist'). In the second data set, the wrist images were labeled according to 22 classes and these classes were tried to be predicted (Twenty two classes=' Wrist 1',' Wrist 2'... ' Wrist 22'). In the study, both segmentation and classification performance in images were tried to be measured with these two data sets. After the data labeled and prepared, the labeled data was trained with Deep Neural Network - (DNN) with the software realized within the scope of the study. In the training phase of the network, the images labelled were given for 1 class (wrist) of

output as an introduction to the neural network built from YOLO architecture powered by ResNet50 and training has been done. Layer structure of the designed network is given in Table 2. As seen in Table 2, the input images are 224x224 in size and have 3 layers. After the convolution process, the sub-image with 112x112x64 activation was obtained. After the activation process, the activation number of 55x55x64 was obtained by max pooling process. Convolution was made using Yolo2 in the output layer. For the other data set used in the study (Twenty two classes dataset), the 150th output layer (YOLO v2 Output) has 22 neurons.

Table 2. *Layer structure of the designed network*

	Name	Type	Activations	Learnables	Total Learnables
1	**Input_1** 224x224x3 images	Image Input	224x224x3	-	0
2	**Conv1** 64 7x7x3	Convolution	112x112x64	Weights 7x7x3x64 Bias 1x1x64	9472
3	**bn_conv1** Batch Normalization	Batch Normalization	112x112x64	Offset 1x1x64 Scale 1x1x64	128
4	**activation_1_relu** ReLU	ReLU	112x112x64	-	0
5	**max_pooling2d_1** 3x3 max pooling	Max Pooling	55x55x64	-	0
6	**res2a_branch1** 256 1x1x64 conv	Convolution	55x55x256	Weights 1x1x64x256 Bias 1x1x256	16640
7	**bn2a_branch1**	Batch Normalization	55x55x256	Offset 1x1x256 Scale 1x1x256	512
8	**res2a_branch2a** 64 1x1x64 conv	Convolution	55x55x64	Weights 1x1x64x64 Bias 1x1x64	4160
9	**bn2a_branch2a**	Batch Normalization	55x55x64	Offset 1x1x64 Scale 1x1x64	128
10	**activation_2_relu** ReLU	ReLU	55x55x64	-	0
			…		
141	**activation_40_relu** ReLU	ReLU	14x14x1024	-	0
142	**yolo2Conv1** 1024 3x3x1024 conv	Convolution	14x14x1024	Weights 3x3x1024x1024 Bias 1x1x1024	9438208
143	**yolo2Batch1** Batch Normalization	Batch Normalization	14x14x1024	Offset 1x1x1024 Scale 1x1x1024	2048
144	**yoloRelu1** ReLU	ReLU	14x14x1024	-	0
145	**yolo2Conv2** 1024 3x3x1024 conv	Convolution	14x14x1024	Weights 3x3x1024x1024 Bias 1x1x1024	9438208
146	**yolo2Batch2** Batch Normalization	Batch Normalization	14x14x1024	Offset 1x1x1024 Scale 1x1x1024	2048
147	**yoloRelu2** ReLU	ReLU	14x14x1024	-	0
148	**yoloClassConv** 4 1x1x1024 conv	Convolution	14x14x42	Weights 42 Bias 1x1x42	43050
149	**yolo2Transform**	Yolo V2 Transform Layer	14x14x42	-	0
150	**yolo2OutputLayer**	Yolo V2 Output Layer	-	-	0

ResNet50 was created in a 150-layered structure because of integration with YOLO. Since the input data of the ResNet network is 224x224x3, the images in the data set have been adapted to this resolution. The training parameters of the designed deep neural network were given in Table 3. The output size is 1 for segmentation of the wrist region and the output size is 22 for classification of the wrist data. The optimizer used in this study is Stochastic gradient descent with momentum. It is an iterative method for optimizing an objective function with suitable smoothness properties. The iteration number for segmentation and classification process is limited to 1300 and 2600, respectively.

Table 3. *Training parameters for the proposed method*

Option	**Segmentation Data Set**	**Classification Data Set**
Epoch (Iteration)	100 (1300)	200 (2600)
Batch Size	16	16
Learning Rate	0,001	0,001
Optimizer	SGDM	SGDM
Input Size	224x224x3	224x224x3
Output Size	**1 (Wrist)**	**22 (Wrist)**
Verbose	True	True
Verbose Frequency	1	1
Shuffle	Never	Never

The computer on which the designed model is trained has the following hardware configuration: Intel i7-2600 3.40 GHz processor running under 64 Bit Windows 10 operating system, 22 GB RAM, GTS450 GDDR5 1GB 128Bit Nvidia GeForce DX11 Graphics Card hardware. In addition, Matlab software was used in coding, training and testing of the model.

The change of Root Mean Square Error (RMSE the square root of the difference between the value obtained and the value obtained during the training of the network) and LOSS (expressing how different the estimate made by the model from the actual value) values are given in Figure 6. In addition, the training time of the segmentation network is 03:46:10, the resulting minimum RMSE value of the network is 0.07 and LOSS value is 5.2e-03. The training time of the classification network is 05:27:30, the resulting minimum RMSE value of the network is 0.077 and LOSS value is 0.088.

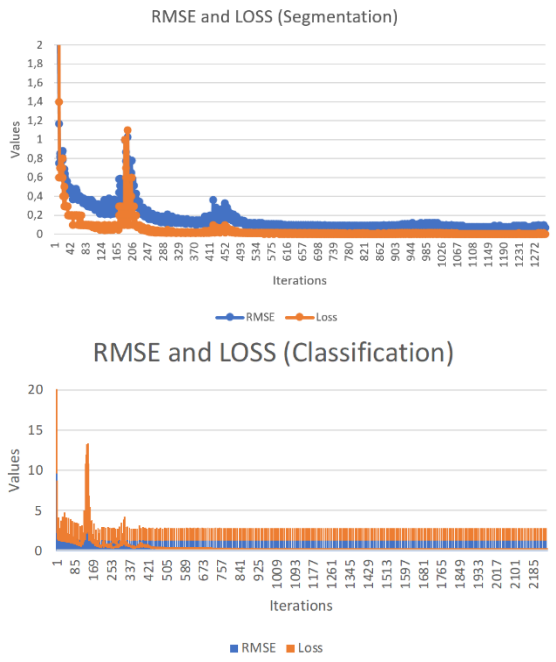


Figure 6. *Change of RMSE and LOSS values during training (Segmentation and Classification).*

4. Results and discussion

For the evaluation of the study, the test process was carried out by using images (30%) that were not shown to the network model whose training was completed. Test images were given as an input to the network and the wrist areas were determined. In determining the region, the minimum value of similarity ratio was taken as 0.70 and the highest similarity value was determined as the result. Sample test images including the detected and marked boxes are given in Figure 7.

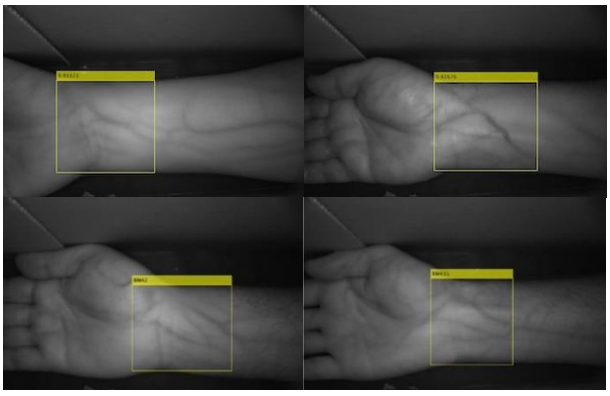


Figure 7. *Detection results*

The percentage values for segmentation of the wrist area, which are formed by input (total of 66 images in the test process), are shown in Figure 8(a). The process of determining the wrist area on an image takes approximately 0.42 seconds. It took approximately 30 seconds to determine the ROI areas of all test images and the ROI's were saved as a file.

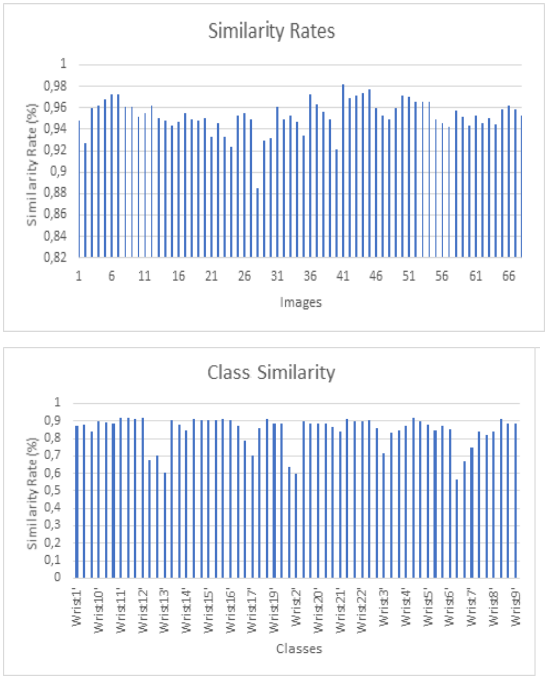


Figure 8. *Similarity rates for segmentation (a) and classification (b)*

The percentage values for classification, which are formed by input (total of 66 images in the test process), are shown in Figure 8(b). The process of classification on an image takes approximately 0.37 seconds. It took approximately 30 seconds to determine the ROI areas of all test images and the ROI's were saved as a file.

5. Conclusion

In this study, a deep learning based classification approach is implemented for wrist print recognition. YOLO architecture with ResNet50 is used for training process in deep learning. As a result of the study, it was seen that the wrist regions were segmented correctly in all test images and the average value of the obtained similarity rates was 95.26%. As a result of the classification process, 61 out of 66 images were classified correctly. Therefore, 92.43% classification successes were obtained. Therefore, it can be said that the deep learning architectures ResNet50 and YOLO are effective in the segmentation of the wrist region.

The obtained error rate for segmentation and classification is 7.57% in our approach. When we compare the values we obtained as a result of this study with previous studies on wrist recognition, it is seen that an average performance is achieved. The main reason for this is the small number of images (7 images per wrist) used for training in the study. It is known that the higher the number of entrance images in deep learning approaches gives the higher the classification performance. Therefore, it is thought that the success will increase by increasing the number of samples.

Acknowledgements

This study was presented as an oral presentation at the 4th International Conference on Engineering Technologies (ICENTE) held in Konya on 19-21 November 2020.

References

[1] Hartung D., Olsen M.A., Xu H., Busch C., "Spectral minutiae for vein pattern recognition," in 2011 International Joint Conference on Biometrics (IJCB), IEEE, (2011), 1-7.

[2] Wang L., Leedham G., Cho S.-Y., "Infrared imaging of hand vein patterns for biometric purposes," IET computer vision, 1, 3, (2007), 113-122.

[3] Das A., Pal U., Ballester M.A.F., Blumenstein M., "A new wrist vein biometric system," in 2014 IEEE Symposium on Computational Intelligence in Biometrics and Identity Management (CIBIM), IEEE, (2014), 68-75.

[4] Nikisins O., Eglitis T., Anjos A., Marcel S., "Fast cross-correlation based wrist vein recognition algorithm with rotation and translation compensation," in 2018 International Workshop on Biometrics and Forensics (IWBF), IEEE, (2018), 1-7.

[5] Uriarte-Antonio J., Hartung D.,. Pascual J.E.S, Sanchez-Reillo R., "Vascular biometrics based on a minutiae extraction approach," in 2011 Carnahan Conference on Security Technology, IEEE, (2011), 1-7.

[6] Chen H., Lu G., Wang R., "A new palm vein matching method based on ICP algorithm," in Proceedings of the 2nd International Conference on Interaction Sciences: Information Technology, Culture and Human, (2009), 1207-1211.

[7] Garcia-Martin R., Sanchez-Reillo R., "Wrist Vascular Biometric Recognition Using a Portable Contactless System," Sensors, 20, 5, (2020), 1469.

[8] Wikipedia. "Residual neural network." Wikipedia. https://en.wikipedia.org/wiki/Residual_neural_network (accessed 15.11.2020, 2020).

[9] He K., Zhang X., Ren S., Sun J., "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition, (2016), 770-778.

[10] Amidi A., Amidi S., Vlachakis D., Megalooikonomou V., Paragios N., Zacharaki E.I., "EnzyNet: enzyme classification using 3D convolutional neural networks on spatial representation," PeerJ, 6, (2018), e4750.

[11] Amidi S., Amidi A., Vlachakis D., Paragios N., Zacharaki E.I., "Automatic single-and multi-label enzymatic function prediction by machine learning," PeerJ, 5, (2017), e3095.

[12] Kuyumcu B., Buluz B., Kömeçoğlu Y., "Author Identification in Turkish Documents with Ridge Regression Analysis," in 2019 27th Signal Processing and Communications Applications Conference (SIU), 2019: IEEE, pp. 1-4.

[13] Redmon J., Farhadi A., "YOLO9000: better, faster, stronger," in Proceedings of the IEEE conference on computer vision and pattern recognition, (2017), 7263-7271.

[14] Nguyen C.C.et al., "Towards real-time smile detection based on faster region convolutional neural network," in 2018 1st International Conference on Multimedia Analysis and Pattern Recognition (MAPR), IEEE, (2018), 1-6.

[15] Redmon J., Divvala S., Girshick R., Farhadi A., "You only look once: Unified, real-time object detection," in Proceedings of the IEEE conference on computer vision and pattern recognition, (2016), 779-788.

MANAS Journal of Engineering
ISSN 1694-7398 | e-ISSN 1694-7398

Volume 9, Issue 1, (2021) Pages 37-44
https://doi.org/10.51354/mjen.854265

A Gas Sensor Design and Heat Transfer Simulation with ZnO and TiO_2 Sensing Layers

Gozde Konuk Ege[1,*], Huseyın Yuce[2], Garıp Genc[3]

[1] Istanbul Gedik University, Istanbul, 34913, Turkey, gozde.konuk@gedik.edu.tr, ORCID: 0000-0001-7349-0416

[2] Marmara University, Istanbul, 34722, Turkey, huseyin@marmara.edu.tr, ORCID: 000-0001-5525-7733

[3] Marmara University, Istanbul, 34722, Turkey, ggenc@marmara.edu.tr, ORCID: 0000-0001-7711-3845

ABSTRACT

Micro Electro-Mechanical System (MEMS) based devices offer innovative approaches in sensor technologies with the advantages of high efficiency and miniaturization. The most important stage in the development of new generation MEMS-based devices is the design and optimization stage. However, device design and optimization processes are developed in a laboratory by empirical approaches. This causes time loss and creates an unnecessary waste of resources. In this study, it is aimed to design and analyze two gas sensors based on ZnO and TiO_2 sensing layers. Electro-thermal analysis of the sensor structure was carried out at room temperature and high temperature (294,15K-573,15K) and heat transfer parameters were compared. According to the simulation results, it is obtained that, as the applied temperature increases to the sensor, the temperature over the sensing layer increases linearly. It is compatible with the literature. The temperature on the ZnO surface increases to three times the TiO_2 surface temperature. The heat transfer results obtained will be used as a guide for device design and optimization in future works. In this way, as a result of numerical analysis, a MEMS-based device will be produced with high accuracy. Thus, time and resources will be saved.

ARTICLE INFO

Research article

Received: 5.01.2021
Accepted: 30.04.2021

Keywords:
gas sensor,
sensing layer,
finite element method
(FEM,heat transfer)

*Corresponding author

1 Introduction

Rapidly increasing environmental pollution and consequently increasing health problems have become one of the most important problems of today. In addition, different gases are frequently used in the manufacturing industry. This situation directly affects to human health. The integration of smart devices developed for the benefit of humanity into our daily life is becoming more and more popular with decades due to their usefulness. Gains in MEMS technology also has been affected sensor technologies. Sensors produced in line with micro technologies have advantages in terms of cost, miniaturization, and sensitivity as smart sensors. In this view, the development of gas detection devices is important for widespread uses, such as monitoring gas accumulation, medical diagnostics, food quality assurance and the safety of industrial processes or domestic systems. Today, in parallel with the technological progress in nano-electronic structuring and the point it has reached, gas sensor technologies continue to progress with similar acceleration [1–5].

Micro/nano technology is a multi-disciplinary technology because of using material science, structure-control design and finite-element method (FEM). Material technology allows the use of new generation materials and semiconductor technologies in miniature device development. Moreover, structural/control design and analysis methods offer high accuracy and analysis advantages with high accuracy.

It was observed for the first time in 1953 that the conductivities of semiconductor materials are changed by gas absorption [6]. Subsequently, the first semiconductor gas sensor patented in 1962 [7]. Afterwards, gas-sensing devices are researched and examined worldwide. Researches on gas detection mechanisms is focused on designing gas detection elements that can detect gases harmful to human health and nature with high performance detection materials as a consequence various gas detection technologies have been developed for use in gas detection processes [8–13].

Table 1. *Gas detection methods used in gas detection processes*

Gas sensors related to detecting material	Gas sensors related to detecting mechanism
Metal oxide semiconductors	Resistance Change
Carbon nanotubes	Optical methods
Polymers	Acoustic methods
	Gas chromatography
	Calorimetric methods

Gas detection processes can be examined in two classes, according to the sensing material and the sensing mechanism, as shown in Table 1. In gas sensors; Metal oxide semiconductors, carbon nanotubes and polymers are used as the material in the gas detection layer. On the other hand, resistance changing, optical methods, acoustic methods (eg; SAW sensor), calorimetric method and gas chromatography can be used as the detection mechanism. Metal oxide-based gas sensors attract researchers' attention thanks to their high detection capability, repeatability and simple manufacturing techniques. MO_X-based gas sensors are used to sense NH_3, H_2, NO_2, H_2S, CO, CO_2, SO_2, O_3, N_2, VOCs, LPG gases [14].

NH_3, H_2, HCl, NO_2, H_2S, CO, CO_2 gases threat human health depending on their conditions of use and release rates of these gases. Table 2 presents that the areas of use or release of these gases and their danger. For example; NH_3 gas appears in the chemical, food and health industry and causes a toxic effect. NO, HCl, CO_2 gases cause respiratory disorders. The release of high levels of H_2S and CO, H_2 gases can result in death.

The working principle of MO_X gas sensors are usually based on the resistivity change of the semiconductor, depending on the gas absorption. When gas analytes come into contact with the sensing layer, the adhesion of oxygen atoms occur on the surface and semiconductor sensing material reacted to the gas analytes. In this way, sensing layer resistivity is decreased according to the analyte concentration. However, the worst feature of MOx-based gas sensors is that they have a high operating temperature. This situation requires high power consumption and an external heat source [33]. Therefore, many studies are carried out to obtain high operating performance at low temperatures. In this study, two MOx-based gas sensors model are designed.

Table 2. *According to the encountered area, and their dangers various hazardous gases*

Hazardous Gases	**The encountered Area**	**Dangers**
NH_3 [15]	Chemical and food industry, medical diagnosis	Toxic effect, vomiting, headache
NO [16–19]	Combustion of fossil fuels, power plants and vehicle engines	Respiratory diseases, asthma
NO_2 [20]	Automobile exhaust fumes, nitric acid production, coal and fuel combustion	Acid rain, environmental pollution
H_2 [21–25]	Renewable energy sources, transport systems and biomedical devices	Explosive, flammable
H_2S [26–29]	Oil/natural gas, geothermal energy and bacterial decomposition	Occupational diseases, death in high concentration
HCl [30]	Semiconductor and chemical industry	Toxic effect, skin burns, respiratory disorders
CO [31]	Fire events	Toxic effect , death
CO_2 [32]	Cellular respiration and burning of fossil fuels	Respiratory disorders, explosion, Global warming, explosion

One of the gas sensors consists of TiO_2 as a sensing layer and another is consist of ZnO. Three-dimensional design and simulation are done by using SolidWorks and COMSOL Multiphysics software. The electro-thermal analysis is done to observe the surface temperature and electrical dispersion along the sensing layer. The effect of TiO_2 and ZnO semiconductors heat transfer is evaluated over the sensing layer at room temperature and different high temperatures.

2 Structure design

Gas sensors are smart devices used to sense many other gases with concentrations ranging from ppm and ppb in the presence of other low concentrations gases [34]. Although it is considered to be an apparently simple operating principle, the gas detection mechanism is quite complicated. MEMS-based gas sensors consist of two main layers as shown in Fig.1. Gas sensing performance and especially sensitivity are controlled these two layers. A chemical or biochemical interface is required to obtain sufficient sensitivity and selectivity for the analyte. This interface is called the sensing layer, this is the top layer. Gas detection mechanism is related to the change of one or more physical properties of the sensing material such as mass, elastic hardness, viscosity, electrical conductivity, and electrical permeability. The conductivity of the sensing material is expected to change according to the gas concentration. It is also desired that this change in conductivity be reversible. The performance of a gas sensor depends on the sensor configuration as well as the interaction of the sensing material with the gas. It is important to understand the interaction between the sensor material and the analyte molecules of basic sensor parameters such as sensitivity and selectivity. If this interaction is weak; sensitivity and selectivity are weakened when the sensor shows good recycling or recycling is weakened while showing good sensitivity and selectivity [35]. Another layer is the transducer layer. This layer converts the sensing analyte signals from the chemical interaction occurring in the sensing layer into an electrical signal.

Figure 1. *MEMS-based gas sensors layer structure*

Sensor specifications, for instance sensitivity, selectivity, response and recovery times, are affected by the properties of the sensing material. In this study, TiO2 and ZnO are preferred as a sensing material. ZnO has high chemical and thermal stability, piezoelectric properties. These provide excellent detection response. It has a low cost. It offers easy production and non-toxic. Thanks to these advantages, they are preferred in gas sensor applications [1]. TiO2 is also stable, low cost and non-toxic semiconductor material. It has unique electrical, optical and catalytic gas sensing material because of its wide bandgap. In smart device applications, the band structure provides superior electronic and optical properties [36]. The designed gas sensors structure is presented in Fig.2.

$LiNbO_3$ wafer is used as a substrate. On the substrate, gold electrodes are evaporated in a vacuum for electrical conductivity. Above the gold electrodes, the sensing layer is sputtered.

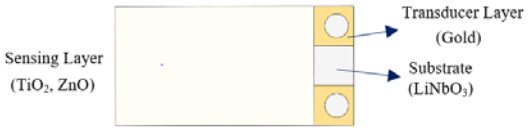


Figure 2. *Gas sensor structure*

Mechanical properties are given in Table 3. The hot plate to be used as a high heat source will be placed where the gas sensor socket in the gas detection chamber.

Table 3. *Mechanical properties of sensor structure*

Material	Size		Thermal Conductivity (W/(m.K))	Density (kg/m3)	Young's Modulus (GPa)
	Length x Wide (mm)	Thickness (mm)			
LiNbO3[37]	6x3	0.5	5.6	4640	170
Gold [38]	6x1	1x10-4	317	19300	70
TiO_2 [39]	5x3	5x10-5	8.9	4230	230
ZnO [40]	5x3	5 x10-5	60	5676	210

3 Electrical conductivity and heat transfer analysis

The heat transfer analysis and joules heating simulation are important for observing the homogeneous heat dispersion on the surface. Homogeneous heat dispersion is one of the determining factors of sensor sensitivity. Seeing that metal oxide gas sensors work according to the resistance change principle, it is expected to be homogeneous and constant at the electrical voltage on the

sensing surface. For electrical conductivity simulation, the electrical field dispersion (E) is determined using Eq.1.;

$$E = -\nabla v \quad (1)$$

The electrical resistivity of most materials changes with temperature. To calculate electrical resistivity, a linear approximation is typically used. For electrical resistivity, (ρ) in certain temperature is determined by Eq.2;

$$\rho(T) = \rho_0(1 + \alpha(T - T_0)) \quad (2)$$

In Eq.2; ρ(T) is electrical resistivity in a certain temperature, ρ0 is electrical resistivity at 20 ºC, α is temperature coefficient at 20 ºC, T0 is 20 ºC and T is a certain temperature.

According to the electric field (E); electrical conductivity of the sensing material,(σ), current density (J) and joule

$$J = \sigma E \quad (3)$$
$$Qj = \nabla J \quad (4)$$

heating power (Qj) are calculated by Eq.3., and Eq.4.; Moreover, the heat transfer is calculated with Eq.5. The heat transfer depends on the heat flux, thermal conductivity, absolute depends on the thermal conductivity, temperature difference, and length of the body. In this equation, q, Δt, and k represent the heat flux, the temperature difference, and the thermal conductivity respectively [41] .

$$q = \frac{k.\Delta t}{L} \quad (5)$$

4 Results and discussion

4.1 Electrical dispersion

The sensing layer electrical dispersion of the gas sensor from room temperature to 573,15K is analyzed. A 9V initial voltage is applied to the gas sensor. In accordance with the resistivity of the sensor material, the average voltage above the sensing layer surface is observed and it is presented in Fig.3. At the ZnO surface, electrical dispersion has three different distinct regions.

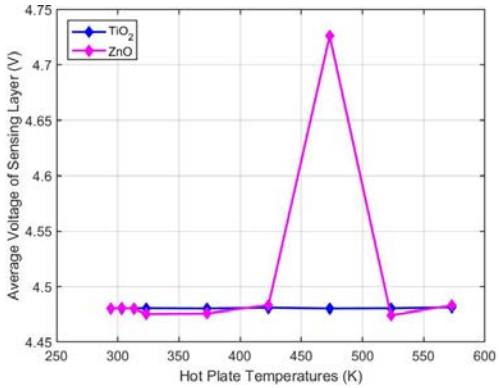


Figure 3. *Average voltage above the sensing layer (TiO_2, ZnO)*

In region I, the average electrical potential is observed steadily from room temperature to 423,15K. ZnO exhibits a negative temperature coefficient of resistance (NTCR) behavior in this region. At second region from 423,15K to 523,15K ZnO indicates a positive temperature coefficient of resistance (PTCR) behavior.

Resistivity level is the lowest in this region. In region III from 523,15K to 573,15K resistivity increases again and ZnO shows a negative temperature coefficient in this region. There are several studies about the temperature-dependent resistivity of ZnO according to NTCR and PTCR behavior.

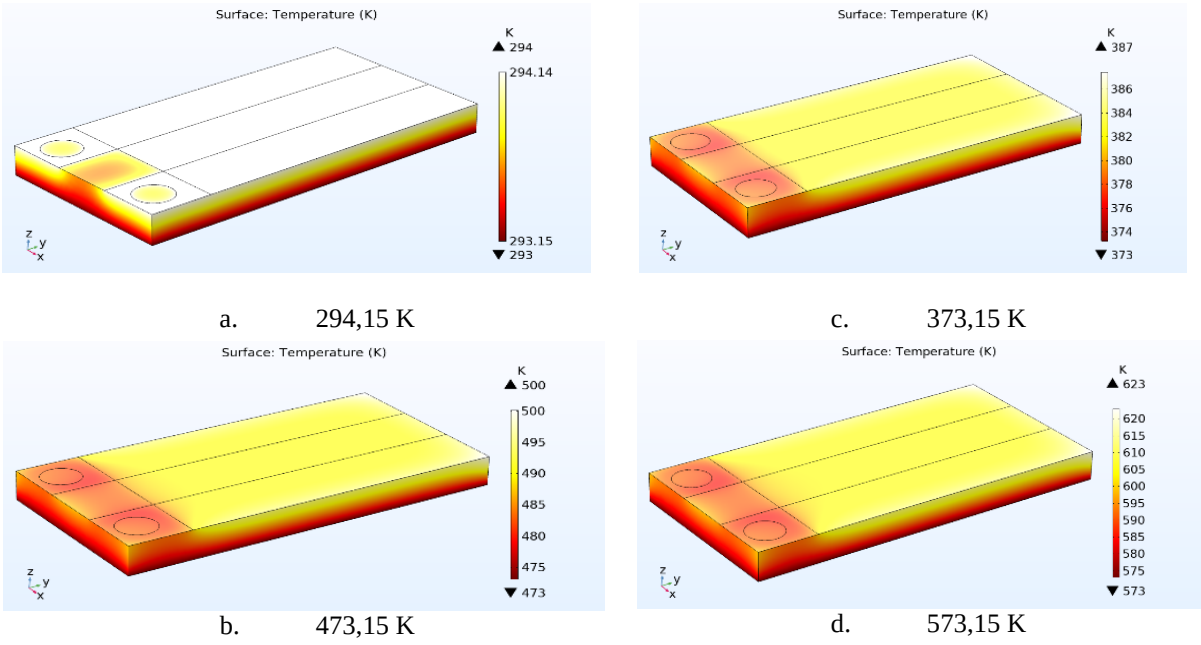
The simulation results obtained are compatible with the literature [42], [43]. On the other hand, the TiO_2 surface electrical dispersion is observed steadily. TiO_2 shows a negative temperature coefficient from room temperature to 573,15K.

4.2 Temperature dispersion

In this simulation; absolute pressure is 1 atm. First of all, the room temperature condition is done and heat dispersion is observed. Then, high temperature conditions (323,15K, 373,15K, 423,15K, 473.15K, 523,15K, 573,15K) are created by hot plate respectively. An initial temperature (T_0=room temperature) is 294,15K and the external temperatures (T_{ext}=hot plate temperature) were applied by increasing the temperature step by step from 323,15K to 573,15K. Fig.4, and Fig.5 shows the temperature profiles of TiO_2 and ZnO sensing layers with a base substrate.

A uniform sensing layer temperature is a requirement for sensitive operation of the sensor. Figure 4 and 5 shows uniform temperature dispersion on the sensing surface at different operating temperatures. Furthermore, Fig.6. shows the comparison of surface temperature at constant electrical voltage 9V. While the temperature on the TiO_2 surface at 473,15K was 489.56 K and 569.86 K on the

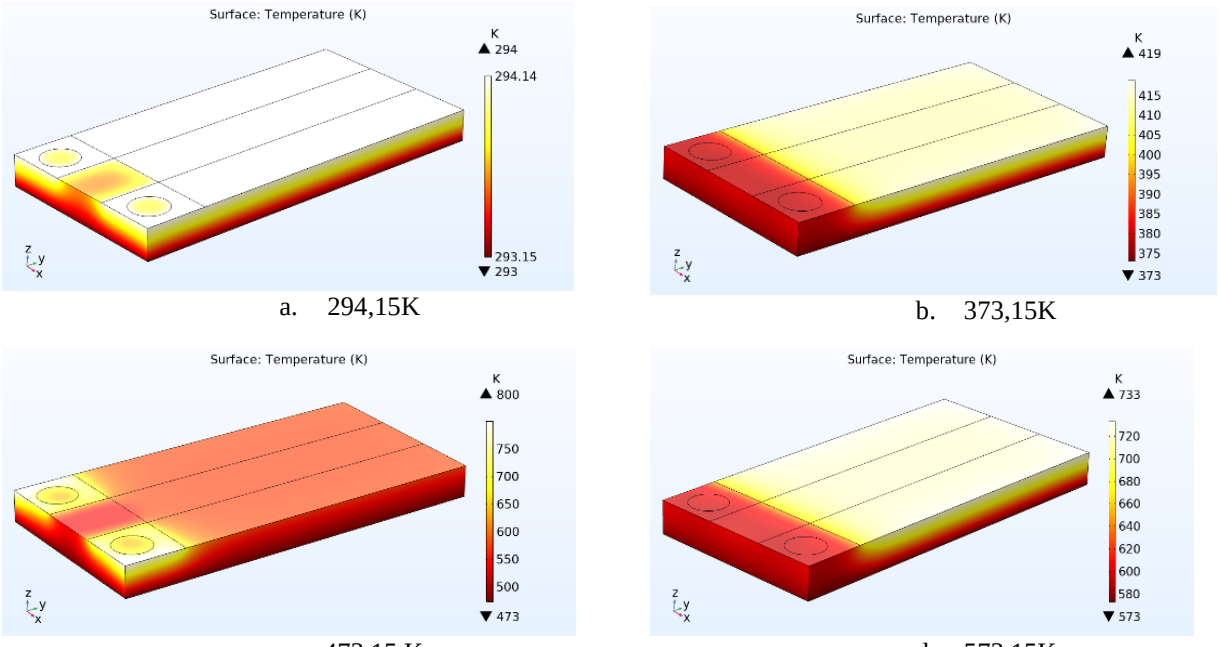
ZnO surface, the temperature on the TiO_2 surface at 573,15K was 603.41 and 715.51 K on the ZnO surface.



a. 294,15 K c. 373,15 K

b. 473,15 K d. 573,15 K

Figure 4. *Temperature dispersion in TiO_2 layer, for the temperature 294,15K (a), 373,15K (b), 473,15K (c), 573,15 K (d)*



a. 294,15K b. 373,15K

c. 473,15 K d. 573,15K

Figure 5. *Temperature dispersion in ZnO layer, for the temperature 294,15K (a), 373,15K (b), 473,15K (c), 573,15K (d)*

The comparison in Fig.6. depicts that sensing surface temperature increase on both surfaces; increased proportionally from the room temperature to 423.15 K.

When 423.15 K exceeds; the temperature increase on ZnO surface gained acceleration compared to TiO_2 surface. While the temperature on the TiO_2 surface at 473,15K was 489.56 K and 569.86 K on the ZnO surface, the temperature on the TiO_2 surface at 573,15K was 603.41 and 715.51 K on the ZnO surface.

According to the simulation results, the optimum operating temperature was determined of ZnO and TiO_2 thin films. Sensing surface temperature rises above at the 473.15K, 523.15K, 573.15K operating temperature.

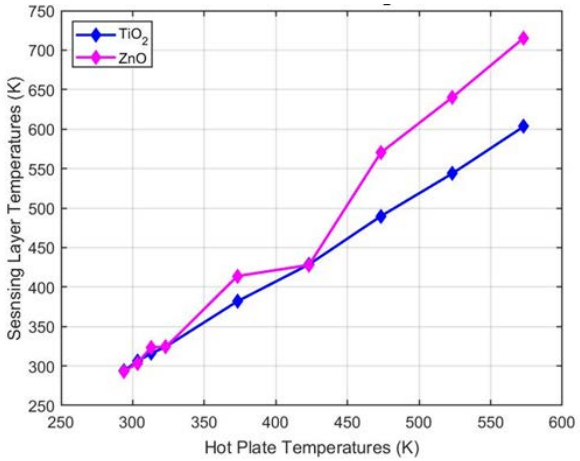


Figure 6. *Sensing layer temperature at different hot plate temperatures*

Consequently, heating inevitably occurs on the surfaces. In this case, the optimum working temperature for both thin-film structures with a thickness of 0.5 µm was determined as 423.15 K.

5 Conclusion

In this study, heat transfer analysis of the MOx-based gas sensor is discussed. Two sensing layer structures that used TiO_2 and ZnO as a material are designed and simulated. To achieve uniform heat dispersion above TiO_2 and ZnO sensing layer at room temperature and high temperature are analyzed by using FEM with COMSOL Multiphysics. The temperature profiles of sensing layers are analyzed and uniform heat dispersion was achieved in both thin films. As stated in simulation results, an optimum working temperature was determined for ZnO and TiO_2 thin films.

Our future works will be focused on the designed metal-oxide/polymer-based gas sensor at room temperature and high temperature and compared simulation results between them. Additionally, we aim to design and fabricate a gas sensor using metal oxide and metal oxide/polymer sensing layer.

Acknowledgment

This work was supported by Scientific Research Projects Coordination Unit of Istanbul Gedik University. Project number "GDK201905-14"

References

[1] V. S. Bhati, M. Hojamberdiev, and M. Kumar, "Enhanced sensing performance of ZnO nanostructures-based gas sensors: A review," *Energy Reports*, no. xxxx, 2019.

[2] K. H. Rahman and A. K. Kar, "Titanium-di-oxide (TiO2) concentration-dependent optical and morphological properties of PAni-TiO2 nanocomposite," *Mater. Sci. Semicond. Process.*, vol. 105, no. April 2019, p. 104745, 2020.

[3] R. Kumar and R. Ghosh, "Selective determination of ammonia, ethanol and acetone by reduced graphene oxide based gas sensors at room temperature," *Sens. Bio-Sensing Res.*, vol. 28, no. January, p. 100336, 2020.

[4] L. Kumar, I. Rawal, A. Kaur, and S. Annapoorni, "Flexible room temperature ammonia sensor based on polyaniline," *Sensors Actuators, B Chem.*, vol. 240, pp. 408–416, 2017.

[5] B. Mondal, M. S. Meetei, J. Das, C. Roy Chaudhuri, and H. Saha, "Quantitative recognition of flammable and toxic gases with artificial neural network using metal oxide gas sensors in embedded platform," *Eng. Sci. Technol. an Int. J.*, vol. 18, no. 2, pp. 229–234, 2015.

[6] W. H. Brattain and J. Bardeen, "Surface properties of germanium," *Bell Syst. Tech. J.*, vol. 32, no. 1, pp. 1–41, 1953.

[7] T. Seiyama, A. Kato, K. Fujiishi, and M. Nagatani, "A new detector for gaseous components using semiconductive thin films.," *Anal. Chem.*, vol. 34, no. 11, pp. 1502–1503, 1962.

[8] C. Wartelle, N. Pereira Rodrigues, M. Koudelka-Hep, and F. Bedioui, "Amperometric fluidic microchip array sensing device for nitric oxide

determination in solution," *Mater. Sci. Eng. C*, vol. 26, no. 2–3, pp. 534–537, 2006.

[9] S.-M. Park, S.-L. Zhang, and J.-S. Huh, "NO Sensing Characteristics of ZnO Nanorod Prepared by Ultrasound Radiation Method," *Korean J. Mater. Res.*, vol. 18, no. 7, pp. 367–372, 2008.

[10] Z. Zhang, C. Yin, L. Yang, J. Jiang, and Y. Guo, "Optimizing the gas sensing characteristics of Co-doped SnO 2 thin film based hydrogen sensor," *J. Alloys Compd.*, vol. 785, pp. 819–825, 2019.

[11] F. Rathgeb and G. Gauglitz, "Optical gas sensors in analytical chemistry: Applications, trends and general comments," *Encycl. Anal. Chem. Appl. Theory Instrum.*, 2006.

[12] G. Korotcenkov and B. K. Cho, "Metal oxide composites in conductometric gas sensors: Achievements and challenges," *Sensors Actuators, B Chem.*, vol. 244, pp. 182–210, 2017.

[13] C. Liu *et al.*, "A high-performance flexible gas sensor based on self-assembled PANI-CeO2 nanocomposite thin film for trace-level NH3 detection at room temperature," *Sensors Actuators, B Chem.*, vol. 261, pp. 587–597, 2018.

[14] S. Mahajan and S. Jagtap, "Metal-oxide semiconductors for carbon monoxide (CO) gas sensing: A review," *Appl. Mater. Today*, vol. 18, p. 100483, 2020.

[15] Y. De Wang, X. H. Wu, Q. Su, Y. F. Li, and Z. L. Zhou, "Ammonia-sensing characteristics of Pt and SiO2 doped SnO2 materials," *Solid. State. Electron.*, vol. 45, no. 2, pp. 347–350, 2001.

[16] S. Ummartyotin and H. Manuspiya, "A critical review on cellulose: From fundamental to an approach on sensor technology," *Renew. Sustain. Energy Rev.*, vol. 41, pp. 402–412, 2015.

[17] M. R. Vilar *et al.*, "Development of nitric oxide sensor for asthma attack prevention," *Mater. Sci. Eng. C*, vol. 26, no. 2–3, pp. 253–259, 2006.

[18] H. BISGAARD, L. LOLAND, and J. A. N. H. ØJ, "NO in exhaled air of asthmatic children is reduced by the leukotriene receptor antagonist montelukast," *Am. J. Respir. Crit. Care Med.*, vol. 160, no. 4, pp. 1227–1231, 1999.

[19] A. D. Smith and D. R. Taylor, "Is exhaled nitric oxide measurement a useful clinical test in asthma?," *Curr. Opin. Allergy Clin. Immunol.*, vol. 5, no. 1, pp. 49–56, 2005.

[20] D. R. Taylor, M. W. Pijnenburg, A. D. Smith, and J. C. de Jongste, "Exhaled nitric oxide measurements: clinical application and interpretation," *Thorax*, vol. 61, no. 9, pp. 817–827, 2006.

[21] C. Grimes *et al.*, "A sentinel sensor network for hydrogen sensing," *Sensors*, vol. 3, no. 3, pp. 69–82, 2003.

[22] M. Z. Jacobson, W. G. Colella, and D. M. Golden, "Cleaning the air and improving health with hydrogen fuel-cell vehicles," *Science (80-.).*, vol. 308, no. 5730, pp. 1901–1905, 2005.

[23] A. M. Bassam, A. B. Phillips, S. R. Turnock, and P. A. Wilson, "Development of a multi-scheme energy management strategy for a hybrid fuel cell driven passenger ship," *Int. J. Hydrogen Energy*, vol. 42, no. 1, pp. 623–635, 2017.

[24] L. Boon-Brett *et al.*, "Identifying performance gaps in hydrogen safety sensor technology for automotive and stationary applications," *Int. J. Hydrogen Energy*, vol. 35, no. 1, pp. 373–384, 2010.

[25] W. J. Buttner, M. B. Post, R. Burgess, and C. Rivkin, "An overview of hydrogen safety sensors and requirements," *Int. J. Hydrogen Energy*, vol. 36, no. 3, pp. 2462–2470, 2011.

[26] K. H. Kim, E. C. Jeon, Y. J. Choi, and Y. S. Koo, "The emission characteristics and the related malodor intensities of gaseous reduced sulfur compounds (RSC) in a large industrial complex," *Atmos. Environ.*, vol. 40, no. 24, pp. 4478–4490, 2006.

[27] H. Kimura, "Hydrogen sulfide: Its production and functions," *Exp. Physiol.*, vol. 96, no. 9, pp. 833–835, 2011.

[28] K. H. Kim, Y. Choi, E. Jeon, and Y. Sunwoo, "Characterization of malodorous sulfur compounds in landfill gas," *Atmos. Environ.*, vol. 39, no. 6, pp. 1103–1112, 2005.

[29] M. N. Hughes, M. N. Centelles, and K. P. Moore, "Making and working with hydrogen sulfide. The chemistry and generation of hydrogen sulfide in vitro and its measurement in vivo: A review," *Free Radic. Biol. Med.*, vol. 47, no. 10, pp. 1346–1353, 2009.

[30] S. C. K. Misra, P. Mathur, M. Yadav, M. K. Tiwari, S. C. Garg, and P. Tripathi, "Preparation and characterization of vacuum deposited semiconducting nanocrystalline polymeric thin film sensors for detection of HCl," *Polymer (Guildf).*, vol. 45, no. 25, pp. 8623–8628, 2004.

[31] C. Yao *et al.*, "Sub-ppm CO detection in a sub-meter-long hollow-core negative curvature fiber using absorption spectroscopy at 2.3 μm," *Sensors Actuators, B Chem.*, vol. 303, no. October 2019, p. 127238, 2020.

[32] S. H. Nimkar, S. P. Agrawal, and S. B. Kondawar, "Fabrication of Electrospun Nanofibers of Titanium Dioxide Intercalated Polyaniline Nanocomposites for CO2 Gas Sensor," *Procedia Mater. Sci.*, vol. 10, no. Cnt 2014, pp. 572–579, 2015.

[33] S. Pandey, "Highly sensitive and selective chemiresistor gas/vapor sensors based on polyaniline nanocomposite: A comprehensive review," *J. Sci. Adv. Mater. Devices*, vol. 1, no. 4, pp. 431–453, 2016.

[34] M. Imran, N. Motta, and M. Shafiei, "Electrospun one-dimensional nanostructures : a new horizon for gas sensing materials," no. 2, 2018.

[35] A. Altindal, "ORGANlK YARI iLETKEN FiLMLERiN KARAKTERiZASYONU VE GAZ SENSORU OLARAK UYGULANMASI," 1999.

[36] Z. Li *et al.*, "Resistive-type hydrogen gas sensor based on TiO2: A review," *Int. J. Hydrogen Energy*, vol. 43, no. 45, pp. 21114–21132, 2018.

[37] "Wafer LiNbO3 Properties," 2020. [Online]. Available: https://unitedcrystals.com/LiNbO3Prop.html. [Accessed: 10-Jan-2020].

[38] K. J. lesker Company, "Gold Properties." Kurt J. Lesker Company.

[39] K. J. L. Company, "Titanium Dioxide (TiO2) Sputtering Targets." Kurt J.Lesker Company, pp. 2–5, 2020.

[40] K. J. L. Company, "Zinc Oxide (ZnO) Sputtering Targets." Kurt J. Lesker Company, pp. 2–5, 2020.

[41] Comsol Multiphysics, "Heat Transfer, Comsol Multiphysics." 2020.

[42] T. K. Roy, D. Sanyal, D. Bhowmick, and A. Chakrabarti, "Temperature dependent resistivity study on zinc oxide and the role of defects," *Mater. Sci. Semicond. Process.*, vol. 16, no. 2, pp. 332–336, 2013.

[43] Y. Zhang and J. Han, "Microstructure and temperature coefficient of resistivity for ZnO ceramics doped with Al2O3," *Mater. Lett.*, vol. 60, no. 20, pp. 2522–2525, 2006.

MANAS Journal of Engineering
ISSN 1694-7398 | e-ISSN 1694-7398

Volume 9, Issue 1, (2021) Pages 45-54
https://doi.org/10.51354/mjen.869736

One Step Ahead Prediction of Ozone Concentration for Determination of Outdoor Air Quality Level

Waleed Khalid M. Mahmood[1,*], Ercan Avşar[2]

[1]Çukurova University, Department of Electrical and Electronic Engineering, Adana, Turkey, waleed.mahmood93720@gmail.com, ORCID:0000-0002-4973-0106
[2]Dokuz Eylül University, Department of Computer Engineering, İzmir, Turkey, ercan.avsar@deu.edu.tr, ORCID: 0000-0002-1356-2753

ABSTRACT

With the rapid spread of urbanization, competent authorities become increasingly anxious from air pollution risks and effect on citizens especially those with respiratory diseases. In this work, performances of six machine learning methods were analyzed for prediction of maximum ozone (O_3) concentration for the next-day. The models make the prediction using concentrations of six atmospheric components (PM2.5, PM10, Ozone (O3), Sulfur Dioxide (SO2), Nitrogen Dioxide (NO2), and Carbon Monoxide (CO)). The utilized machine learning methods are multilayer perception (MLP), Support Vector Regression (SVM), k-Nearest Neighbor (K-NN), Random Forests (RF), Gradient Boosting (GB), and Elastic Net (EN). After the predictions made by these models, the predicted values were further processed to be classified into one of the six air quality levels defined by United States Environmental Protection Agency. The prediction performances of the models as well as their corresponding classification results were analyzed. It was shown that MLP model gives the lowest RMSE of 2246 for prediction step while SVR achieved the highest accuracy score of 0.790.

ARTICLE INFO

Research article

Received: 28.01.2021
Accepted: 16.04.2021

Keywords:
machine learning, time-series forecasting, regression methods, sequence-to-sequence, air quality index.

**Corresponding author*

1 Introduction

The increase in industrialization and urbanization has effects on the natural balance, this consequently affect human, animal, and plant wellbeing. Air pollution is one of the results of this situation and is considered to be the fourth of the biggest killer in the world after tobacco, high blood pressure and poor diet [1]. The surveys confirm the possibility of an increased number of deaths 40% from outdoor air pollution than today, therefore the United Nations Sustainable Development to aim to put strategy for the next two decades to limit air pollution hazards [2]. The Organization for Economic Cooperation and Development (OECD) has expected in 2012 that Outdoor Air Pollution will endure increasing towards 2050, and eventually will be the main cause of environmentally-related deaths worldwide [3].
Ambient Air Pollution is one of the mainsprings that worsens respiratory disease like Asthma and Chronic Obstructive Pulmonary Disease (COPD); therefore, air pollution has classified as a threat factor for people with these disease. According to the Global Asthma Network (GAN) studies, there are approximately 334 million asthma patients worldwide and accounts for 383 thousand deaths annually [4].
The World Health Organization (WHO) in 2016 has published an estimate survey that ambient air pollution accounts for 4.2 million premature death worldwide and more than 18% of deaths were courses COPD [5, 6].
According to the surveys, outdoor air pollution considers the major of many substantial adverse to universal public health, therefore, the level of hazard requires the cooperation of all researchers to study, analyze and reduce the risk of outdoor air pollution on the health of chronic respiratory patients [7]. In addition to civilized awareness, people are becoming more attentive due to changes in air quality and their impact on respiratory patients health. Although the children and young adults spend most of the time indoors because of COVID-19, the air quality-related with outdoor environments is still important for Asthma and COPD patients because outdoor environment air has a major impact on the indoor air quality. The reason for such an impact is that there is a lack of efficacy in preventing particulate matter transmission from the outside environment to the inside environment [8].
The significance of air quality forecasting is increasing rapidly for establishing citizen's superior quality and safety life to adapt with growing and developing of the modern era, and for

providing an influential and appropriate environment for sensing research. Therefore, this work aims to explain the potential ways to set up a time series forecast model architecture, which to facilitate the establishment of an airtight future model aim to support respiratory disease patients.

Majority of the previous studies rely on six important pollutants of the atmospheric components (PM2.5, PM10, Ozone (O3), Sulfur Dioxide (SO2), Nitrogen Dioxide (NO2), and Carbon Monoxide (CO)) to health problems. O3 is formed by complex chemical processes having a high impact on the environment. As a result of O3 exposure, approximately more than 21,000 premature deaths were reported in Europe, and 1.1 million in worldwide premature deaths, many of death have respiratory diseases [9]. Therefore, O3 is selected as the target pollutant to be predicted in this study. For this purpose, six regression models that are artificial neural networks (ANN), support vector machines (SVM), k-nearest neighbor (KNN), random forests (RF), gradient boosted decision tree (GBDT), and elastic net (EN), utilized for establishing a time series-forecasting model. The prediction models were trained by using several measurements of the pollutant gases and the daily maximum O3 level for the next day was forecast using these models. For the classification method, the predicted values obtained by training data were classified to five levels of air quality (good, moderate, unhealthy for sensitive groups, unhealthy, very unhealthy). United States Environmental Protection Agency (US EPA) standard equation was used calculate the air quality level and hence to classify the predicted value. These classification results were used to generate a confusion matrix, eventually.

2 Literature Review

In recent times, air pollution threats have been studied widely by the researchers, several of them have contributed to investigate interesting solutions with the potential of reducing the risks. Most of them consider machine-learning methods as the ideal solution to develop a time series forecast model. In this section, we will discuss some of the machine learning methods of previous work.

In earlier years, many researchers have tended to compare regression models. For example, N. K. Ahmed et al. presented a paper to compare the major regression models for building the architecture of time-series forecast [10]. The comparisons of models are substantial movement to highlight the appropriate model to establish a time-series forecast, also, provide solutions to time-series forecasting problems. In 2010, Ping-Feng Pai et al. illustrated the significance of using SVM with time series forecasting and compared the performance with Autoregressive Integrated Moving Average (ARIMA) [11]. Involvement of SVM in building the architecture of sequence-to-sequence forecast provide a major effect on the generalization performance, also doesn't take much time to implement. The substantial common solution in the growth era is establishing a hybrid model to eliminate the main problems in time-series forecast. In 2013, F. S. de Albuquerque Filho et al. suggested an intelligent hybrid system for time-series forecasting to predict the levels of pollution of four components (CO, SO2, PM10, and NO2) in the atmosphere. The hybrid model built by an ANN-MLP with a particle swarm optimization (PSO) algorithm [12]. ANN-MLP model considers an appropriate model for time-series forecast, however, it is a good idea to reduce errors in a model by using a provide potential in PSO algorithm and combined with MLP. Concerning regression tree methods, James R. Lloyd clarified a susceptibility to perform time-series methods by using a GD model to forecast hourly loads of a US utility [13]. Then, H. Tyralis and G. Papacharalampous focused on clearly the potential of using an RF model on time-seres forecasting [14]. Regression tree methods have proven competent to time-series forecasting, however, in future are expected to provide better prediction accuracy.

Due to rapid of industrialization and economic boom in the urban cities of worldwide, Sarajevo capital city of Bosnia has suffered from a high level of pollution during the winter session. In 2016, Maja M. Dedovic et al. presented a research paper urges to reduce the accumulated risks to Sarajevo by recourse to sequence-to-sequence forecast method. They suggested a model predict the concentration of pm10 particles from 2010 to 2013 together with the environmental measures like temperature, humidity, wind speed and pressure. The proposed model depended on an ANN model for training a dataset. In model evaluation, the proposed model output a forecasting result with R^2 from 0.5 to 0.9 of total years [15]. Bing-Chun Liu et al, proposed an AQI forecast model to predict air pollution in three urban cities in China. The main goal of the proposed model is an enhancement of air quality forecasting results via minimizing the errors of machine learning prediction algorithms. The machine learning regression method used in the proposed model is SVM for predict air quality index. They illustrated suffers, problems and struggles with air pollution for each city. The proposed regression model achieved high-performance efficiency, besides, MSE results achieved by training model is between 106.22 to 128.70 of total cities [16]. In 2018, Hong Zheng et al. presented a paper highlighting on rising of air pollution and reduce its impact in China. proposed a new approach of multiple kernel learning (MKL) with support vector classifier (SVC) to train air quality dataset. MKL-SVC performance compared with various machine learning models had widely used, like SVM, RF, ANN-MLP, ARIMA, and long short-term memory (LSTM). The proposed model shown high performance beyond the other models, also they praised SVM, RF, and ANN-MLP performance these presented very good performance than the sequence-to-sequence models LSTM and ARIMA. MKL-SVC model succeeded accuracy of 0.972 with MSE of 0.030 [17]. F. Martinez et al. illustrated the potential to train a time-series dataset by KNN model. It is possible to use KNN model in time series architecture, however, it is not the most suitable among the previously implemented models [18]. In 2020, K.Maheshwari and S. Lamba, they proposed time series forecast model. The

proposed model aimed to minimize the effects of air pollution emission. the proposed model forecast future concentrations of pm2.5 depend on six machine learning regressors models for training a dataset, were Linear Regression, ANN-MLP, K-NN, Decision Tree, RF and Stochastic Gradient Descent. The proposed model achieved an accuracy of 95.5% [19]. F. Shen et al. proposed a prediction model to overcome the limitation in time-series forecasting. The combination of EN model and high order Fuzzy cognitive maps (FCMs) shown predictable of a time-series dataset with less error compared with other regression models [20]. EN model is a combination of lasso regression and ridge regression, it has proven efficiency of predict time series compare with lasso and Ridge.

3 Materials and Methods

3.1 Time Series Dataset

Dataset was obtained from The World Air Quality Index project; they have published it with the spreading out of COVID-19 [21]. Dataset Includes 380 major cities in the worldwide, has been collected from the several stations since 2016 until now it is updated daily. The World Air Quality Dataset contains the minimum, the maximum, the median and the standard deviation for daily measurements of the six air pollutant species ($PM_{2.5}$, PM_{10}, SO_2, NO_2, CO and O_3). Among these species, the unit for O_3 and CO is parts per million (ppm), while SO_2 and NO_2 are measured in parts per billion (ppb). On the other hand, the unit for the particulate matter pollutants, PM_{10} and $PM_{2.5}$ is micrograms per m^3 ($\mu g/m^3$). The atmospheric components like temperature, humidity, wind speed, wind gust, dew, and pressure are not included in the dataset, hence not utilized in this study. A sample part of the dataset is provided in Table 1. In this work, Istanbul was chosen as the city whose air quality to be predicted for the next day. After pre-processing steps, the dataset size has obtained consists of 1013 row and 56 columns. The data between the dates from 26.12.2016 to 14.10.2020 is used as the dataset. The reason for selecting this time interval is to be able to include the data belonging to covid-19 situation. In that case, the future expansions to the current study may cover detailed analysis about the temporal effects of covid-19 on the outdoor air pollution.

Table 1. An illustrative table of dataset contents.

Date: From 12/26/2016 To 10/15/2020	**Country: Turkey**	**City**	**Species**	**Air Pollutant Species (S)**			
				Min	**Max**	**Median**	**Variance**
12/26/2016 12:00:00 AM	TR	IST	SO_2 (ppb)	0.6	20.8	3.6	150.56
12/26/2016 12:00:00 AM	TR	IST	PM_{10} (ppm)	1	119	32	2965.21
12/26/2016 12:00:00 AM	TR	IST	O_3 ($\mu g/m^3$)	0.5	18.3	2.5	136.67
12/26/2016 12:00:00 AM	TR	IST	NO_2 (ppb)	6.9	76.1	28.8	2093
12/26/2016 12:00:00 AM	TR	IST	CO ($\mu g/m^3$)	0.1	59.3	7.6	1997.7
12/26/2016 12:00:00 AM	TR	IST	$PM_{2.5}$ (ppm)	2	751	85	61407.9

3.2 Air Quality Index

After performing prediction by the models, the predicted values are assigned an AQI level and then compared with the AQI level of the actual value. Air Quality Index (AQI) is a required index that represents air quality based on air pollutant concentrations at a certain period. AQI schedule varies depending on the country, each country containing a specific Index. AQI Accredited by United States Environmental Protection Agency (US EPA) standard has divided into 6 groups. Each group symbolized by different colors and a standardized public health advisory depending on the concentration of air pollution [22]. The US EPA standard method for calculating AQI is given in Equation (1) and the levels for AQI indexes are illustrated in Table 2. The green color indicates the best and changes gradually to maroon indicating the worst. The orange color has been reported to be the threshold for risky situations for Sensitive Groups (lung disease, older adults and children).

Equation (1): The US EPA standard equation to calculate air quality index.

$$I_p = \frac{I_{HI} - I_{LO}}{BH_{HI} - BH_{LO}} \left(C_p - BH_{LO} \right) + I_{LO} \quad (1)$$

Where I_p = the index for pollutant p

C_p = the truncated concentration of pollutant p

BH_{HI} = the concentration breakpoint that is greater than or equal to C_p

BH_{LO} = the concentration breakpoint that is less than or equal to C_p

I_{HI} = the AQI value corresponding to BH_{HI}

I_{LO} = the AQI value corresponding to BH_{LO}

Table 2. AQI levels.

AQI	Air Pollution Levels	Health Implications
0-50	Good	Air quality is satisfactory, no risk.
50-100	Moderate	Air quality is acceptable.
100-150	Unhealthy for Sensitivity Group	Members of sensitive groups may experience health effects.
150-200	Unhealthy	Members of the public may experience health effects, for sensitive groups will be more seriouse health effects.
200-300	Very Unhealthy	The risk of health effects is increased for everyone; health alert.
300-500	Hazardous	Health warning for emergency conditions.

3.3 Time-Series Dataset Preprocessing

Initial step, Istanbul city dataset was extracted from the public main dataset, and then the six essential components for atmospheric pollution are separated from other atmospheric components. Eventually, a time-series data containing minimum, maximum, median, and variance of daily measurements of six pollutants for 37 months of duration is obtained. This dataset was prepared for the time-series prediction task by performing the steps explained in the following subsection.

3.3.1 One-Step Ahead Forecast

The purpose of time-series prediction is estimating the future information by past and current information samples. Therefore, one-step forecast strategy (in this problem it may be defined as "next day predict strategy") is a supervised learning task. The main goal of this strategy is generating forecasting dataset from current data by reframing the original dataset to derive historical dataset (t-1, t-2,.….t-n) and next-day (t+1) dataset, then merge all of them to gain entry time-series dataset with target dataset, therefore, predicting the future values (t+1, t+3,…. t+ n) (Figure 1).

Figure 1. *One-step Ahead strategy.*

3.3.2 One-hot Encoding

The main step in preprocessing dataset is a coding method, for getting the better results in machine learning (ML) based forecasting. One-Hot Encoding was used to encode time-series dataset. One-Hot Encoding employs N-bits status, each independent category of entry has one bit. Bits represented by "one" or "zero" and only one of the entries will be valid with a state of "one". As shown in Table 3, One-Hot Encoding was used to encode weekdays, a 7-bits encoding had used to differentiate each day. For example, for a sample that corresponds to measurements for Wednesday, the one-hot feature vector is 0001000.

Table 3. *One-Hot Encoding scheme for weekdays.*

days	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
12/25/2016	**1**	**0**	**0**	**0**	**0**	**0**	**0**
12/26/2016	**0**	**1**	**0**	**0**	**0**	**0**	**0**
12/27/2016	**0**	**0**	**1**	**0**	**0**	**0**	**0**
12/28/2016	**0**	**0**	**0**	**1**	**0**	**0**	**0**
12/29/2016	**0**	**0**	**0**	**0**	**1**	**0**	**0**
12/30/2016	**0**	**0**	**0**	**0**	**0**	**1**	**0**
12/31/2016	**0**	**0**	**0**	**0**	**0**	**0**	**1**

3.3.3 Standardization

In general, the machine learning models require to rescale input dataset to be sure that the input features are measured on the same scale, especially when all data on a completely different scale. Standardization or called Z-score normalization refers to transform input dataset to zero mean and unit standard deviation to be on the form of a normal distribution. The formula for applying z-score normalization is given in Equation (2) where X denotes the input feature, μ and $\boldsymbol{\sigma}$ are mean and standard deviation of the feature, respectively. This transformation is applied to all features other than one-hot encoded binary features.

$$\boldsymbol{z} = \frac{\mathbf{x}-\mu}{\sigma} \tag{2}$$

3.3.4 Walk-Forward Validation

The common methods in machine learning such train-test split and k-fold cross-validation are avoided here because it is not a valid approach for time series datasets. The dataset was divided into training and testing subsets based on train test split, which esteems the temporal order of observations. To eliminate common errors in time series forecasting, the prediction of test set should not rely solely on the training set. Common errors were avoided by performing Walk-Forward validation methods. Walk-Forward validation approach is predicting one value from the test set for each iteration at a time. After predicting the value taken from the test values will add this value to training set or widow set (number of training data each iteration) for the next iteration. Training and predicting of data that will continue until the end of all iteration steps, consequently, the number of iteration should be equal to the number of tests. There are two types of Walk-Forward validation, Rolling Walk Forward and Anchored Walk Forward. In this study, Rolling Walk Forward (also named as rolling window) method is used. Rolling window approach is the in-sample periods “roll”; meaning that the window size is steady during all iteration steps [23]. In other words, the number of samples in the training set will not change during the prediction process, as shown in Figure 4.

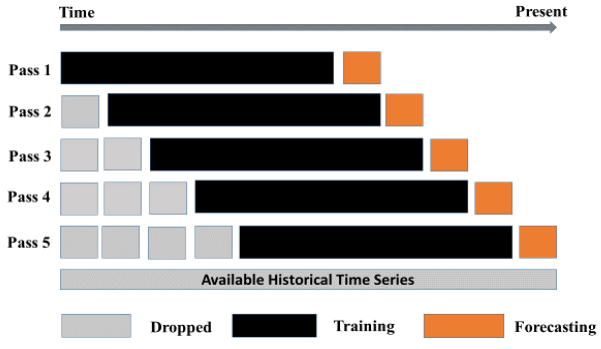


Figure 2. *Illustration Rolling Walk Forward validation.*

3.4 Regression Models

Six regression models were used to perform the time series forecasting. The forecasting models used a past and current time step to next-day forecast, in turn, to compare with the future data we have already specified as preliminarily target sets. The comparison for measurement of our proposed model's performance with the MSE and R^2 for regression methods, and accuracy, F-score, Recall, and Precision for classification methods. In this section, we will clarify applied regression models for the sequence-to-sequence forecast**.**

3.4.1 Support Vector Regression

Support Vector Machines (SVMs) are a satisfactory way to perform on regression problems based on Vapnik-Chervonenkis (VC) theory [24], it evolved over the last decade to generalize to become appealable on yet-to-be-seen data. As a supervised machine learning methods, it has been proven to be most effective in a real time-series forecast [25]. The prediction functions $f(x)$ for linear and non-linear regression applications had derived from one-dimensional example shown in equations (3, 4). If we assume that $f(x)$ is non-linear, thus, our aim is mapping data to higher dimensional "feature" space or called kernel space, which achieved by Kernel function $\phi(x)$ [26]. Kernel functions types are Gaussian function and polynomial, and hyperbolic tangent. For our proposed model chosen Gaussian function or radial basis function (RBF), as shown in Equation (5), where$\|\boldsymbol{x}-\mu\|$ is the squared Euclidean distance of two-feature vector and $\boldsymbol{\gamma}$ is gamma. We aim to find the best weight (w) and threshold (b). Parameter C is the regularization parameter setting the margin of the decision function in linear and non-linear SVM. If we encourage a large margin, should select a small value of parameter C, consequently will get a lower misclassification rate and vice versa.

$$\boldsymbol{f(x)} = (\boldsymbol{w}.\boldsymbol{x}) + \boldsymbol{b} \tag{3}$$

$$\boldsymbol{f(x)} = (\boldsymbol{w}.\boldsymbol{\phi(x)}) + \boldsymbol{b} \tag{4}$$

$$\boldsymbol{\phi(x,\mu)} = \boldsymbol{exp}(-\boldsymbol{\gamma}\|\boldsymbol{x}-\mu\|^{\mathbf{2}}) \tag{5}$$

3.4.2 K-Nearest Neighbors

K-Nearest Neighbors (KNN) regression is one of lazy learning or instance-based learning (memory-based learning) that use for time-series forecast. KNN is a simple non-parametric algorithm that based on determining the most k- nearest or similar of new samples (test instance samples with an unidentified target) to the training instance (feature and its target samples) according to the similarity measure or a distance metric [27]. The k parameter represents the number of neighbors to consider for determining the predicted value

of a test sample. For KNN regression, the average of closest k samples is calculated as the prediction. Equation (6) shows the mathematical equation that used to calculate the distance metric.

$$d_{ij} = \sqrt{\sum_{m=1}^{n}(x_{im} - x_{jm})^2} \quad (6)$$

Where d_{ij} represent the distance metric between specified dataset, m is the specific time, x_{im} is the training instances, i, j is the dimensionality of the feature space and x_{jm} is the testing data with an unidentified target.

3.4.3 Multilayer Perceptron

Fundamentally, Artificial Neural network (ANN) is a group of nodes or units connected with each other like neurons in a human brain, each node is able to transfer data to the other starting from the input layer going through the hidden layer, and reaching the output layer. Multilayer Perceptron (MLP) is an ANN belong to feedforward neural network class, which means the input data move to only forward directions [28]. MLP has a wide agreement that it is the appropriate model for time-series forecasting regression because it provides a non-linear tool for resolving regression problems, and is able to approximate any smooth function without prior presumption about data distribution. MLP model was tuned by adjusting number of hidden layer units, which is the first and most effective parameter that permits us to set a number of layers and nodes possessed by the neural network.

3.4.4 Regression Trees

Regression Trees (RT) are based on the principle of splitting the input parameters space into special independent and non-overlapping regions according to set rules. Many alternatives work on similar prediction principle with better predictive performance and a completely different way to create trees are random forest (RF), and gradient boosting (GB). RF algorithm approach is forming trees independently, with maximum depth. RF structure permits to minimize the variance of a huge number of complex decision trees (high depth) [29]. GB is a numerical optimization algorithm that works on approach adding a new decision tree at each step iteratively to minimize the loss function [30]. The RF and GB regression models was tuned by adjusting the number of trees, which is the major effective parameter that determine the number of sequential trees for best forecasting.

3.4.5 Elastic Net

Elastic net is a linear regression model that considers a combination of lasso regression (least absolute shrinkage and selection operator) and Ridge regression. Lasso is a penalized minimal squares method proposed by Tibshirani (1996), which imposes an L1-penalty on the regression coefficients [31]. The aim of EN method is reducing the loss function that given below:

$$L_{en}(\hat{\boldsymbol{\beta}}) = \frac{\sum_{i=1}^{n}\left(y_i - x_i^J\hat{\beta}\right)^2}{2n} + \lambda\left(\frac{1-\alpha}{2}\sum_{j-1}^{m}\hat{\boldsymbol{\beta}}_j^2 + \alpha\sum_{j-1}^{m}\left|\hat{\boldsymbol{\beta}}_j\right|\right) \quad (7)$$

Where λ is the model parameter, **x** is the input matrix, **y** is the target vector and $\boldsymbol{\beta}$ is the weight vector. Alpha (α) is the mixing parameter that combines Ridge and Lasso methods. When $\alpha = 0$ the elastic net will corresponds to simple ridge regression, while $\alpha = 1$ to lasso regression.

3.5 *Performance metrics*

In order to predict the air quality level for the next day, fist the ozone concentration is first predicted then this predicted value is assigned a level using Equation 1. Hence, different performance measures are calculated for each of these steps. Since the first step is a regression problem, root mean squared error (RMSE) and R^2 values are calculated according to the equations given below:

$$\boldsymbol{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(Y_{Ti} - Y_{pi})^2}{n}} \quad (8)$$

Where n is number of non-missing data point, Y_{Ti} and Y_{pi} denote the observed and predicted values, respectively.

$$\boldsymbol{R^2} = 1 - \frac{RSS}{TSS} \quad (9)$$

Where R^2 is the coefficient of determination.

RSS is the residual sum of squares = $\sum_i(Y_{Ti} - Y_{pi})^2$ (10)

TSS is the total sum of square = $\sum_i(Y_{Ti} - Y_{mi})^2$ (11)

And Y_{mi} is denote the number of the Y_{Ti}.

The assigned level of air quality is then compared by its actual level. In order to do it for all the predictions, a multi class confusion matrix is generated because five different levels of air quality are defined. A generic confusion matrix considering k classes is given in Figure 3.

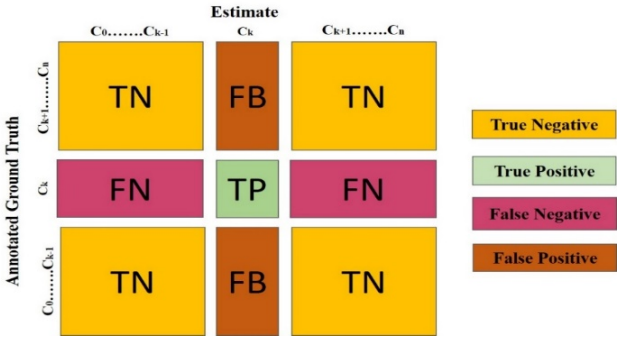


Figure 3. *The confusion matrix for multi-class.*

For the final classification part, accuracy, precision, recall, and f1-score are calculated according to the equations given below:

$$\boldsymbol{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN} \quad (12)$$

$$\boldsymbol{Precision} = \frac{TP}{TP+FP} \quad (13)$$

$$\boldsymbol{Recall} = \frac{TP}{TP+FN} \quad (14)$$

$$\boldsymbol{F1\ Score} = \frac{2 * (Recall * Precision)}{(Recall + Precision)} \quad (15)$$

4 Results and Discussion

In the previous section, the dataset size and concerned region were mentioned, and the methods were implemented to process the dataset as well as the architecture of the proposed algorithm for the time-series forecasting. In this section, the results obtained with the mentioned methods are presented. Since these methods have different parameters that influence the prediction performance, a gird of parameters for the methods are used in testing of the methods. These parameters are given in Table 4. For the number of neighbors (k) in KNN model, three values are used and the lowest RMSE was obtained when K=3. In the SVM model stage, the linear method is shown with regularization parameter equal one achieved the best RMSE result compared with the non-linear method. As for MLP, several of hidden layer standard and the maximum number of iteration was examined. In an MLP with two hidden layers, where the first layer has 14 units and the second has 7 units, yielded the smallest RMSE when the maximum training iterations was set to 300. In the EN model, three alpha values were examined and 0.9 was found to be the most suitable one compared with others. As for the number of trees in regression trees models, three values are examined for both RF and GB, 100 and 300 are determined to be the best number of trees for RF and GD, respectively. The obtained RMSE and R^2 for the regression methods with the best parameters are given in Table 5. As indicated earlier, the predicted values are categorized into different levels of air quality and then compared with the actual level. The results associated with this classification step is given in Table 6.

Table 4. *The examined methods and the corresponding parameters.*

Method	Parameters	Tried values for the parameters
KNN	K	3,5,7
SVM	Kernel type	Linear, Gaussian (RBF)
	C	1,3,7,550,700
	gamma	0.001, 0.0008, 0.0003
RF	Number of estimators	100, 300, 500
EN	Alpha	0.1, 0.5, 0.9
MLP	Number of maximum iterations	200, 300 , 400, 500 , 700, 1000
	Hidden layer size	(10,5), (14,7), (20,10), (50,25), (25,), (50,), (70,), (100,), (150,), (200,), (250,), (300,)
GB	Number of estimators	300, 500, 700

Table 5. *Regression results for air quality prediction.*

	KNN	SVM	RF	EN	MLP	GB
RMSE	2594	2284	2548	2250	**2246**	2745
$\boldsymbol{R^2}$	0.45	**0.52**	0.46	**0.52**	**0.52**	0.42

Table 6. *Classification results for air quality prediction.*

	KNN	SVR	RF	EN	MLP	GB
Accuracy	0.756	**0.790**	0.713	0.726	0.697	0.709
recall	0.756	**0.790**	0.713	0.726	0.697	0.709
F1-score	0.757	**0.788**	0.736	0.743	0.716	0.732
Precision	0.76	**0.786**	0.772	**0.786**	0.765	0.760

The performance measures for regression models in Table 5 are close to each other. However, it is clearly seen that MLP performs the prediction with lowest error and three methods (SVM, EN and MLP) has the highest R^2 score. In Figure 4, the actual and predicted values on the test set are given on the same plot for visual comparison. In addition, these values are presented as a scatter plot in Figure 5. Both of these figures depict the results obtained by MLP model, which outputs the lowest RMSE value. It is obvious from these figures that the majority of errors are caused by the samples where the actual O_3 concentration is very high. This is probably due to the fact that the samples with high O_3 value are very rare. As a result, it becomes difficult for the learning methods to model the samples with such behavior. On the other hand, the classification results in Table 6 show that the accuracy and recall values for all of the methods are identical. This implies balanced classification models, meaning that the ability of the models for classifying the positive samples and the negative

samples are the same. This may be due to the indirect class assignment of the test samples. In other words, since the corresponding O_3 values are predicted for test samples firstly and next, the corresponding class labels are calculated according to Equation 1, the classification results may have such a balance. Besides, it notable that predictions obtained by SVR method yield the highest classification result for all of the metrics. The corresponding confusion matrix for SVR predictions is given in Figure 6.

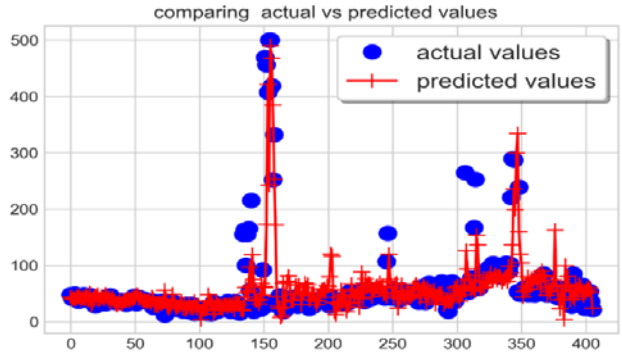


Figure 4. *Actual and predicted values for MLP model*

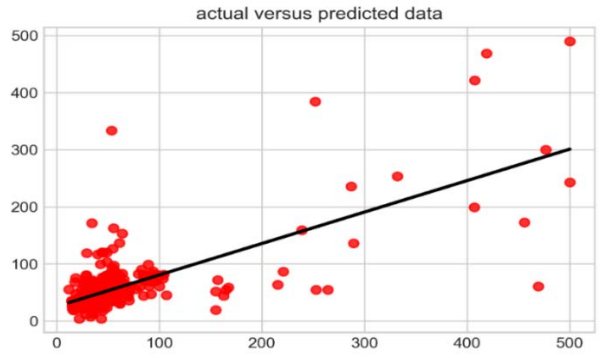


Figure 5. *Actual versus predicted values for MLP models in a scatter plot*

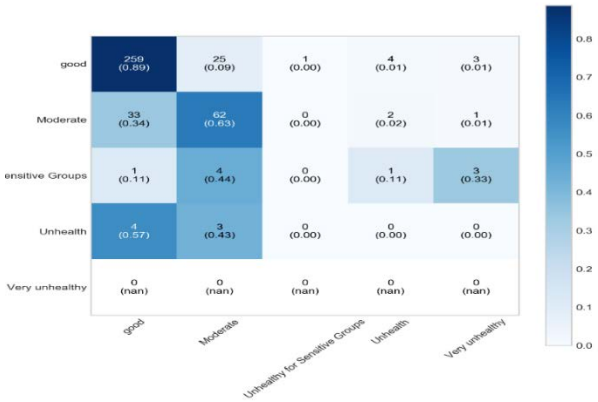


Figure 6. *The confusion matrix for SVR predictions. Rows and columns indicate actual and predicted values, respectively.*

In the related literature, there are other studies in which the future values of air pollutant concentrations are predicted. Majority of these studies try to solve this problem by means of regression methods only. In other words, no further calculations for air quality level assignment are made in these works. Furthermore, only a limited number of methods with predefined parameters are used in these works. For example, P. García Nieto et al. presented a prediction model for forecasting the time-series dataset of PM based on SVM, VARMA, ARIMA and ANN-MLP [32]. The lowest RMSE measure was achieved by SVM as 2.061. In the study proposed by Z. Meng, a prediction model to forecast ground O_3 level based on machine learning methods [33]. SVM, decision tree, RF, and logistic regression models were utilized to train the O_3 dataset which contains binary target values as high or low ozone concentration. In other words, no regression model was fit on the data and the obtained classification accuracies ranged between 0.8 and 0.949. Another study about classification of air quality levels considers five different stages for the AQI [34]. The study uses Naïve Bayes and decision trees for the classification task where 91.99% of accuracy was observed with decision trees. Even though the accuracy score is higher than those obtained in this work, the F1-score performance was not reported. Therefore, it becomes impossible to analyze the performance in case of a class imbalance problem in the dataset. In a recent study, ANN method is utilized for predicting four pollutant species that are NO_2, PM_{10}, $PM_{2.5}$, and O_3. The results were enhanced through an operation called real-time-corrections and it was shown that prediction performance of O_3 can have RMSE and R^2 scores of 0.187 and 0.79, respectively [35]. However, in that work, no information regarding the unit of the O_3 measurement is provided. Therefore, the RMSE results is not comparable with those obtained in this study. In addition, performances of different machine learning methods have not been compared and suitable parameters for the ANN model has not been searched in that work. Even though there are some studies in which performances of various methods have been analyzed and compared together [36, 37], a two-step method (class assignment followed by regression) for air quality level detection has not been proposed earlier.

5 Conclusion

In this paper, several machine learning methods were analyzed to predict maximum O_3 concentration for the next-day as an indicator of outdoor air quality. The predictions were made for one-step ahead in the test set and the training was performed using rolling walk forward validation method. Therefore, a distinct training process was performed for each test sample. Total of six different machine learning methods, namely, KNN, SVM, RF, EN, MLP, and GB were used for time-series prediction experiments. Next, the predicted values were assigned a level for the air quality. This assignment was made using the formula proposed by United States Environmental Protection Agency. As a result, the problem was transformed into a multi-class classification problem. For

the first prediction step, MLP was observed to achieve the lowest error while SVR predictions were better classified into the air quality levels. Therefore, it may be concluded that low prediction accuracy does not always imply better representation for air quality levels. Another noticeable point in the predictions is that error related with low pollution values is lower but the models make more error for the samples corresponding to high pollution values. This means that increasing the number of samples for high pollution measurements in the dataset may enable the models learn the patterns for such samples.

References

[1] Niall McCarthy, "Air Pollution Contributed to More Than 6 Million Deaths in 2016, Data journalist covering technological", societal and media topics, 2016.

[2] P.Rafaj, G.Kiesewetter , T.Gul, W.Schoppa, J.Cofala, Z.Klimont, P.Purohit, C.Heyes, M.Amann, J.Borken-Kleefeld, L.Cozzi. Outlook for clean air in the context of sustainable development goals. : Global Environmental Change, September 2018.

[3] OECD (2012), OECD Environmental Outlook to 2050, OECD PUBLISHING. https://dx.doi.org/10.1787/9789264122246-en.

[4] Aloke Ghoshal, Pradyut Waghray, George Dsouza, Mahip Saluja, Mayank Agarwal, Ashish Goyal, Sneha Limaye, Akash Balki, Sudhir Bhatnagar, Manish Jain, Sharad Tikkiwal, Abhijit Vaidya, Meena Lopez, Rashmi Hegde, Jaideep Gogtay, "Real-world evaluation of the clinical safety and efficacy of fluticasone/formoterol FDC via the Revolizer in patients with persistent asthma in India", On 25 November 2019, 10.1016/j.pupt.2019.101869.

[5] Burden of disease from ambient air pollution for 2016, 1211 Geneva 27, World Health Organization 2018. https://www.who.int.

[6] World Health Organization, Ambient air pollution: a global assessment of exposure and burden of disease, 2016, https://apps.who.int/iris/handle/10665/250141.

[7] X. Li, Ling Jin, and H. Kan, Air pollution: a global problem needs local fixes, 25 JUNE 2019, china, https://doi.org/10.1038/d41586-019-01960-7.

[8] Blondeau, P., Iordache, V., Poupard, O., Genin, D., Allard, F., 2005. Relationship between outdoor and indoor air quality in eight French schools. Indoor Air 15, 2–12, 10.1111/j.1600-0668.2004.00263.

[9] Brian S. Freeman, G. Taylor, B. Gharabaghi, and Jesse Thé, forecasting air quality time series using deep learning, 24 May 2018. https://doi.org/10.1080/10962247.2018.1459956.

[10] Nesreen K. Ahmed, Amir F. Atiya , N.El Gayar &H. El-Shishiny, An Empirical Comparison of Machine Learning Models for Time Series Forecasting, 15 Sep 2010. https://doi.org/10.1080/07474938.2010.481556.

[11] Ping-Feng Pai, Kuo-Ping Lin, Chi-Shen Lin, and Ping-Teng Chang, Time series forecasting by a seasonal support vector regression model, June 2010, https://doi.org/10.1016/j.eswa.2009.11.076.

[12] Francisco S. de Albuquerque Filho, Francisco Madeiro e Sérgio M. M. Fernandes, Paulo S. G., de Mattos Neto, and Tiago A. E. Ferreira, Time-series forecasting of pollutant concentration levels using particle swarm optimization and artificial neural networks, Paulo 2013. http://dx.doi.org/10.1590/S0100-40422013000600007.

[13] James R. Lloyd, GEFCom2012 hierarchical load forecasting: Gradient boosting machines and Gaussian processes, 16 August 2013. https://doi.org/10.1016/j.ijforecast.2013.07.002.

[14] H.Tyralis, and G.Papacharalampous, Variable Selection in Time Series Forecasting Using Random Forests, 4 October 2017. https://doi.org/10.3390/a10040114.

[15] M. M. Dedovic, S. Avdakovic, I. Turkovic, N. Dautbasic, and T. Konjic, Forecasting PM10 concentrations using neural networks and system for improving air quality, 08 December 2016, 10.1109/BIHTEL.2016.7775721.

[16] Bing-Chun Liu, A. Binaykia, P. Chang, M.K. Tiwari, C.-C. Tsao, urban air quality forecasting based on multi-dimensional collaborative Support Vector Regression (SVR): A case study of Beijing-Tianjin-Shijiazhuang. July 14, 2017, https://doi.org/10.1371/journal.pone.0179763.

[17] H.Zheng ,H. Li, X. Lu, and T. Ruan, A Multiple Kernel Learning Approach for Air Quality Prediction, 12 Jun 2018, https://doi.org/10.1155/2018/3506394.

[18] F. Martínez, M. P. Frías, F. Charte and A. J. Rivera, Time Series Forecasting with KNN in R: the tsfknn Package, December 2019, ISSN 2073-4859.

[19] K.Maheshwari and S. Lamba, Air Quality Prediction using Supervised Regression Model, 03 February 2020, 10.1109/ICICT46931.2019.8977694.

[20] Fang Shen, Jing Liu, and Kai Wu, Multivariate Time Series Forecasting based on Elastic Net and High-Order Fuzzy Cognitive Maps: A Case Study on Human Action Prediction through EEG Signals, 29 May 2020. 10.1109/TFUZZ.2020.299851.

[21] World's Air Pollution: Real-time Air Quality Index, http://waqi.info/.

[22] Eusebio Jarauta-Bragulat, Carme Hervada-Sala, Juan Jose Egozcue, Air Quality Index Revisited from a Compositional Point of View, published online: 23 May 2015 © International Association for Mathematical Geosciences 2015 .

[23] A. Sanjivanrao More, D.Sunil Ranaware, B. D. Wamane, and G. S. Salunkhe, Enhancement in Financial Time Series Prediction with Feature Extraction in Text Mining Techniques, Nov 2019, 2395-0056, International Research Journal of Engineering and Technology (IRJET).

[24] V. N. Vapnik, An overview of statistical learning theory, Sept. 1999.10.1109/72.788640.

[25] N. I. Pankevych, R.Sankar, Time Series Prediction Using Support Vector Machines: A Survey, 24 April 2009. 10.1109/MCI.2009.932254.

[26] M. Awad, R.Khanna, Efficient Learning Machines (chapter: Support Vector Regression, Pages 67-80), 27 April 2015. https://doi.org/10.1007/978-1-4302-5990-9

[27] F.Martínez, M. P. Frías, M. D. Pérez, and A. J. Rivera, a methodology for applying k-nearest neighbor to time series forecasting, 21 NOV 2019. https://doi.org/10.1007/s10462-017-9593-z.

[28] J. Kukkonen, L. Partanen, A. Karppinen, J. Ruuskanen, H. Junninen, M. Kolehmainen, H. Niska, S. Dorling, T. Chatterton, R. Foxall, and G.Cawle, extensive evaluation of neural network models for the prediction of NO2 and PM10 concentrations, compared with a deterministic modelling system and measurements in central Helsinki, 22 Aug 2003. https://doi.org/10.1016/S1352-2310(03)00583-1.

[29] S. Touzani, J. Granderson, and S. Fernandes, Gradient boosting machine for modelling the energy consumption of commercial buildings, Nov 2017. https://doi.org/10.1016/j.enbuild.2017.11.039.

[30] Max Kuhn, Kjell Johnson, Applied Predictive Modeling, New York 2013, https://doi.org/10.1007/978-1-4614-6849-3.

[31] Hui Zou, and Trevor Hasti, Regularization and variable selection via the elastic net, 09 March 2005, https://doi.org/10.1111/j.1467-9868.2005.00503.x.

[32] P. García Nieto, F. Sánchez Lasheras, E. García-Gonzalo, F. de Cos Juez, PM10 concentration forecasting in the metropolitan area of Oviedo (Northern Spain) using models based on SVM, MLP, VARMA and ARIMA: A case study, 2018, *Science of the Total Environment*, https://doi.org/10.1016/j.scitotenv.2017.11.291.

[33] Z. Meng, Ground Ozone Level Prediction Using Machine Learning, 2019, *Journal of Software Engineering and Applications*, 10.4236/jsea.2019.1210026.

[34] R. Waman Gore, D. S. Deshpande, An approach for classification of health risks based on air quality levels, 01 December 2017, India, 2017 1st International Conference on Intelligent Systems and Information Management (ICISIM), 10.1109/ICISIM.2017.8122148.

[35] S. Agarwal, S. R. Sharma, Md H. Suresh Rahman, et al., Air quality forecasting using artificial neural networks with real time dynamic error correction in highly polluted regions, 2020, Science of the Total Environment, https://doi.org/10.1016/j.scitotenv.2020.139454.

[36] Doreswamy, K. Harish Kumar, Y. Km, I. Gad, Forecasting Air Pollution Particulate Matter (PM2.5) Using Machine Learning Regression Models, 2020, *Procedia Computer Science*, https://doi.org/10.1016/j.procs.2020.04.221.

[37] S. Ameer, M. Ali Shah, A. Khan; H. Song, C. Maple, S. Ul Islam, M. N. Asghar, Comparative Analysis of Machine Learning Techniques for Predicting Air Quality in Smart Cities, 2019, 10.1109/ACCESS.2019.2925082.

[38] W. Leong, R. Kelani, Z. Ahmad, Prediction of air pollution index (API) using support vector machine (SVM), 2020, Journal of Environmental Chemical Engineering, https://doi.org/10.1016/j.procs.2020.04.221.

[39] Z. Yu, and Z. Niu; W. Tang, Deep Learning for Daily Peak Load Forecasting—A Novel Gated Recurrent Neural Network Combining Dynamic Time Warping, 29 January 2019, 10.1109/ACCESS.2019.2895604.

[40] C. S. Malley, D. K. Henze, Johan C.I. Kuylenstierna, H. W. Vallack, Y. Davila, S. C. Anenberg, M. C. Turner, and M. R. Ashmore, Updated Global Estimates of Respiratory Mortality in Adults ≥30 Years of Age Attributable to Long-Term Ozone Expos, 28 August 2017, https://doi.org/10.1289/EHP1390.

[41] Nan-Hung Hsieh, Chung-Min Liao, Fluctuations in air pollution give risk warning signals of asthma hospitalization, August 2013, https://doi.org/10.1016/j.atmosenv.2013.04.043.

[42] S. Du, T. Li, and Shi-Jinn Horng,Time Series Forecasting Using Sequence-to-Sequence Deep Learning Framework, 02 May 2019, 10.1109/PAAP.2018.00037.

[43] T.Liu, A. K. H. Lau, K. Sandbrink, J. C. H. Fung,Time Series Forecasting of Air Quality Based On Regional Numerical Modeling in Hong Kong, 24 March 2018, https://doi.org/10.1002/2017JD028052.

[44] NICOLÒ BALDON, Time series Forecast of Call volume in Call Centre using Statistical and Machine Learning Methods, Sweden 2019, urn:nbn:se:kth:diva-265002.

[45] S. G. Gocheva-Ilieva, A. V. Ivanov, D. S.Voynikova, and D. T. Boyadzhiev, Time series analysis and forecasting for air pollution in small urban area: an SARIMA and factor analysis approach, 25 September 2013, https://doi.org/10.1007/s00477-013-0800-4.

[46] G. Papacharalampous, H. Tyralis, and D. Koutsoyiannis, Univariate Time Series Forecasting of Temperature and Precipitation with a Focus on Machine Learning algorithms: a Multiple-Case Study from Greece, 29 November 2018, https://doi.org/10.1007/s11269-018-2155-6.

MANAS Journal of Engineering
ISSN 1694-7398 | e-ISSN 1694-7398

Volume 9, Issue 1, (2021) Pages 55-65
https://doi.org/10.51354/mjen.872956

Technologies based on energy savings for OLED devices

Mohammed Albaba[1,*], Meryem Sena Akkus[2]

[1,*] Electrical and Electronics Engineering Department, Ankara Yildirim Beyazit University, Ankara/Turkey, 195105403@ybu.edu.tr, ORCID: 0000-0002-0833-8191
[2] Central Research Laboratory Application and Research Center, Ankara Yildirim Beyazit University, Ankara/Turkey, msakkus@ybu.edu.tr, ORCID: 0000-0003-2550-550X

ABSTRACT

Recently, scientists are paying more attention to the Organic Light Emitting Diode (OLED) technology as it is being used in devices and displays to play videos and show photos with high resolution. This technology is used in products such as mobile phones, televisions, laptops, etc. To make the energy consumed less, new methods were shown up to prevent high energy consumption while presenting videos and photos on OLED devices and displays without losing their details and quality, one of the methods is a deep learning-based technique which is related to artificial intelligence. In this review paper, the last methods were discussed as well as their results. Saturation, brightness, contrast, and luminance are factors that impacting energy consumption. In terms of OLED mobile phones, there were a few studies that concentrated on turning off the unnecessary pixels which will be black as default, and as a result, the lifetime of batteries will be extended. Also, for OLED mobile phones, a web browser called Chameleon was presented as it has some modes to save the energy consumed while surfing the internet by remapping the displayed colors of the website.

ARTICLE INFO

Research article

Received: 2.02.2021
Accepted: 25.06.2021

Keywords:
OLED Displays,
Power Consumption,
Deep Learning Method,
Chameleon Web Browser,
Reducing Power

*Corresponding author

1 Introduction

Organic Light-Emitting Diode (OLED) display is a new technology that showed better results than liquid crystal displays (LCD) in terms of brightness, good view, and the ability to make flexible displays from it [1]. In the last few years, a lot of display industries started to merge OLED displays in their products such as smartphones, TVs, wearable devices, and much more. Also, transparent OLED display (TOLED) was presented lately, as many companies started to merge it in their goods like military products, vehicle construction, and other industries [2]. To decrease the power consumption in OLED displays, some methods were made in terms of hardware enhancement and image processing [3,4,5]. This review paper talks about the next methods, Unsupervised PCCE Technique, Region Of Interest (ROI) Extraction & Region Of Noninterest (NON-ROI) Dimming, A Web Browser For Mobile OLED Displays; Chameleon, and Darkening OLED Display Partially, that were used to achieve the desired goal. In terms of mobile devices, it was found out that screens consuming the highest ratio of the battery's power by comparing it with other components with a percentage between 38-50% [6,7]. The advantage of OLED displays is that each pixel emits light by itself without needing a backlight, unlike LCDs. Therefore, it easier to save power consumed by turning off the unnecessary pixel according to the displayed image or video as every pixel emits the green, red and blue color [8, 9]. According to Dong and Zhong, the power consumption is low in black color and highest in white color [10, 11]. Some researchers advised that changing color sets may affect the energy consumed by OLED and make it less by 40% [12].

In the experience of web browsing using smartphones with OLED displays, it is known that webpages have different color designs, especially white color which is 80% [13], some of them use GIFs and flash videos which increase the power demand from the battery to display them on OLED displays. Mobile devices that have OLED displays are facing some problems in terms of energy saving and researchers say that losing energy could be less when the reduction of brightness is applied, but the user will not be satisfied [14]. Some methods were implemented to reduce power consumption while keeping and preserving the details of images displayed on OLED display. One of the methods is depending on the power-constrained contrast enhancement (PCCE) implemented by using one of deep learning algorithms [15]. Another method is by darkening the image by reducing the luminance of some pixels according to the displayed image on OLED display, in other words, remapping the colors of the image [8]. One more method depended on the dark scheme

which generates a part of black color on OLED screen with a specific percentage [16].

2 Methods and Implementations

In this section, some methods will be explained in terms of saving power while preserving the displayed details of the image according to some studies. Also, other studies will be explained on mobile devices that use the OLED display.

2.1 Unsupervised PCCE Technique

This method relates to one of deep learning algorithms directly that called Convolutional Neural Network (CNN). Deep learning outperforms the other methods in terms of image enhancements [17,18]. It is used to enhance power consumption by reducing it while viewing the most improved photo in terms of brightness and details on an OLED display [15]. This display's total dissipated power (TDP) that has N pixels is expressed according to [4,19,20] as follows:

$$P = \sum_{i=1}^{N}\left(w_o + w_R R_i^{\gamma} + w_G G_i^{\gamma} + w_B B_i^{\gamma}\right) \quad (1)$$

Whereas R_i, G_i and B_i are the intensities of red, green and blue color for the *i* pixel, wo constant value of power giving to for the non-pixel section of display, W_R, W_G, W_B and γ are the values of coefficients that are depending on the panel.

This method requires a dataset to be implemented, the dataset was chosen to be Berkeley Segmentation Dataset (BSD) [21]. To train a deep learning CNN model, two components must be considered which are the network architecture design and the loss function. Context aggregation network (CAN) [22] was employed as the base network because it has the feature of demanding a small number of network parameters. For the loss function, training without the reference image was made as it is related to unsupervised learning. By doing this, the enhancement of image's contrast is applied, and at the same time, power consumption is being saved [15]. The method divides the image into two images, luminance, and color. The luminance's intensity reduces to decrease the power demand, of course, the contrast in this way improved by using the deep learning method, CNN, and after that, the resulted photo will be added to the color image as shown in Figure 1 [15].

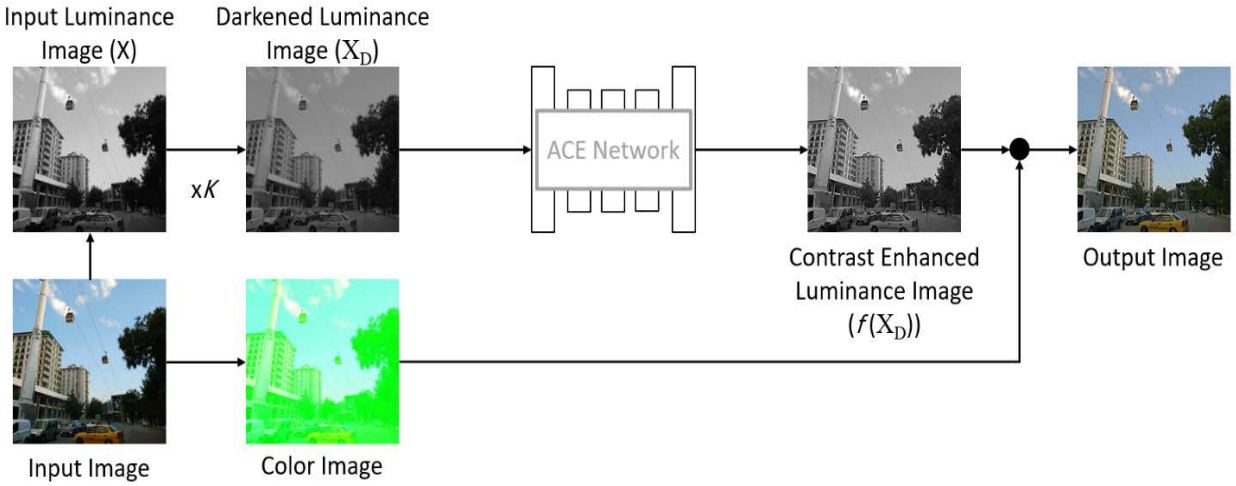


Figure 1. *The image before and after implementing CNN method [Adapted from reference 15].*

After reducing the luminance of the input image, the rate of power saving (R) will be reduced as well as shown in Eq. (2). P_{dim} and P_{in} are TDP of dimmed and input images, and k is a ratio of reducing the luminance image's intensity. By implementing this, the image quality may reduce and lose its details, for that the adaptive contrast enhancement (ACE) was introduced to improve the contrast of darkening image [15].

$$R = 1 - \frac{P_{dim}}{P_{in}} = 1 - \frac{\sum_i^N (kY_i)^{\gamma}}{\sum_i^N (Y_i)^{\gamma}} = 1 - k^{\gamma} \quad (2)$$

Whereas R_i, G_i, and B_i are the intensities of red, green, and blue color for the *i* pixel, Wo constant value of power giving to for the non-pixel section of the display, W_R, W_G, W_B, and γ are the values of coefficients that are depending on the panel.

As shown in Figure 2, the ACE network architecture contains some layers from the input image which is L_0 to the output image that is L_d. Through the process, the image is being processed by applying the CNN method on it to reach the best image details with less power demand [15].

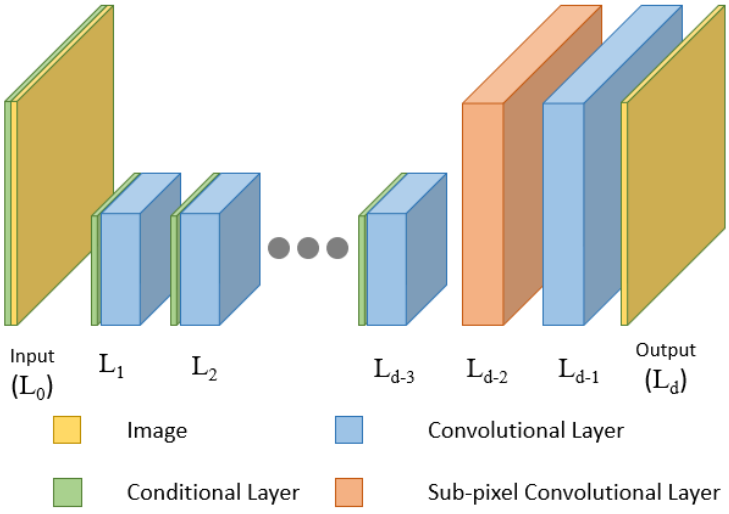


Figure 2. *The ACE network architecture [Adapted from reference 15].*

In terms of the loss function, there are three factors that make the total loss function which is power loss L_p, similarity loss L_s and contrast loss L_c. The first factor makes sure that the power is reduced after improving contrast, the second factor sets the changing of the image, and the third factor is enhancing the output image's contrast. The sum of the three losses is giving total loss as it is illustrated in Eq. 3. [15]. Hyper parameters that control-related importance to every factor is λ_P, λ_S, and λ_C.

$$L_{total} = \lambda_P L_P + \lambda_s L_s + \lambda_c L_c \quad (3)$$

For training the ACE network, 300k iterations were trained with a 0.0001 learning rate [23, 24]. Adam optimizer was used as well [25]. The ACE network has seven layers, and the rate of power-saving (R) was between [0.01, 0.8] for every image in the training section for every iteration, while in the test section, the range was between [0.1 - 0.7] with a step of 0.1.

The process is done by using a single Titan X GPU with the usage of Tensorflow. The dataset is the BSDS500 dataset [21] which has 500 images, it is randomly separated into the training section and the testing section as 400 images and 100 images, respectively. As a result, the training process is good and it is done without the problem of over-fitting.

In terms of video displaying, an experiment was applied for two movies which are 'The Shawshank Redemption' and 'Avatar' [26]. An average luminance with the value of R in the ACE network was employed for the frame of the input video. In dark scenes, the R-value is approximately 0 to stop reducing brightness, whereas it has a large value for bright scenes and frames to reduce their brightness in order to reduce power consumption [15].

2.2 Region Of Interest (ROI) Extraction & Region Of Noninterest (NON-ROI) Dimming

Here, by using extraction and dimming features and applying them to the image, the power consumption reduces. In addition, the details of the image save as well.

Dong & Zhong were the first who modeled the formulas for optimizing and modeling power for the OLED displays, they designed Eq. 4 which shows the power contributed for one pixel in terms of three colors are red, green, and blue, whereas f(R), h(G), and k(B) are expressing the consumed power for each color in the pixel; red, green and blue, respectively.

Eq. 5 shows the full power consumption that the OLED display consumes for n pixels. There is a constant value which is C that is expressing the stable power consumption required by control chips [10].

$$P_{pixel}(R,G,B) = f(R) + h(G) + k(B) \quad (4)$$

$$P = C + \sum_{i=1}^{n} \{f(R_i) + h(G_i) + k(B_i)\} \quad (5)$$

It depends on getting rid of the un-useful details while keeping the area of interest in the photo. It has two levels; the first level is extracting the region of interest (ROI) by using the detection model on the photo. The second level is by changing the saturation and luminance degrees that are related to the noninterest region (NON-ROI) of the photo. By implementing this technique, the consumed power reduces as the NON-ROI is getting dimmed by changing its parameters which are saturation and luminance.

To use the abstraction method on ROI, the Canny [27] algorithm was used, also, and for changing luminance and saturation values for NON-ROI, ALS (adjusting luminance & saturation) method was used.

In terms of ROI, the Canny method is designed for a gray image to find the edges of it and because of the improvements of this industry, it uses to detect the edges of color photos [27]. So, to apply this method, by using Eq. 6, the photo converts to gray [8].

$$P_{Gray} = P_R * 0.299 + P_G * 0.587 + P_B * 0.114 \quad (6)$$

After that, the resulted gray photo will be smoothed. Then, the areas of images with high and sharp gray intensity will be found, and a gradient will be applied to them, this step is called non-maximum suppression. Then after, and to get rid of the noise in the photo, it is a substantial thing to filter the undesired low gradient value. Finally, the pixels will be connected to each other forming the edges.
In terms of NON-ROI, and by changing the values of saturation and luminance, the power consumption will be reduced. Keeping the quality and details of the photo are important in this case. These factors will be applied to NON-ROI related to the photo to make sure that the photo still has almost the same requirements for the user vision.

For changing luminance value, it is known that the pixel's luminance is related to the red, green, and blue colors. When changing pixel's luminance, their values reduce power consumption. The luminance for a pixel is the sum of values of R, G, and B colors for a pixel, Eq. 7 shows the function of sum. Adjustment of luminance is expressed by Eq. 8. P_{pixel} is prime luminance for one pixel, whereas P'_{pixel} is expressing luminance after adjusting it. The parameter of adjusting is X which has a value between {0-1}. Adjustment regions' number in the noninterest region in the image is expressed by N. Region number of pixels located is n_i.

$$P_{pixel}(S) = f(R, G, B) \quad (7)$$

$$P'_{pixel}(S) = f(R, G, B) * (1 - X * (\frac{n_i}{N})) \quad (8)$$

To change the value of saturation, and according to OLED displays, the blue color component was found that it is consuming higher power than other color components as shown in Figure 3. So, when adjusting the saturation value of the photo, it means that the values of the components that consuming higher power will be reduced and the bright image will be dimmed as well. The method's name is Adjusting Luminance & Saturation (ALS) [8]. For videos and mobile games, the darkening method is a good technique to be implemented in order to save more energy [28, 29].

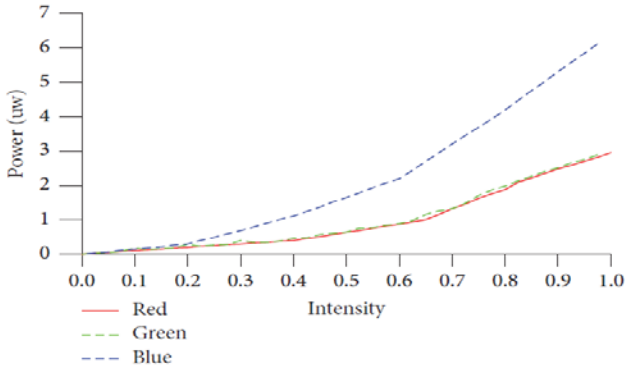


Figure 3. *Power consumed by red, green, and blue components of OLED pixel [Reproduced with permission from reference 8. Copyright 2016, Deguang Li et al. Mobile information systems]*.

2.3 A Web Browser For Mobile OLED Displays; Chameleon

In terms of internet usage and surfing various websites on mobile OLED displays, a web browser color-adaptive is presented. OLED displays are consuming a lot of energy when the main color is white, which is a power-hungry color. It is obvious that the black color consumes much less power.
For one pixel *I*, the total power consumption is expressed by Eq. 9, meanwhile, Eq. 10 is to calculate the power consumption for an OLED display of *N* pixels [30].

$$P_i = a.R_i + b.G_i + c.B_i \quad (9)$$

$$P = \sum_{i=1}^{N} P_i = \sum_{i=1}^{N} (a.R_i + b.G_i + c.B_i) \quad (10)$$

As it is noticeable, most of websites are depending on white color in their color design for displaying their websites' contents. From that point, the Chameleon web browser raised up to recolor the GUI objects of websites, while keeping images without changes. The GUI objects' colors are being transformed by counting, mapping, and painting them again in a way that keeps the consistency of their colors for each website as is illustrated in Figure 4 [31].

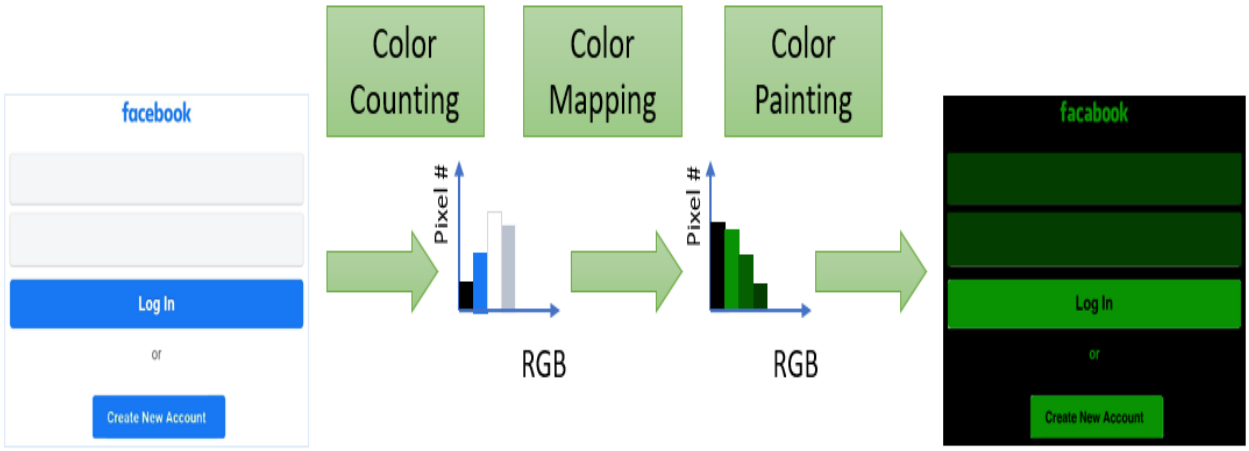


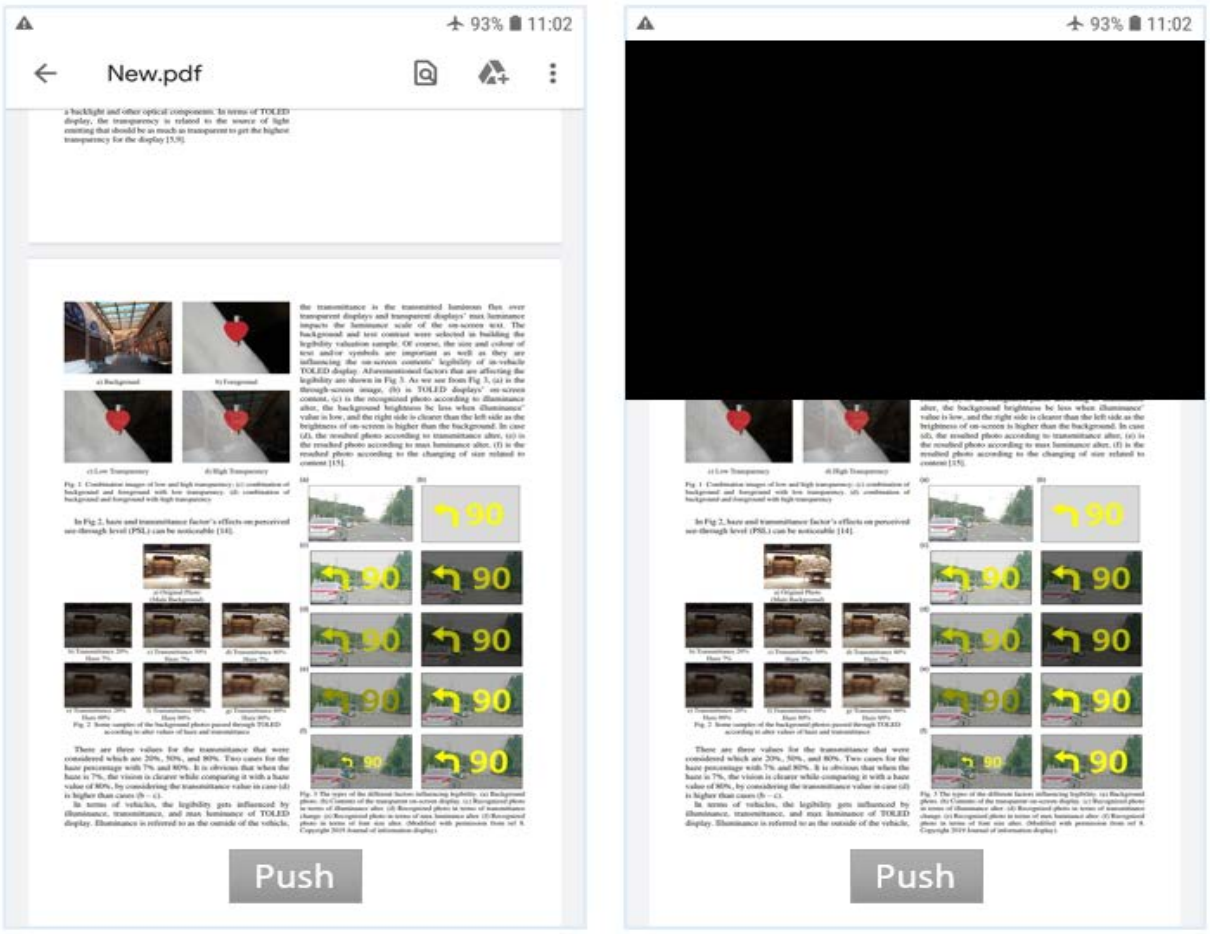
Figure 4. *GUI objects' color transforming. [Adapted from reference 31]*

In terms of video support, it is easy to make the video darker by putting a dark layer that has an opacity level on the video with the same size as it. As a result, power consumption will decrease and battery life will increase [31].

2.4 Darkening OLED Display Partially

As OLED displays depending on the colors that are being displayed, a new method was discovered to reduce the power consumption of OLED displays by darkening them partially. In this method, the black color which is known as a nonconsumer power color, active to darken the display by a specific ratio according to the OLED display. In other words, the pixels which include the diodes turn off.

This method made some experiments by analysing and modifying the android OS internal code. It is well known that when using mobile phones, a drawing process is being activated to show the new contents of the display which consumes more power, especially when scrolling the screen up and down [32]. By implementing this method, and by editing the source code of the Android operating system, it is possible to change the displaying parameters and the displaying dimensions according to the whole size of the OLED display, as it is related to the GPU drawing [33], by dividing the screen into the active and darkened screen as it is illustrated in Figure 5. In other words, it gets rid of the unnecessary computations which are happening in the android OS to view the new contents after scrolling up and down [16]. The study was made also by taking brightness into consideration, low and high brightness intensity, for both cases which are full display and partial display on OLED screens [16].



(a) (b)

Figure 5. *(a) Viewing full screen on whole OLED display (b) Viewing partial screen with 30% darkening display from top [Adapted from reference 16].*

Some studies [34, 35, 36] concentrated on Dynamic Voltage Scaling (DVS) to reduce the power consumed by OLED display when showing a video as the model is dominate the OLED display's input voltage.

3 Results and Discussions

The results of the unsupervised power-constrained contrast enhancement (PCCE) technique, region of interest (ROI) extraction & region of noninterest (NON-ROI) Dimming method, and Chameleon method studies show in this section.

3.1 Unsupervised PCCE Technique

The conventional methods prevent selecting the requested (R) value, so they selected the parameters to get that value inversely, so the produced image gets the desired value of (R=0.5). When this value selects, power consumption reduces to its half as the intensity of the image's luminance reduces. As shown in Figure 6, the optical comparison is made between four methods, three of them are conventional and the fourth one is the proposed method. In Figure 6, (a) is the input image, (b) is HPCCE [4], (c) is Chang et al [37], (d) is Chondro et al [5], and (e) is the proposed method [15]. By comparing the produced images, we can easily notice that the proposed method has the best result by preserving the details of the image and improving the contrast for the fixed value of (R). The input image in Figure 6 (a) was taken from BSDS500 test sets of the dataset [38].

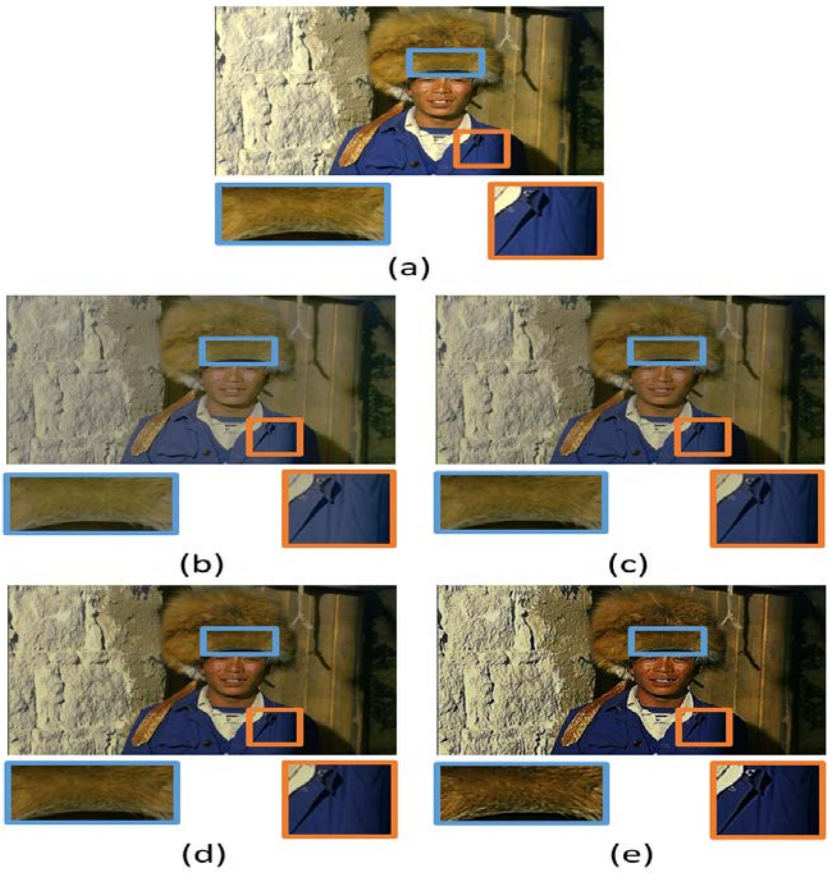


Figure 6. *The optical comparison of the conventional and proposed method in BSDS500 test sets; (a) original photo, (b) HPCCE [4], (c) Chang et al [37], (d) Chondro et al [5], (e) proposed method [15] [Adapted from reference 15].*

As mentioned before, the loss function is a combination of three factors, which are power loss L_p, similarity loss L_s, and contrast loss L_c. To check the performance of each factor, training was made on the ACE network without each factor. The proposed method is the best result which is the combination of the three factors is working together to get the superior image as it shows the high contrast, image details and at the same time limits the power consumption.

3.2 Region Of Interest (ROI) Extraction & Region Of Noninterest (NON-ROI) Dimming

By using the proposed method of adjusting luminance & saturation (ALS) and comparing it with the other two methods which are adjusting luminance (AL) and adjusting saturation (AS) in terms of power consumption, the best method selected for future usage. The previous methods applied to (Seagull) photos [8]. The result shows in Figure 7. The display model was selected to be the aμOLED-32028-P1 AMOLED (P1 Active Matrix Organic Light Emitting Diode) with a power meter of HOIKI 3334. The resulted photo after implementing the three methods illustrates the differences between each of them in a visual way [8].

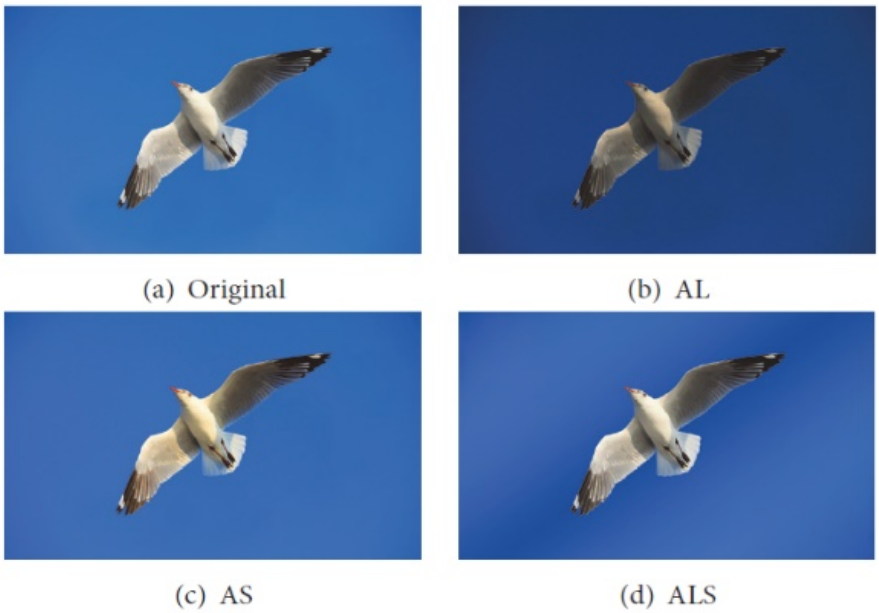


Figure 7. *Images processed by the three methods: (b) Using adjusting luminance, (c) Using adjusting saturation, (d) Using adjusting luminance & saturation [Reproduced with permission from reference 8. Copyright 2016, Deguang Li et al. Mobile information systems]*.

When using adjust luminance method, the power reduced by a percentage of 26.68% in (b) case. In (c) case, the power decreased by 25.96%. In the proposed method which is (d) case, power consumption of 26% reduced.

It was found that the adjusting luminance & saturation method is effective in terms of saving power when the content of photos has blue or white colors with a high ratio. However, it is not much effective, when the major color of photos is grey.

3.3 A Web Browser For Mobile OLED Displays; Chameleon

Like every application, there are various options and modes to use and select from inside the program. Chameleon web browser offers four modes to select from which are; Dark, Green, Arbitrary, and Inversion mode as shown in Figure 8.

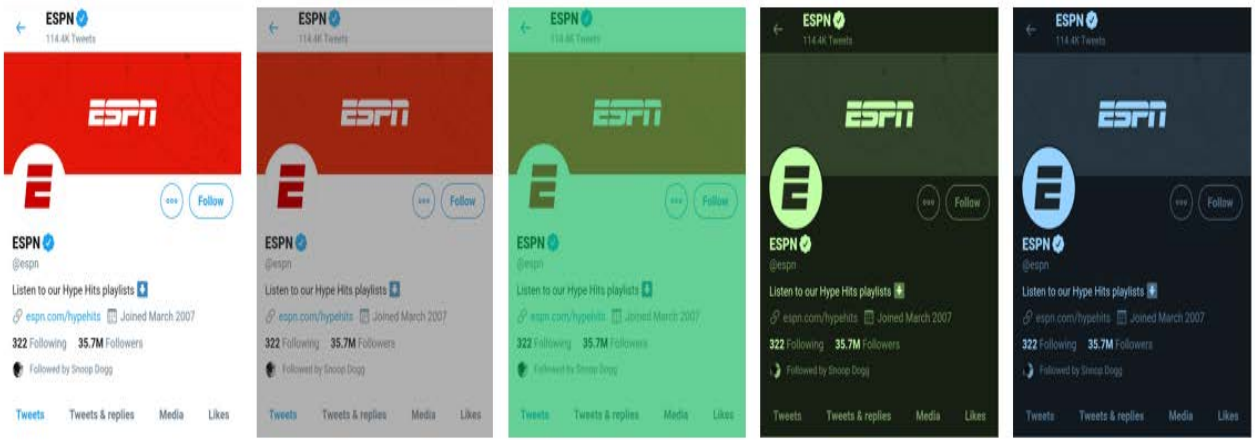


Figure 8. *Original view and other four modes for displaying [Adapted from reference 31]*.

It was found that the power consumption decreased by 25% in dark mode, 34% in green mode, 66% in Inversion mode, and 72% in Arbitrary mode. Inversion mode uses blue color. In this result, the Arbitrary mode is the best in terms of saving power. Overall, and in this method, while surfing the internet and websites, 41% of the power consumption of the system was reduced.

3.4 Darkening OLED Display Partially

The presented method promises to achieve more by reducing the power consumption of OLED displays. By applying it on Nexus 6, and after shutting down some services like Bluetooth, Wi-Fi, GPS and putting it in airplane mode, the results came out. The method was able to reduce the power by 11.1% when low brightness was activated, and in high brightness mode the power reduction was 10.4%.

4 Conclusion

Some studies were made to enhance saving power consumption for OLED displays. Energy-saving methods for OLED displays commonly used and information about the application of the methods is given in this review paper. This paper shows four various techniques which showed positive and promising results. The first work is by using an unsupervised deep learning method on PCCE with great results for the OLED display. This method is the convolutional neural network (CNN) which is used in the ACE network, and for training it. A new loss function is proposed which is a combination of three factors, power loss, similarity loss, and contrast loss. The result was superior in terms of contrast, details, and power consumption by comparing it with other methods. The second study depended on extracting edges and changing the values of luminance and saturation. The quality of photos was preserved and power consumption was reduced by 22.5%. The third study was made by using a color adaptive web browser on mobile OLED devices, it showed a perfect performance in terms of power reduction which was reduced by 41% for the system power. The fourth and last study explained the idea of changing the computation of the drawing area of mobile OLED display devices, by altering the android operating system's code and getting rid of the unnecessary computation on the dimmed area. The reduction of power consumption reached 11.1% when the brightness was low, and 10.4% when high brightness was activated.

As a future work in terms of methods used, more advanced deep learning methodologies may be used to get more accurate results. In addition, some improvements could be done to the internet browser to make it more reliable by decreasing consumed power. Also, Android and iPhone systems' developers may possibly add some features that one of them is letting users darken a specific part of their phones' display by developing their internal system.

References

[1] S. Kang, "Image-Quality-Based Power Control Technique for Organic Light Emitting Diode Displays," in Journal of Display Technology, vol. 11, no. 1, pp. 104-109, Jan. 2015, doi: https://10.1109/JDT.2014.2363086

[2] ALBABA, Mohammed; AKKUS, Meryem Sena. NEXT GENERATION DISPLAY TECHNOLOGY: TRANSPARENT ORGANIC LIGHT EMITTING DIODES. Selcuk-Teknik Dergisi, [S.l.], p. 168-180, dec. 2020. ISSN 1302-6178. Available at: http://sutod.selcuk.edu.tr/sutod/article/view/524. Date accessed: 29 Dec. 2020.

[3] C. Chae, H. Le, K. Lee, G. Cho and G. Cho, "A Single-Inductor Step-Up DC-DC Switching Converter With Bipolar Outputs for Active Matrix OLED Mobile Display Panels," in IEEE Journal of Solid-State Circuits, vol. 44, no. 2, pp. 509-524, Feb. 2009, doi: https://10.1109/JSSC.2008.2010986

[4] C. Lee, C. Lee, Y. Lee and C. Kim, "Power-Constrained Contrast Enhancement for Emissive Displays Based on Histogram Equalization," in IEEE Transactions on Image Processing, vol. 21, no. 1, pp. 80-93, Jan. 2012, doi: https://10.1109/TIP.2011.2159387

[5] P. Chondro, C. Chang, S. Ruan and C. Shen, "Advanced Multimedia Power-Saving Method Using a Dynamic Pixel Dimmer on AMOLED Displays," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 28, no. 9, pp. 2200-2209, Sept. 2018, doi: https://10.1109/TCSVT.2017.2723600

[6] A. Carroll and G. Heiser, "An analysis of power consumption in a smartphone," in Proceedings of the USENIX Conference on USENIX Annual Technical Conference, vol. 14, p. 21, Boston, Mass, USA, June 2010.

[7] X. Chen, Y.Chen, Z.Ma, andF. C. A. Fernandes, "How is energy consumed in smartphone display applications?" in Proceedings of the 14th Workshop on Mobile Computing Systems and Applications (ACM HotMobile '13), Jekyll Island, Georgia, USA, 26-27 February 2013.

[8] Deguang Li, Bing Guo, Yan Shen, Junke Li, Yanhui Huang, "Making Image More Energy Efficient for OLED Smart Devices", Mobile Information Systems, vol. 2016, Article ID 6575931, 8 pages, 2016. https://doi.org/10.1155/2016/6575931

[9] S. Kim, S. Hyun, T. Heo, D. Im and J. Huh, "Blind: Power saving color transform method for OLED displays," 2016 IEEE International Conference on Consumer Electronics (ICCE), 2016, pp. 500-501, doi: https://10.1109/ICCE.2016.7430704

[10] M. Dong and L. Zhong, "Power Modeling and Optimization for OLED Displays," in IEEE Transactions on Mobile Computing, vol. 11, no. 9, pp. 1587-1599, Sept. 2012, doi: https://10.1109/TMC.2011.167

[11] Wee, Tan & Balan, Rajesh. (2012) Adaptive Display Power Management for OLED Displays. ACM SIGCOMM Computer Communication Review. 42. 10.1145/2342480.2342487

[12] R. Mertens, "Choosing color sets can reduce OLED power consumption by up to 40%," [Online]. Jul 09, 2009, Available: https://www.oled-info.com/choosing-color-sets-can-reduce-oled-power-consumption-40 Date accessed: 15 June. 2021

[13] A. Laaperi, "Disruptive Factors in the OLED Business Ecosystem," Information Display Magazine, vol. 25, pp. 8-13, Sept. 2009.

[14] Minyoung Park and Minseok Song, "Saving Power in Video Playback on OLED Displays by Acceptable Changes to Perceived Brightness," J. Display Technol. 12, 483-490 (2016)

[15] Y. -G. Shin, S. Park, Y. -J. Yeo, M. -J. Yoo and S. -J. Ko, "Unsupervised Deep Contrast Enhancement With Power Constraint for OLED Displays," in IEEE Transactions on Image Processing, vol. 29, pp. 2834-2844, 2020, doi: https://10.1109/TIP.2019.2953352

[16] S. BoHum, K. Soontae. Enhancing Energy Efficiency on OLED Display Equipped Mobile Device.

[17] K. Zhang, W. Zuo, Y. Chen, D. Meng and L. Zhang, "Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising," in IEEE Transactions on Image Processing, vol. 26, no. 7, pp. 3142-3155, July 2017, doi: https://10.1109/TIP.2017.2662206

[18] O. Oktay et al., "Anatomically Constrained Neural Networks (ACNNs): Application to Cardiac Image Enhancement and Segmentation," in IEEE Transactions on Medical Imaging, vol. 37, no. 2, pp. 384-395, Feb. 2018, doi: https://10.1109/TMI.2017.2743464

[19] Y. Nam, D. Choi and B. C. Song, "Power-Constrained Contrast Enhancement Algorithm Using Multiscale Retinex for OLED Display," in IEEE Transactions on Image Processing, vol. 23, no. 8, pp. 3308-3320, Aug. 2014, doi: https://10.1109/TIP.2014.2324288

[20] D. Jahier Pagliari, E. Macii and M. Poncino, "LAPSE: Low-Overhead Adaptive Power Saving and Contrast Enhancement for OLEDs," in IEEE Transactions on Image Processing, vol. 27, no. 9, pp. 4623-4637, Sept. 2018, doi: https://10.1109/TIP.2018.2844722

[21] P. Arbeláez, M. Maire, C. Fowlkes and J. Malik, "Contour Detection and Hierarchical Image Segmentation," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 5, pp. 898-916, May 2011, doi: https://10.1109/TPAMI.2010.161

[22] F. Yu and V. Koltun, "Multi-Scale Context Aggregation by Dilated Convolutions," in International Conference on Learning Representations (ICLR), May 2016.

[23] D. Chen, J. Liao, L. Yuan, N. Yu and G. Hua, "Coherent Online Video Style Transfer," 2017 IEEE International Conference on Computer Vision (ICCV), 2017, pp. 1114-1123, doi: https://10.1109/ICCV.2017.126

[24] Z. Yi, H. Zhang, P. Tan and M. Gong, "DualGAN: Unsupervised Dual Learning for Image-to-Image Translation," 2017 IEEE International Conference on Computer Vision (ICCV), 2017, pp. 2868-2876, doi: https://10.1109/ICCV.2017.310

[25] Kingma, D.P., Ba, J.L., "Adam: a method for stochastic optimization". In: International Conference on Learning Representations, pp. 1–15 (2015).

[26] C. Lee, C. Lee, Y. Lee and C. Kim, "Power-Constrained Contrast Enhancement for Emissive Displays Based on Histogram Equalization," in IEEE Transactions on Image Processing, vol. 21, no. 1, pp. 80-93, Jan. 2012, doi: https://10.1109/TIP.2011.2159387

[27] L. Ding andA.Goshtasby, "On the canny edge detector," Pattern Recognition, vol. 34, no. 3, pp. 721–725, 2001. ISSN 0031-3203, doi: https://doi.org/10.1016/S0031-3203(00)00023-6

[28] T. K.Wee andR. K. Balan, "Adaptive display power management for OLED displays," in Proceedings of the 1st ACM SIGCOMM Workshop on Mobile Gaming (MobiGames '12), pp. 25–30, Helsinki, Finland, August 2012. doi: https://doi.org/10.1145/2377677.2377770

[29] X. Chen, Y. Chen and Chun Jason Xue, "DaTuM: Dynamic tone mapping technique for OLED display power saving based on video classification," 2015 52nd ACM/EDAC/IEEE Design Automation Conference (DAC), 2015, pp. 1-6, doi: https://10.1145/2744769.2744814

[30] Mian Dong, Y. K. Choi and L. Zhong, "Power modeling of graphical user interfaces on OLED displays," 2009

46th ACM/IEEE Design Automation Conference, 2009, pp. 652-657.

[31] M. Dong and L. Zhong, "Chameleon: A Color-Adaptive Web Browser for Mobile OLED Displays," in IEEE Transactions on Mobile Computing, vol. 11, no. 5, pp. 724-738, May 2012, doi: 10.1109/TMC.2012.40. https://www.sigmobile.org/mobisys/2011/program.html

[32] J. Yu et al., "Sensing Human-Screen Interaction for Energy-Efficient Frame Rate Adaptation on Smartphones," in IEEE Transactions on Mobile Computing, vol. 14, no. 8, pp. 1698-1711, 1 Aug. 2015, doi: https://10.1109/TMC.2014.2352862

[33] Android Developers. Hardware acceleration. [Online]. Available: https://developer.android.com/guide/topics/graphics/hardware-accel.html Date accessed: 19 Nov. 2020

[34] Xiang Chen, Jian Zeng, Yiran Chen, Wei Zhang and Hai Li, "Fine-grained dynamic voltage scaling on OLED display," 17th Asia and South Pacific Design Automation Conference, 2012, pp. 807-812, doi: https://10.1109/ASPDAC.2012.6165066

[35] X. Chen, J. Zheng, Y. Chen, M. Zhao and C. J. Xue, "Quality-retaining OLED dynamic voltage scaling for video streaming applications on mobile devices," DAC Design Automation Conference 2012, 2012, pp. 1000-1005.

[36] B.-H. Lee and Y.-J. Kim, “Dynamic Voltage Scaling Using Scene Change Detection for Video Playback on Mobile AMOLED Displays,” in Proceedings of the 2016 International Symposium on Low Power Electronics and Design. Association for Computing Machinery, New York, NY, USA, 2016, pp. 302–307. doi: https://doi.org/10.1145/2934583.2934640

[37] T. Chang, S. S. Xu and S. Su, "SSIM-Based Quality-on-Demand Energy-Saving Schemes for OLED Displays," in IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 46, no. 5, pp. 623-635, May 2016, doi: https://10.1109/TSMC.2015.2455020

[38] P. Arbeláez, M. Maire, C. Fowlkes and J. Malik, "Contour Detection and Hierarchical Image Segmentation," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 5, pp. 898-916, May 2011, doi: https://10.1109/TPAMI.2010.161

MANAS Journal of Engineering
ISSN 1694-7398 | e-ISSN 1694-7398

Volume 9, Issue 1, (2021) Pages 66- 71
https://doi.org/10.51354/mjen.907396

Potential geoheritage assessment; Dereiçi travertines, Başkale, Van (east anatolian Turkey)

Çetin Yeşilova

Van Yüzüncü Yıl University, Faculty of Engineering, Department of Geological Engineering, Van 65000, Turkey, cetinyesilova@gmail.com, ORCID: 0000-0002-8884-0842

ABSTRACT

Dereiçi travertines famous for their Pamukkale-like conveyors are located to the south of the Başkale district of Van. The travertines are located after the 30th km of Başkale - Hakkâri highway and after 12 km by the road separating to the west. In this study, the morphological analyzes of travertines bearing geological heritages and their evaluation as "geo-tourism" were investigated. The oldest travertines in the region are the Layer-Type Traverteners at the base. On the Layer Type Travertines are the shear cracks developed in the K 10 B and K 40 D directions due to regional compressions and the Ridge Type Travertines formed by the waters from this crack. Finally, the waters draining down the slope from the cracks opening in the northeastern slope of the Ridge Type Travertines form the old and current Terrace Type Travertines.There are also small-sized Cave Type Travertines in the void spaces of various sizes developed in back-type travertines. When considered in potential of geoheritage, the region draws attention with two Ridge Type Travertines, a traverten bridge, some non-active terraced travertine, and a waterfall along a shear crack.In addition, the appearance of the natural walls formed by the downwind water due to the inclination of the Ridge Type Travertines along the northeast patch is also worth seeing. Dereiçi Travertines is one of the most important regions to be evaluated and protected from the perspective of geoheritage. However, these geological assets have not been found to be interesting, but at the same time, they have recently become the targets of the managerships looking for marble industry. Moreover, because of the increasingly decreasing hydrothermal waters coming out of the cracks, some of these terraces are dried up and some of them are coming to the stage of drying.The measures to be taken in order to continue natural formation are discussed in this article.

ARTICLE INFO

Research article

Received: 31.03.2021
Accepted: 17.05.2021

Keywords:
Dereiçi,
travertine,
geoheritage,
Başkale,
Van

1. Introduction

Geological units that record all the evolution of the time they were formed are called geological sites (geocytes) [1]. These geosites can be a rock sequence, fossil assemblages, structure, texture or a specific region [2, 3]. Geosites that facing the threat of extinction are called "geological heritage" [2, 3]. Special nature protection zones that contain the same or different types of geological heritage are called geoparks [1]. Geoparks have an important place and tourism potential in this sense.

There are many studies on the geoparks in Turkey [3-8]. In addition, important studies have been carried out for the recognition and protection of geoparks [9-11]. In recent years, Çiftçi and Güngör have worked to standardize Geoparks and inventory studied in these areas [12-15].

In this study, as suggested by Çiftçi and Güngör, it is aimed to contribute to the preparation of the "geotourism oriented" geological heritage inventory of the Dereiçi (Başkale) region [13]. De Lima et al. (2010) stated that in order to present such inventories, subject, scope, scale and purpose of use should be determined. Such a study can only be carried out within the scope of an envisaged geopark study [16].Visual features come to the fore in geomorphological elements. These visual features are partly dependent on subjective evaluations. However, such structures also have a scientific value as representing any period of geological history. As a suggested in Çiftçi and Güngör, geosites within the scope of a geopark are classified according to their scientific and visual value, get a certain score and their weight in the geopark is controlled in this way [13].

2. Regional geology

Dereiçi Village is located approximately 23 km southeast of Başkale district of Van (Fig. 1). The basement in the study area are represented by Paleozoic - Mesozoic metamorphic rocks belonging to the Bitlis Massif (Fig. 1). Bitlis Massif is composed of metamorphic rocks in the P-T conditions ranging from low to high, it is one of Turkey's largest and most important massifs. According to many researchers [17-20], this massif is divided into two different main units as Lower Metamorphites and Upper Metamorphites. Upper Cretaceous aged Yüksekova Complex unconformably overlies Bitlis metamorphics. This complex was firstly described by Perinçek [21].

The Yüksekova Complex is the dominant unit in the study area and has a wide distribution in the regions of Yılan Village, Meşeli Village, Boğaz Hill and Billurik River. This complex, represented by very different rock assemblages in Hakkari and its vicinity, was studied under the name of three different rock groups in the study area. These are granitic rocks, dioritic rocks and aplitic veins cutting these rocks. The Durankaya Complex unconformably overlies this complex and outcrops over a large area in the study area. The Durankaya Complex was first described by Perinçek [22]. The type section of the unit is seen around Durankaya Village, 22 km southwest of Hakkari Province, and is represented by shale, sandstone and limestone. It also contains serpentinite, gabbro, amphibolite and marble levels in places. Shales are extremely deformed and gray in color. Limestones with a thickness of 700 m are claret red and red in color and in places in the form of clayey limestone. The unit was deposited in slope-deep sea environment [22]. The travertines and current alluviums, which are related to NE-SW trending strike-slip faults, formed as a result of tectonic movements in the region, constitute the youngest units in the study area (Fig. 1).

The aim of this study is to determine the geomorphological geoheritage elements of the travertine formations [23], which were partially examined and defined before, to evaluate them in terms of geotourism and to discuss the conservation conditions.

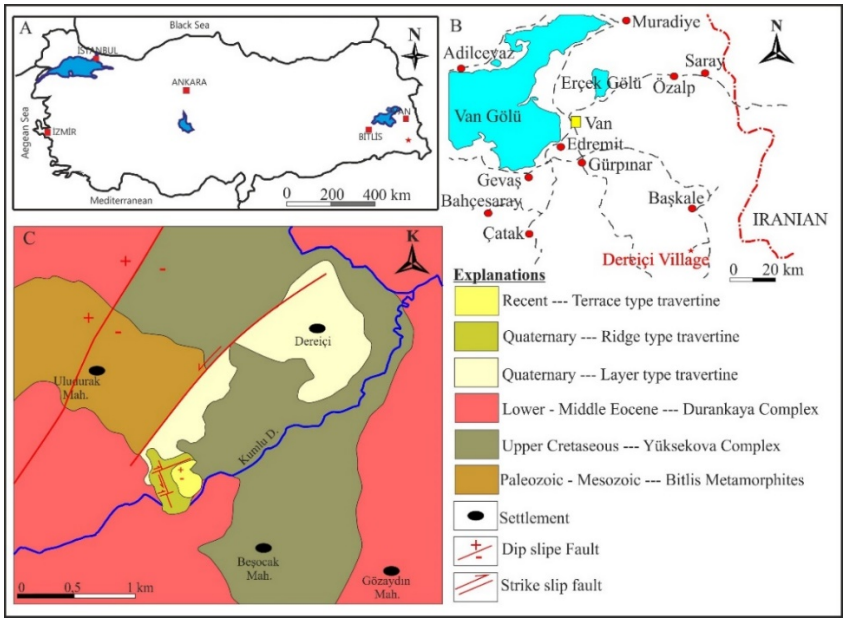


Figure 1. *Of the study area, Turkey (A) and in Van Province (B) map showing the location. C, Geological map of the study area. The red star shows the location of the travertines.*

3. Results

According to the studies conducted on the travertines in the region in previous years [23], there are three different geomorphological processes and three different travertine formation:

1. "Layer Type Travertines", most of which are formed by crystalline crustal facies,
2. An mode-I crack and "Ridge Type Travertines" developed due to this crack.
3. "Terrace Type Travertines", which developed due to mode-I cracks and still continue to form(Fig. 2). These formations will be introduced sequentially from south to north.

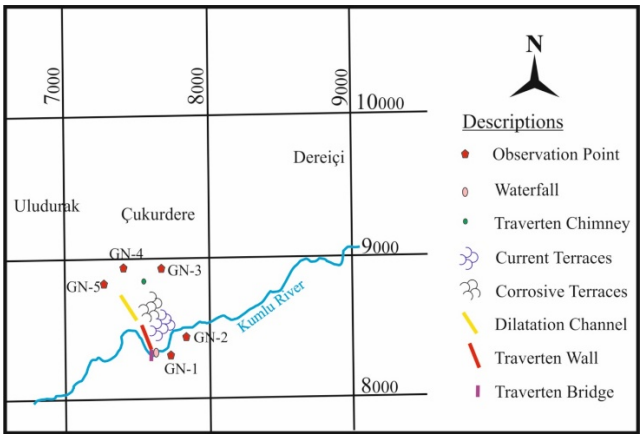


Figure 2. *Potential geoheritage elements and locations in the study area.*

3.1. Mode-I crack, walking channel and travertine wall

It is located on a slope at the southernmost point of Dereiçi Travertines and consists of "ridge type" travertines formed by the waters flowing through a crack opened in the N 40^0 E direction (Fig. 3A). The southeastern parts of the crack opened up in time and turned into a channel with a width of 47-68 cm and a depth of 40-80 cm (Fig. 3B). On the northeast slope of the travertines, the waters flowing along the sloping surface formed a steep wall on this slope (Fig. 3C). The wall is 80 m long and has a height of 3 to 5 m and a width of 1.2 to 3.5 m. Travertines consist of light brown - beige colored layers varying between 0.7 - 10 cm. The fact that they are generally composed of gas bubble and crystalline crustal facies indicates that these travertines are precipitated by gases-rich hot waters [24].

3.2. Travertine chimney

It is a small fracture that develops in the area and a travertine formation that grows upward through the fracture. It is located in front of the slope to the north of the walking channel (Fig. 2). Its thickness is between 1.3 and 3 m, its height is 2.7 m and its width is around 2 m. This formation stands as if it witnessed history. It consists of cream beige crystalline crust facies (Fig.3D).

Figure 3. *A. Mode-I crack forming on the ridge type travertines, B. Channel structure formed in ridge type travertines, C. Wall formed in back type travertines, D. Travertine chimney.*

3.3. Travertine bridge

It is a rare beautiful travertine bridge formed by the opening crack in the N 40^0 E direction over the existing stream in the area (Fig. 2, 4A). The travertine bridge was formed as a result of the carving and giving way of the current cracks in the ridge type travertine, which blocks the front of the river. The bridge is 5 m high, 3.3 m wide and 1.5 - 2.5 m thick.

3.4. Waterfall

The travertine waterfall, which was formed as a result of the discharge of the waters coming out of this bridge along a steep slope in the north of the travertine bridge, presents a natural beauty in itself (Fig. 4B). The waterfall has 3 steps. Its total height is 6 m, and its width is around 3 m. Although the amount of water in the waterfall increases in rainy seasons, there is a continuous flow.

3.5. Travertine Pool.

Terrace type travertines formed by old and current water outlets form these pools (Fig. 4C D). These draw attention with their similarities to Pamukkale Travertines. Pisolitic type travertine formation is observed in the terrace pools of travertines. These waters, which are concentrated in volatile components, form "crystalline crust" on the terrace edges and "gas-bubble" and "paper thin–raft type" travertines in the travertine pools. More than 60% of the terrace type travertines dryed out as a result of the decrease in the water forming them and the change of direction (Fig. 4E). If no action is taken, in the future, all of these travertines will start to dry and go through the wear process. The pools range from a few cm wide to 2 - 3 m wide. The deepest pool is around 50 cm. At the exit of the waters forming the current pools, bad-smelling gases also exit. From the discolorations on the edges of the travertine pools, it is understood that the gases released are rich in sulfur (Fig. 4F).

Figure 4. *A. Travertine bridge, B. Waterfall structure formed on steep slopes, C. Terrace of travertine, D. Travertine pool, E. The travertine terraces that have begun to wear out, F. Discoloration that develops depending on the amount of sulfur in the water.*

4. Conclusion

In studies conducted on travertines, *crystalline crust, paper-thin raft type, gas buble, shurb type* and *pisolitic travertine* facies were found. Both their morphology and facies characteristics show that Dereiçi Travertines were formed by the flow of hydrothermal waters rich in volatile components down the slope to the surface along a shear crack..
The travertines in the region have a great potential in terms of geotourism. However, both short-term and long-term studies need to be done.

In the short term; Marble cutting and excavation operations have been identified by some investors in the region (Fig. 5A). For this reason, in order to protect and maintain the region, constant meetings have been held with the Governorship, Relevant municipalities and non-governmental organizations. In this context, the Law No. 2863, Article 13 / A of the Decree No. 644 (LAW; 2013) [25]; It has been taken under protection within the scope of the "Regulation on Procedures and Principles Regarding the Determination, Registration and Approval of Protected Areas" (LAW; 2012) [26], published in the Official Gazette dated 19.07.2012 and numbered 28358 (Official Gazette dated 04.03.2020 and numbered 31058, decision number: 2195). However, in addition to taking the area under protection, there should be restrictions on daily visitors' activities.

Only 40% of the Dereiçi Travertines are still up-to-date (Fig. 5B). The gradually decreasing water exit in the region will cause the travertines to dry and then erosion. With the detailed geological, hydrogeological and geophysical studies to be done, the chemistry, depth, reserve etc. of the underground water. must be determined. By bringing these underground waters to the surface, the regeneration of travertines can be continued without disturbing the naturalness of the travertines.

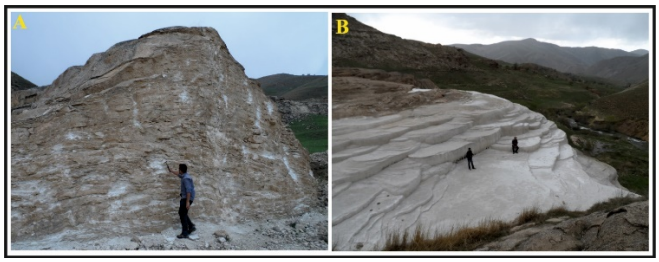


Figure 5. *A. The damages caused by the marble operators to the tarvertines and natural structure in the region, B. Terrace type travertines that are still being formed.*

Access to the travertines is via a 20 km stabilized road from the main road. This road is under heavy snow cover most of the year. Therefore, it is recommended that this road be improved and asphalted as soon as possible.

In the long term; These are divided into two as sociological and tourism studies. Sociological studies of the area where the travertines are located should be conducted by sociologists who are experts in their field, to raise awareness about the travertines in the region, and to organize meetings with local administrations and investors in the region about the geotourism potential of travertines. Studies should be carried out to reveal the importance and added value potential of the said travertines for tourism agencies operating in the national and Lake Van Basin, with studies to be organized by tourism professionals who are experts in their fields. In addition, even if it is simple, the preparation of the signs that introduce both the geological and geomorphological features of these travertines will contribute to the awareness to be created on the local people.
In the Lake Van Basin and its eastern part, dozens of geological heritage elements that have both scientific value and touristic potential can be counted at once, but unfortunately this has no practical benefit. With the detailed studies to be done, the inventories of the geological and cultural heritage elements in the region should be prepared in accordance with the "geological heritage accreditation" stated by Çiftçi and Güngör (2017b). These inventories, supported by scientific data, can only be contributed to the sustainable

development of the people of the region and country tourism when they are owned by the local governments in the region.

Acknowledgements.

This study was carried out within the scope of Yüzüncü Yıl University Scientific Research Projects project no "2014-MİM-B074". I would like to thank Cahit Helvacı, Yahya Çiftçi, Mustafa Açlan, Yıldırım Güngör and Pelin Güngör Yeşilova for their contribution in the preparation of the article and field studies.

References

[1]. Kazancı N., Özgen Erdem N., Erturaç MK., "Kültürel Jeoloji ve Jeolojik Miras; Yerbilimlerinin Yeni Açılımları", Türkiye Jeoloji Bülteni, 60/1 (2017), 1-16.

[2]. ProGeo Group., "A first attemt at a geosites frameworkfor Europe -an lUGS initiative to support recognition of World heritage and European geodiversity". Geologica Balcanica 28, (1998), 5-32.

[3]. Kazancı N., "Dünyada ve Türkiye'de Jeosit-Jeopark-Jeomiras Olgusuna Yaklaşımlar". Kızılcahamam-Çamlıdere Jeopark ve Jeoturizm Projesi Raporu", (2010a) 76p, Ankara.

[4]. Kazancı N., "Milli Parklarda Jeolojik Miras 1, Soğuksu Milli Parkı (Kızılcahamam-Ankara)", Çevre, Atmosfer, Yer ve Deniz Bilimleri Araştırma Grubu, (2007) 61p, Ankara.

[5]. Kazancı N., "Jeolojik Koruma (Kavram ve Terimler)", Jemirko ve TMMOB Jeoloji Mühendisleri Odası Yayınları, (2010b) 60p, Ankara.

[6]. Kazancı N. Şaroğlu, F., "Türkiye Jeositleri Çatı Listesi". 62. Türkiye Jeoloji Kurultayı (13-17 Nisan 2009) Bildiri Özleri Kitabı-I, Jeoloji Mühendisleri Odası, Ankara, (2009) 266-267.

[7]. Gürler G., Timur E., "Jeoparkların Koruma Kullanım Yöntemlerinin Belirlenmesi; Karapınar Potansiyel Jeopark Alanı İçin Bir Değerlendirme, Türkiye". Proceedings of The Second International Symposium on Development Within Geoparks Environmental Protection and Education, Lushan, Jiangxi Province, China, 12-15 June, (2007).

[8]. Koçan N., "Ekoturizm ve Sürdürülebilir Kalkınma: Kızılcahamam-Çamlıdere (Ankara) Jeopark ve Jeoturizm Projesi". Karadeniz Fen Bilimleri Dergisi, 2/6, (2012a) 69-82, Trabzon.

[9]. Gürler M., "Doğal koruma alanlarının belirlenmesi ve planlanmasında jeoloji. JMO Semineri; Aynı yazı ve yazar": Birlik Haberleri, Sh. (1997) 46-48.

[10]. Gürler M., "Ekolojik planlamada jeolojik oluşumlar için öneri bir sınıflama modeli". TMMOB Birlik Haberleri, Mayıs-Haziran Sayısı, (1999), Ankara.

[11]. Gürler, M., "Anıt nitelikli jeolojik oluşumlar ve koruma çalışmaları". JMO Mavi Gezegen Popüler Bilim Dergisi, S. 4, (2001) 10-11, Ankara.

[12]. Çiftçi Y., Güngör Y., "Jeosit Tanımlama ve Jeopark Düzenleme Standartları Üzerine Bir Çalışma", 67. TJK. Bildiri özleri kitabı, (2014) p. 422, Ankara.

[13]. Çiftçi Y., Güngör Y., "Jeopark Projeleri Kapsamındaki Doğal ve Kültürel Miras Unsurları İçin Standart Gösterim". Maden Tetkik ve Arama Dergisi, 153, (2016) 223-238.

[14]. Çiftçi Y., Güngör Y., "Jeolojik Miras Envanteri Hazırlamada Terminoloji, Yöntem, Kurumsal Sahiplenme ve Eşgüdüm". 70. Türkiye Jeoloji Kurultayı Bildiri Özleri, (2017a), p. 620 - 621, 10-14 Nisan 2017, Ankara

[15]. Çiftçi Y., Güngör Y., "Dünya'da ve Türkiye'de Jeolojik Miras Envanteri Hazırlama ve Jeokoruma Çalışmaları: Akreditasyon, Ulusal Kodlama Sistemi ve Etkileşimli Veri Tabanı Önerisi". 70. Türkiye Jeoloji Kurultayı Bildiri Özleri, (2017b), p. 252 - 253, 10-14 Nisan 2017, Ankara

[16]. De Lima F.F.,Brilha J.B., Salamuni E., "Inventorying Geological Heritage in large territories: a methodological proposal applied to Brazil". Geoheritage, 2, (2010) 91-99.

[17]. Yılmaz O.,. "Cacas bölgesi (Bitlis Masifi) kayaçlarının petrografik ve stratigrafik incelenmesi". Türkiye Jeoloji Kurumu Bülteni, 18/1, (1975) 33-41.

[18]. Tolluoğlu A. Ü., "Mutki (Bitlis) yöresi metamorfiklerinin petrografisi/petrolojisi". Yüksek Mühendislik Tezi, Hacettepe Üniversitesi, (1981), 86 p. Ankara (yayımlanmamış).

[19]. Şengün M., "Tatvan güneyinin (Bitlis masifi) jeolojik/ petrografik incelenmesi": Doktora tezi, (1984). 157p. (yayımlanmamış).

[20]. Göncüoğlu, M. C. ve Turhan, N., "Bitlis Metamorfik kuşağı orta bölümünün temel jeolojisi". MTA Rapor No. (1985), 7707 (yayımlanmamış).

[21]. Perinçek D., "The Geology of Hazro-Korudağ-Çüngüş-Maden-Ergani-Elazığ-Malatya area": Guide Book, TJK yayını, (1979), Ankara

[22]. Perinçek D., "Hakkari ili ve dolayının stratigrafisi, GDA Türkiye": TPJD Bülteni, 2/1, (1990), 21-68.

[23]. Yeşilova Ç., Açlan, M., Güngör Yeşilova, P., "Van Gölü Çevresindeki Travertenlerin Jeoturizm Açısından Değerlendirilmesi", Doğu Anadolu Jeoloji Sempozyumu , (2015), 252 - 253 p , Van, Türkiye , 07 - 12 Eylül 2015

[24]. Ayaz E., "Travertenlerde Gözlenen Morfolojik Yapılar ve Tabiat Varlığı Olarak Önemleri". Cumhuriyet Üniversitesi Mühendislik Fakültesi Dergisi, Seri A-Yerbilimleri 19/2, (2002), 123-134, Aralık 2002

[25]. ÇŞB., "Korunan Alanların Tespit, Tescil ve Onayına İlişkin Usul ve Esaslara Dair Yönetmelik", (2012), Çevre ve Şehircilik Bakanlığı, Tabiat Varlıklarını Koruma Genel Müdürlüğü.

[26]. ÇŞB., "Doğal Sit Alanlarının Değerlendirilmesine İlişkin Yönetmelik", (2013), Çevre ve Şehircilik Bakanlığı, Tabiat Varlıklarını Koruma Genel Müdürlüğü.

MANAS Journal of Engineering
ISSN 1694-7398 | e-ISSN 1694-7398

Volume 9 (Issue 1) (2021) Pages 72-84
https://doi.org/10.51354/mjen.809844

Capacitated Multiple Allocation Hub Covering Flow Problem

Nazmi SENER[1,*], Orhan FEYZİOGLU[2]

[1] Department of Industrial Engineering, Galatasaray University, Istanbul, Turkey

[2] Department of Industrial Engineering, Galatasaray University, Istanbul, Turkey

ABSTRACT

The aim of the Capacitated Multiple Allocation Hub Covering Flow Problem is to find the optimal design for hub-and-spoke networks while taking into account hub opening and demand routing costs. Every network node has the potential to be a hub and demand from an origin to a destination must be sent through at least one hub. The network is incomplete in the sense that the maximum allowed or coverage distance between any opened hub and demand origin/destination is predefined. It is assumed that there is a cost saving to route demand via hubs due to consolidation. Another important issue is the consideration of capacity restrictions imposed on network links and opened hubs. The problem is developed as a mixed-integer linear optimization problem. According to the results obtained from computational experiments, we show that taking into account both flow related costs and capacities of network components concurrently is very important to have a cost effective design.

ARTICLE INFO

Research article

Received: 15.11.2020
Accepted: 25.01.2021

Keywords:
Facility Location, Network Design, Hub-and-Spoke Networks, Capacitated Hub Covering, Multiple Allocation

**Corresponding author*

1. Introduction

Locating economically attractive hub facilities through which demand flows are to be routed from origins to destinations is the subject of hub location problems (HLP). As for example, flows can be associated with passengers or freight, and origins/destinations with cities in the physical world. Flow form an origin to a destination can be assured with a direct trip (bypassing hubs) or via paths that visit hubs. At hubs, traffic arriving from several origins can be aggregated or they can be disaggregated to depart several destinations. On the one hand, fewer links would be required to connect origins and destinations with this redirection compared to direct connections. On the other hand, economies of scales can be achieved through consolidation of flows at hubs and thus costs related to flows can be reduced. The main decisions for HLP are to locate hub nodes and route flows so as to satisfy the demand. The network for a HLP consists of nodes so called as hub, origin and destination and arcs connecting hubs to other hubs, origins to hubs, hubs to destinations, and occasionally origins directly to destinations. Conform to the traditional location theory, locations of facilities and linking of points in the network to facilities are two distinctive decisions. Within this context, the generic HLP can be regarded as a network design problem with location. Campbell and O'Kelly [1] summarize the differentiating characteristics of HLP as in the following:

1. Demand is related with origin-destination (OD) node pairs and not with individual nodes,
2. Demand flows pass through hub nodes,
3. Location of hub nodes must be identified,
4. Routing flows via hub nodes is a requirement or has a benefit,
5. Problem objective is formulated based on the hub locations and flow routing.

This definition covers a wide range of problems which have not been viewed as HLP. Campbell [2] defines two additional important features of HLP that formed a base for further research:

1. Hub nodes that can be visited on a path linking an OD node pair can be at most two,
2. Flow from an origin directly to a destination is not allowed.

HLP can be classified according to the following characteristics [3]:

- *Solution space for locating hub nodes*: all network nodes, a subset of network nodes or the continuous space (the domain is a plane or a sphere).

- *Number of hubs to locate*: single or multiple, specified in advance or determined at the optimum solution.
- *Node allocation*: To a single hub node or single allocation (S), or to multiple hub nodes or multiple allocation (M).
- *Capacity of a hub/link*: capacitated or uncapacitated.
- *Cost of establishing hub nodes*: N/A, fixed or variable.
- *Cost of allocating nodes*: N/A, fixed or variable.
- *Objective*: Minimizing the maximum path cost that is due to the routing of demand flows from origins to destinations (minmax), or minimizing the total cost of establishing hubs and assigning other nodes to hubs (minsum).

The main application areas of HLP are in transportation and telecommunications. Freight or passengers are carried by vehicles on infrastructures such as roadways, railways, airways or waterways in transportation hub location problems, and hub nodes are located by taking into consideration the associated distance based travel time and/or cost. Meanwhile, hub facilities such as routers, switches and concentrators are located to provide communication among a set of nodes in telecommunications hub location problems. As the electronic data is moved using physical links (cables) or though the air (microwaves), there may not be significant distance based traveling costs in this type of network. Both transportation and communication networks often have very similar abstract models, but the operations, relevant costs, service measures, and constraints are quite different. Routing protocols for transportation and telecommunication usually differ, as large communication networks require packet switching while individual traffic units (passengers/shipments) are usually not divisible in transportation networks.

In this paper, the design of incomplete hub-and-spoke networks are investigated where every network node has the potential to be a hub and the demand from an origin to a destination must be sent from at least one hub. Meanwhile, the subnetwork formed by hubs is assumed to be complete. The aim to minimize the total cost due to the design (opened hubs) and operations (transportation). Another important issue is the consideration of capacity restrictions imposed on network links and opened hubs. As we will show in the sequel, taking into account both flow related costs and capacities of network components concurrently for the hub covering problem (HCP) is very important to have a cost effective design. We provide a detailed literature survey for HLP and HCP in the next section. We introduce the notation and mathematical formulation of the problem in Section 3. Based on two well-known benchmark data set from the literature, numerical trials are carried out to determine the response of our mathematical model to variations of parameters. Results of these experiments together with some practical insights are provided in Section 4. The last section of the paper includes some concluding remarks and perspectives.

2. Literature Survey

The first study including similar concepts to HLP is due to Hakimi [4]. Application of the hub location for aviation is discussed by Toh and Higgins [5]. O'Kelly [6, 7] authored pioneering papers introducing the mathematical formulation and solution methods for HLP. Many papers have been published with a significantly increasing trend since then. There exist comprehensive literature surveys on hub location problems [1, 3, 8–11], so we refer the interested reader to these works and the references therein. HLP has different types such as hub covering problem (HCP), hub center problem, hub median problem, and hub arc location problem. In this study, we are focused on a particular case of HCP designed as the hub covering flow problem (HCFP).

HCP was first proposed by Campbell [2] in analogy to the set covering problem. In this work, hub set cover and maximal hub cover problems with single and multiple allocation are provided and three different coverage cases are enlisted. Let us denote two hub nodes as k and l. Then an OD node pair (i, j) is covered if the length(s) given in the following cases do(es) not exceed a preselected particular value: (1) length of the path (i, k, l, j); (2) length of each link on the path (i, k, l, j); and (3) lengths of links (i, k) and (l, j). In our study, we take into consideration the last covering case.

Yetis Kara and Tansel [12] give a novel formulation for SHCP that is different from the hub set cover problem introduced in Campbell [2]. In terms of average CPU times and storage requirements, the linearization of this model performs pretty better than the best performing linearization of the Campbell [2]'s model. In their numerical experiments, they used the data set that is created based on the Civil Aeronautics Board (CAB) Survey of 1970 airline data in the US. Ernst et al.[13] suggest a new compact formulation for the uncapacitated p-SHCP based on the coverage radius concept. Auxiliary variables are included to this formulation to determine the distance between each hub and the furthest node assigned to it. Despite being slightly weaker than the others, it

was shown empirically to perform better. Wagner [14] proposes new model formulations for S/MHCP which include quantity dependent and/or independent transportation times. The author uses AP and CAB data sets to experimentally show that models with quantity independent transportation times perform better. AP data set represents distribution divisions of the Australian Post [15].

Weng and Wang [16] give a new model for multiple allocation HCP and propose to solve large instances with scatter search and genetic algorithm. Qu and Weng [17] propose an evolutionary algoritm based on route reconnecting to solve a new reformulation of maximal MHCP. A single allocation HCP for cargo delivery applications is investigated by Alumur Alev and Yetis Kara [18]. In this study, the assumption of completely connected hub network is relaxed and the objective is to design a network constrained by a specified path traveling time. A closely related model for the SHCP over incomplete hub networks is introduced in Calik et al. [19]. The model objective is locate hub nodes, establish interhub links and allocate non-hub nodes to hub nodes such that the travel time on the path joining any OD pair is less than a specified time bound. An efficient tabu-search based heuristic is proposed and its performance is tested on TR and CAB data sets.

All non-hub nodes must be covered by hub nodes in HCP, but in most of the literature, the cost of transportation between OD node pairs is not accounted for. This can be an important shortcoming. For example, transportation costs should not be overlooked by passenger airlines when their operational hub locations are to be identified, given that the cost efficiency is a key for the survival of these companies. In consequence, Lowe and Sim [20] come up with HCFP which aims to design hub-and-spoke networks at a minimum cost by locating hub nodes and routing demand flows through these nodes given coverage constraints. Design costs they consider are fixed hub opening costs and variable transportation costs. Alumur et al. [21] consider transportation travel times and costs together while formulating multimodal HCP. In their model, different transportation modes between hub nodes and different types of service time commitments for paths joining OD node pairs are allowed.

Capacities of hubs and links of the each route is a less considered issue in the literature, especially for HCP. Campbell [22] initially propose a mixed integer linear optimization model for the capacitated MHLP with four indexed variable. Aykin [23] presents the capacitated hub-and-spoke network design problem with fixed capacity. Flows between OD node pairs can bypass hubs in the associated mathematical model. As the solution procedure, the author provides a branch-and-bound procedure and a heuristic methodology dividing the solutions set on the basis of hub placements. Bryan [24] extends HLP model proposed in O'Kelly and Bryan [25] in many different ways. The base multiple allocation model unequivocally represents scale economies by permitting interhub expenses to be a component of streams with per unit cost diminishing as streams increment.In the extended models, the effect of imposing minimum and maximum flow limits on interhub links is studied. A new formulation for the capacitated SHLP is investigated by Ernst and Krishnamoorthy [26]. They develop heuristic approach for its algorithm based on random descent and simulated annealing.

Ebery et al.[27] provide formulations with three indexed variable for the capacitated MHLP. They incorporate the upper bound obtained from an efficient heuristic in a linear optimization based branch and bound solution scheme to solve large instances. Marín[28] also consider new formulations for capacitated MHLP, and obtain better computational results with the given resolution techniques. As a natural extension of the uncapacitated one-stop (no interhub flows) hub-and-spoke model, Sasaki and Fukushima [29] give a new formulation for one-stop capacitated model. More precisely, arc and hub capacity constraints are included to the model, and a branch-and-bound based exact solution method is used to solve this model. A capacitated SHLP encountered in the design of telecommunications networks is investigated by Carello et al. [30]. In this problem, the subnetwork formed by hub nodes is required to be fully connected and the traffic passing through each hub node is limited by capacity constraints. The goal is to limit the amount of fixed expenses of opening and preparing hubs and linking expenses of introducing on each edge the capacity expected to route the traffic on the edge itself (modular capacity). A local search methodology is introduced and various metaheuristic procedures have been developed based on this approach.

The same problem is investigated by Yaman and Carello [31]. Objective of the associated model minimizes the total fixed cost of adding hubs and determining the needed capacity on each link. Capacity on an edge is adjusted by setting the number of edges of modular capacity. Moreover, hub capacity restricts the amount of flow that can pass through a hub rather than the incoming flow. They adapt the branch and cut algorithm and develop a two-level heuristic for this problem. In a later study, Yaman [32] focus on the star p-HLP with modular link capacities where each node is assigned to a single hub and each hub is served to a single central hub.

Rodríguez-Martín and Salazar-González[33] develop a

formulation of capacitated MHLP for incomplete hub networks where both hubs and links are capacitated. The authors propose an efficient nested two level algorithm depending on Benders decomposition to solve the problem. Another approach to include capacity related issues to HLP is given in [34]. In this work, authors formulate a bi-objective SHLP such that one of the objectives minimize the inflow processing times at the hubs instead of using hard capacity constraints for hubs to limit the inflow. Mohammadi et al. [35] provide a new multiobjective model for the capacitated SHCP and solve it by multi-objective imperialist competitive algorithm. Contreras et al. [36] consider a capacitated HLP in which hub capacities are not parameters but decision variables and implement Benders decomposition algorithm to solve it. Sedehzadeh et al.[37] study a multi-product multi-mode capacitated SHCP using queue approach. One of the objective is to minimize the sum of hub node opening costs and transportation costs, while the second one minimizes the maximum transfer time for each product and between each OD node pair by considering transportation and waiting times. This multi-objective programming problem is solved by multi-objective parallel simulated annealing algorithm. Karimia et al. [38] present a tabu search algorithm to solve the multi-modal capacitated p-SHCP over fully interconnected networks. Allocations of non-hub nodes to hubs and locations of hubs are to be found such that the travel time between any OD node pair is less than or equal a given time bound. Merakli and Yaman [39] consider a capacitated MHLP with hose demand uncertainty. Both the feasibility of the solutions and the total cost are affected by the demand uncertainty, since hub nodes are capacitated and transportation cost is a function of demand. Hoff et al. [40] solve the capacitated modular SHLP (capacity of interhub edges is increased in a modular fashion) via different combination of heuristics which is not taken into account before.

Demir et al. [41] introduce a multi-objective linear optimization model for the capacitated MHLP. They use both link and hub capacities in their model. For solving the model, they develop a multi-objective evolutionary algorithm which is a NSGA-II based heuristic. Danach et. al [42] combine vehicle capacity idea with p-SHLP. For solving the novel model, they develop a hybrid hyper-heuristic algorithm which is a usage of Lagrangian relaxation within a reinforced learning framework. Taherkhani et al. [43] consider profit based modelling framework with capacitated HLP with various demand classes with deterministic and stochastic versions. They use a novel fastened version of Benders decomposition for deterministic version. For stochastic version, sample average approximation algorithm is employed with the benefits of the improved Benders decomposition. Butun et. al. [44] formulate the capacitated directed cycle HLP considering congestion. They first linearize the model after that use a tabu search heuristics for solving the linearized model. From the survey above, multiple allocation studies are generally not studied. On the other hand, it is assumed that all possible hubs must be selected from a subset of nodes. Generally, capacity is taken as hub capacity. There is a limited number of studies investigate link capacity. Because of these, we first introduce a novel multiple allocation hub covering problem including transportation costs and we also add to the model both hub and/or link capacity constraints.

3. Mathematical Formulation

Four models of the multiple allocation hub covering flow problem (MHCFP) is introduced in this section. These mathematical models are: uncapacitated MHCFP or UMHCFP, link capacitated MHCFP or MHCFP-1, hub capacitated MHCFP or MHCFP-2 and lastly, link and hub capacitated MHCFP or MHCFP-3. These models are based on the formulation of Ebery et al. [27] and Lowe and Sim[20]. A common notation for these problems is given below:

$\mathcal{V}$	set of nodes
h_{ij}	demand flow originating from node $i \in \mathcal{V}$ destined for node $j \in \mathcal{V}$
H	total amount of flow to be sent
O_i	total demand originating from node $i \in \mathcal{V}$
D_j	total demand destined for node $j \in \mathcal{V}$
ω_{ik}	flow capacity of the links connecting nodes $i, k \in \mathcal{V}$
Γ_k	flow capacity of hub $k \in \mathcal{V}$
f_i	cost of opening a hub at node $i \in \mathcal{V}$
c_{ij}	unit flow cost for the link connecting nodes $i, j \in \mathcal{V}$
d_{ij}	length of the link connecting nodes $i, j \in \mathcal{V}$
$\mathbf{A}$	node coverage matrix (A_{ij} is 1 if node $j \in \mathcal{V}$ can be covered by node $i \in \mathcal{V}$ and 0 otherwise)
$\mathbf{B}$	path coverage matrix ($B_{ikj} = A_{ik}A_{kj}$)
α	interhub flow cost discount factor with $\alpha \in (0, 1)$

It is clear that $O_i = \sum_{j\in\mathcal{V}} h_{ij}$, $D_j = \sum_{i\in\mathcal{V}} h_{ij}$ and $H = \sum_{i,\, j\in\mathcal{V}} h_{ij}$. Hub opening decision is related with the binary decision variable, x_k, which equals to 1 if a hub is set to node k, 0 otherwise. Variable z_{ik} denotes the amount of flow sent from node i to hub node k, q_{ilj} the amount of flow sent from node i to node j through hub node l, and y_{ikl} to the amount of flow sent from node i via hub nodes k and l.

We introduce first the formulation of model **UMHCFP** below:

$$\min \sum_{k\in\mathcal{V}} f_k x_k + \alpha \sum_{i,k,l\in\mathcal{V}} c_{kl} B_{ikl} y_{ikl} + \sum_{i,k\in\mathcal{V}} c_{ik} A_{ik} z_{ik} + \sum_{i,l,j\in\mathcal{V}} c_{lj} B_{ilj} q_{ilj} \quad (1)$$

$$\text{s.t.} \sum_{k\in\mathcal{V}} A_{ik} z_{ik} = O_i \qquad i \in \mathcal{V}, \quad (2)$$

$$\sum_{l\in\mathcal{V}} B_{ilj} q_{ilj} = h_{ij} \qquad i,j \in \mathcal{V}, \quad (3)$$

$$\sum_{l\in\mathcal{V}} B_{ikl} y_{ikl} + \sum_{j\in\mathcal{V}} B_{ikj} q_{ikj} - \sum_{l\in\mathcal{V}} B_{ilk} y_{ilk} = A_{ik} z_{ik} \qquad i,k \in \mathcal{V}, \quad (4)$$

$$\sum_{i\in\mathcal{V}} B_{ilj} q_{ilj} \le D_j x_l \qquad l,j \in \mathcal{V}, \quad (5)$$

$$A_{ik} z_{ik} \le O_i x_k \qquad i,k \in \mathcal{V}, \quad (6)$$

$$q_{ilj}, y_{ikl}, z_{ik} \ge 0 \quad x_k \in \{0,1\} \qquad i,k,l,j \in \mathcal{V}. \quad (7)$$

The objective in Eq.(1) is to minimize the total cost of opening hubs and routing demand through network links by considering interhub flow cost discount factor. Eq.(2) ensures that all the demand originating from node $i \in \mathcal{V}$ is transported through hubs. Eq.(3) guarantees that the demand originating form node $i \in \mathcal{V}$ destined for node $j \in \mathcal{V}$ is transported through hubs. Eq.(4) corresponds to the flow conservation constraints at each hub. Eq.(5) and Eq.(6) together ensure no demand is transported directly between non-hub nodes. Finally, Eq.(7) shows the type of decision variables.

We also make use of the following constraints in this study:

$$z_{ik} \le \omega_{ik}(1 - x_i) + H x_i \qquad i,k \in \mathcal{V}, \quad (8)$$

$$\sum_{i\in\mathcal{V}} q_{ilj} \le \omega_{lj}(1 - x_j) + H x_j \qquad l,j \in \mathcal{V} \quad (9)$$

$$\sum_{i\in\mathcal{V}} z_{ik} \le \Gamma_k \qquad k \in \mathcal{V}. \quad (10)$$

Constraints in Eq.(8) guarantee that the amount of flow on the link connecting node i and hub node k does not exceed the link capacity ω_{ik}. In a similar fashion, constraints in Eq.(9) do not allow an amount of flow to be transported from hub node l to node j surpassing the link capacity ω_{lj}. Constraints in Eq.(10) restrict the inflow to any hub according to its capacity. When Eqs.(8-10) are added to UMHCFP given in Eqs.(1-7), the following variant models can be built:

MHCFP-1 : Eqs.(1-7), Eq.(8), Eq.(9),
MHCFP-2 : Eqs.(1-7), Eq.(10),
MHCFP-3 : Eqs.(1-7), Eq.(8), Eq.(9), Eq.(10).

To compare numerical results, we also associate mathematical models **UMHCP**, **MHCP-1**, **MHCP-2** and **MHCP-3** with UMHCFP, MHCFP-1, MHCFP-2, MHCFP-3 respectively where transportation related costs in the objective function Eq.(1) are omitted.

4. Computational Experiments

We have applied well-known and openly available datasets in our numerical study (CAB dataset with 25 nodes [45] and TR data set with 81 nodes [46]). These benchmark network data sets are all available through OR library [45]. TR data set is complete such that *unit flow costs*, *hub opening costs*, *network links' lengths* and *demand flows* are all provided. CAB data set only includes links lengths and demand flows between nodes. Accordingly, we assume that unit flow costs are proportional to link lengths such that $c_{ij} = d_{ij}/25,000$ for all $i,j \in \mathcal{V}$ and hub opening costs f_i for all $i \in \mathcal{V}$ are set all equal to 10,000, 20,000 or 30,000 for a given instance. In order to obtain reasonable results in terms of the number of opened hubs, original hub opening costs for TR data set are all multiplied with 300. CAB data set has symmetrical structure such that $h_{ij} = h_{ji}$ and $d_{ij} = d_{ji}$ for all $i,j \in \mathcal{V}$. TR data set has not this symmetrical structure.

Interhub flow cost discount factor α is set to 0.2, 0.5 or 0.8 for CAB data set, and 0.4, 0.6 or 0.8 for TR data set. The hub or *node coverage radius* is obtained by multiplying the *coverage ratio* Δ with length of the longest link of the network. Δ should be selected such that the existing network does not contain disconnected sub-networks. Hence, Δ is set to 0.6, 0.7 or 0.8 for CAB data set and 0.55, 0.65 or 0.75 for TR data set. Then, each element A_{ij} of the *node coverage matrix* is fixed to 1 if the node coverage radius is greater than or equal to the length of the link connecting nodes i and j, and 0 otherwise.

Link capacities, ω_{ik} for all $i,k \in \mathcal{V}$, are not included in the original CAB and TR data sets, so we develop

a procedure to identify them which we briefly expose here. First, UMHCFP is solved to optimality to obtain optimum link flows. As there is no upperbound on the amount of flow that can be sent between hubs, we exclude interhub flow amounts among the optimum link flows and calculate the average (μ) and standard deviation (σ) of the remaining link flow values. Finally, assuming that the link flows are normally distributed, all of links' capacities ω_{ik} $i, k \in \mathcal{V}$ are set equal to $\omega_p = \mu + \zeta_p \times \sigma$ where ζ_p is the z-score corresponding to probability p with $p = \{0.70, 0.80, 0.90\}$. Original data sets do not contain also *hub capacities*, so we simply designate all Γ_k $k \in \mathcal{V}$ equal to a fraction (10%, 20% or 30%) of the total demand H where we denote these fractions as $\Gamma_{0.10}$, $\Gamma_{0.20}$ and $\Gamma_{0.30}$ respectively.

To demonstrate the usefulness of proposed MHCFP models, namely UMHCFP, MHCFP-1, MHCFP-2 or MHCFP-3, we introduce two indicators: the number of opened hubs (NH) and the percent of cost reduction (IMP). For a given setting of parameters, NH is obtained by solving one of proposed MHCFP models optimally. Let us denote the optimum objective function value of a MHCFP model as $v^{\star}_{MHCFP}$, and the value of (objective) function in Eq.(1) calculated by using the optimum solution of the associated MHCP model as v_{MHCP}. Then, IMP is calculated such that

$$\text{IMP} = 100 \times \frac{v_{MHCP} - v^{\star}_{MHCFP}}{v_{MHCP}}.$$

Table 1 summarizes execution times to reach optimum solutions for four models and two data sets. For each model and data set, several instances are formed by the combinations of model parameters and solved to optimality. For example, 27 different instances are formed by varying $\alpha = \{0.20, 0.50, 0.80\}$, $\Delta = \{0.60, 0.70, 0.80\}$ and $f = \{10,000, 20,000, 30,000\}$ for UMHCFP and CAB data set. All instances have been solved with GAMS 24.9.4 [47] via solver CPLEX 12.7.1 on dual Intel Xeon E5-2670 (2.6 GHz) processor and 32 GB of RAM workstation running Windows Server 2012 R2-64 bits. While the network size for TR data set is more than the triple compared to CAB data set, the solution times reported for it in average CPU times are much more smaller except for MHCFP-1 interestingly. In fact, it can observed that there is a great discrepancy in execution times for CAB data set by just looking to maximum and minimum CPU times in Table 1. This can be mainly attributed to varying hub opening costs. As hub opening costs are included in the original TR data set, we did not need to generate them and setup experiments with different cost sets accordingly. This is exactly what we have done in case of CAB data set since hub opening costs were not provided. Accordingly, the number of instances solved for CAB data set is three times more than TR data set as three different cost sets are considered.

Table 1: Solution time statistics (CPU times in seconds) for models and data sets

Statistics for CAB data set

	Num. of Instances	Mean Time	Max. Time	Min. Time
UMHCFP	27	31.53	111.33	3.73
MHCFP-1	81	65.02	196.09	6.69
MHCFP-2	81	82.08	739.16	4.19
MHCPF-3	243	107.04	1127.27	6.73

Statistics for TR data set

	Num. of Instances	Mean Time	Max. Time	Min. Time
UMHCFP	9	26.54	39.27	16.80
MHCFP-1	27	109.50	332.97	47.03
MHCFP-2	27	36.45	72.73	18.11
MHCFP-3	81	68.87	125.19	35.81

Results for UMHCFP given in Tables 2 and 6 indicate that both the optimum number of hubs NH and the cost improvement ratio IMP significantly increase as α and Δ decrease, and reach their highest values when these two model parameters attain their lowest values. We have to first note that the optimum solution of UMHCP requires only one hub to be opened for all combinations of α and Δ for both CAB and TR data sets. Then, this outcome can be easily interpreted as follows. For UMHCFP, the decrease in α favors more flows between hubs and this stimulates the opening of new hubs to benefit from this cost reduction. In fact, α is one of the most influential factor on NH for all models. Meanwhile, IMP is more affected by the decrease in the coverage ratio Δ. For two data sets, the added cost for increasing the number of hubs is overshadowed by the cost saving due to the increasing interhub flows, and the total cost is reduced accordingly. For CAB data set, increasing hub opening costs obviously decreases NH as the total hub opening cost becomes consequential compared to the total flow cost. In fact, the single objective of UMHCP is to set as few as possible nodes to hubs while routing all the flow demand without any regard for transportation costs. This approach is insufficient as it leads to design hub-and-spoke networks resulting in higher total cost.

Table 2: Summary of UMHCFP Results (CAB data set)

f_i	α	Δ					
		0.80		0.70		0.60	
		NH	IMP	NH	IMP	NH	IMP
10,000	0.80	6	31.67	6	31.43	5	42.62
	0.50	7	42.74	7	42.11	8	50.61
	0.20	9	56.34	9	54.89	10	60.42
20,000	0.80	4	24.79	4	24.38	4	36.28
	0.05	5	32.67	5	32.13	5	41.92
	0.20	5	44.45	5	43.59	5	50.09
30,000	0.08	3	20.08	3	19.89	3	32.17
	0.50	4	25.64	4	24.86	4	35.63
	0.20	5	36.44	5	35.60	5	43.39

The effects of imposing bounds on the flow passing thru network links and hubs are shown in Tables3-5 and 7-9. The results given in those tables are obtained by solving MHCFP-1, MHCFP-2 and MHCFP-3 respectiveley to optimality for several combinations of α, Δ, ω and Γ. Compared with the results of the unbounded model UMHCFP, it is clear that IMP rises further as bounds become more restrictive. In other words, it becomes more critical to consider transportation costs in hub covering problems and thus the *usefulness* of our proposed models increases.

Table 3: Summary of MHCFP-1 Results (CAB data set)

ω		$\omega_{0.90}$						$\omega_{0.80}$						$\omega_{0.70}$					
Δ		0.50		0.50		0.50		0.50		0.50		0.50		0.50		0.50		0.50	
f_i	α	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP
10,000	0.80	6	50	7	64	7	59	7	65	7	54	7	59	7	52	7	56	7	57
	0.50	7	54	7	66	8	61	8	65	7	55	8	61	9	56	8	56	8	58
	0.20	10	61	10	70	11	64	10	68	10	59	11	66	11	62	11	59	12	60
20,000	0.80	4	46	4	60	4	55	4	61	4	49	5	56	5	61	5	42	5	57
	0.50	5	46	5	57	5	54	6	59	6	47	6	56	6	60	7	44	6	57
	0.20	6	51	7	57	7	55	7	60	7	49	7	58	7	62	7	50	7	58
30,000	0.80	3	50	3	58	4	54	3	59	3	47	4	57	4	59	4	54	4	58
	0.50	4	48	4	57	4	52	4	55	4	43	4	53	4	57	5	49	5	55
	0.20	5	49	5	58	5	53	5	54	5	42	5	50	6	56	6	46	6	53

Table 4: Summary of MHCFP-2 Results (CAB data set)

Γ		$\Gamma_{0.30}$						$\Gamma_{0.20}$						$\Gamma_{0.10}$					
Δ		0.80		0.70		0.60		0.80		0.70		0.60		0.80		0.70		0.60	
f_i	α	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP
10,000	0.80	6	61	7	65	7	67	7	64	7	64	7	61	10	61	10	66	10	58
	0.50	7	62	7	67	8	68	8	66	8	66	8	62	12	63	12	66	12	58
	0.20	9	66	9	70	10	70	10	70	10	69	11	65	12	67	12	67	13	60
20,000	0.80	4	57	4	62	4	64	5	57	5	60	6	56	10	51	10	61	10	53
	0.50	6	56	6	61	5	63	6	57	6	60	6	56	10	52	10	59	10	51
	0.20	6	58	6	63	5	63	6	59	6	62	6	58	10	56	10	58	10	50
30,000	0.80	4	55	4	54	4	62	5	59	5	57	5	53	10	50	10	54	10	48
	0.50	4	53	4	51	4	60	5	58	5	55	6	52	10	50	10	52	10	46
	0.20	5	53	5	52	5	59	6	59	6	57	6	53	10	51	10	50	10	45

Table 5: Summary of MHCFP-3 Results (CAB data set)

Γ			$\Gamma_{0.30}$						$\Gamma_{0.20}$						$\Gamma_{0.10}$					
Δ			0.80		0.70		0.60		0.80		0.70		0.60		0.80		0.70		0.60	
ω	f_i	α	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP
$\omega_{0.90}$	10,000	0.80	6	71	7	70	7	61	7	56	7	59	7	57	10	68	10	68	10	59
		0.50	7	72	7	72	8	63	8	59	8	61	8	58	12	69	12	68	12	59
		0.20	10	74	10	76	11	66	10	65	10	65	11	62	12	71	12	70	13	61
	20,000	0.80	4	58	4	54	4	58	5	52	6	55	6	56	10	63	10	63	10	53
		0.50	6	58	6	53	6	57	6	53	6	55	6	55	10	62	10	62	10	52
		0.20	6	60	7	54	7	58	6	56	7	58	7	55	10	62	10	62	10	51
	30,000	0.80	4	66	4	51	4	61	5	49	5	52	6	51	10	60	10	55	10	50
		0.50	4	64	4	49	4	58	6	48	6	51	6	51	10	58	10	54	10	48
		0.20	5	63	5	50	5	57	6	51	6	53	6	51	10	57	10	53	10	47
$\omega_{0.80}$	10,000	0.80	7	64	7	65	7	67	7	67	7	60	7	63	10	67	10	65	10	56
		0.50	8	66	7	67	8	67	9	68	8	61	8	63	12	68	12	64	12	57
		0.20	10	70	10	71	11	68	10	71	10	64	11	66	12	70	12	66	13	60
	20,000	0.80	4	57	5	56	5	62	6	64	6	57	6	49	10	62	10	59	10	51
		0.50	6	56	6	56	6	61	6	63	7	56	6	47	10	61	10	58	10	50
		0.20	7	59	7	59	7	60	7	64	7	57	7	45	10	61	10	57	10	50
	30,000	0.80	4	52	4	63	4	60	5	51	5	48	6	53	10	55	10	54	10	50
		0.50	4	50	5	61	5	57	6	48	6	45	6	52	10	55	10	52	10	48
		0.20	6	51	5	62	5	55	6	46	6	45	6	52	10	56	10	51	10	47
$\omega_{0.70}$	10,000	0.80	7	56	7	58	7	64	7	68	7	54	7	62	10	67	10	64	10	67
		0.50	9	58	8	60	8	65	9	69	9	55	8	63	12	68	12	64	12	65
		0.20	11	64	11	64	12	68	11	71	11	59	12	65	12	71	12	67	13	64
	20,000	0.80	5	52	5	54	5	63	6	62	6	62	6	57	10	61	10	59	10	62
		0.50	6	51	7	53	6	61	7	61	7	61	6	56	10	61	10	59	10	58
		0.20	7	54	7	55	7	62	7	61	7	61	7	57	10	62	10	59	10	55
	30,000	0.80	4	56	4	58	4	59	5	52	5	47	6	46	10	57	10	49	10	53
		0.50	5	54	5	56	5	56	6	51	6	43	6	44	10	56	10	47	10	49
		0.02	6	55	6	57	6	55	6	52	6	41	6	42	10	57	10	47	10	47

Table 6: Summary of UMHCFP Results (TR data set)

α	Δ					
	0.75		0.65		0.55	
	NH	IMP	NH	IMP	NH	IMP
0.80	7	42.97	7	76.85	11	79.33
0.60	7	44.61	15	77.77	15	80.35
0.40	21	47.96	24	79.68	25	82.01

Table 7: Summary of MHCFP-1 Results (TR data set)

ω	α	Δ					
		0.75		0.65		0.55	
		NH	IMP	NH	IMP	NH	IMP
$\omega_{0.90}$	0.80	8	79.13	8	78.10	8	87.90
	0.60	10	76.18	10	75.15	10	85.44
	0.40	25	73.00	25	72.08	25	82.12
$\omega_{0.80}$	0.80	9	77.62	9	75.83	9	87.28
	0.60	13	74.74	13	73.00	13	84.80
	0.40	26	71.75	26	70.22	26	81.48
$\omega_{0.70}$	0.80	9	85.46	9	85.43	9	82.40
	0.60	13	83.25	13	83.22	13	80.16
	0.40	28	80.53	28	80.51	28	77.65

Table 8: Summary of MHCFP-2 Results (TR data set)

Γ	α	Δ					
		0.75		0.65		0.55	
		NH	IMP	NH	IMP	NH	IMP
$\Gamma_{0.30}$	0.80	9	92.11	12	93.00	11	90.95
	0.60	13	90.63	16	91.82	17	89.83
	0.40	24	88.61	26	90.40	28	88.62
$\Gamma_{0.20}$	0.80	12	95.74	13	93.98	14	93.44
	0.60	14	94.82	18	92.78	18	92.20
	0.40	26	93.41	28	91.13	29	90.49
$\Gamma_{0.10}$	0.80	18	97.21	17	96.81	17	94.53
	0.60	23	96.67	25	96.21	24	93.59
	0.40	33	95.87	33	95.37	34	92.31

Table 9: Summary of MHCFP-3 Results (TR data set)

Γ		$\Gamma_{0.30}$						$\Gamma_{0.20}$						$\Gamma_{0.10}$					
Δ		0.75		0.65		0.55		0.75		0.65		0.55		0.75		0.65		0.55	
ω	α	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP	NH	IMP
$\omega_{0.90}$	0.80	10	93	12	92	14	92	13	96	13	93	14	92	18	97	17	97	17	93
	0.60	13	91	17	91	18	90	15	95	18	92	18	90	25	96	28	97	25	92
	0.40	26	89	28	90	30	89	27	94	29	91	30	88	34	95	34	96	34	91
$\omega_{0.80}$	0.80	11	95	12	95	14	89	13	96	13	94	14	90	19	96	18	96	18	94
	0.60	14	94	17	94	18	88	16	95	19	93	18	88	25	96	28	96	24	92
	0.40	26	92	28	93	30	87	27	94	29	91	30	87	34	95	34	95	34	91
$\omega_{0.70}$	0.80	13	95	13	94	14	89	14	96	14	94	14	89	20	96	19	97	19	93
	0.60	16	94	19	93	18	88	17	95	20	93	19	87	25	96	28	96	25	92
	0.40	27	92	28	91	30	87	27	93	30	91	30	85	35	95	34	95	34	91

Link capacities have less consequence for IMP compared to hub capacities. This can be observed when results in Tables 3-4 for CAB data set and Tables 7-8 for TR data set are contrasted. We can also derive the conclusion that hub capacities are a dominant factor for IMP according to the results presented in Tables 4-5 and Tables 8-9. Moreover, IMP is seriously decreased when the cost of opening hubs increases and the capacities of hubs decrease simultaneously according to Table 4-5. From Tables 4-5, it is apparent that the rise of hub opening costs decreases NH, as expected. The other most influential factor on the value of NH is the capacity of hubs, while link capacities have a marginal effect on it. The decrease in Δ has no or slightly increasing consequence on NH for all models. As α is reduced, NH can dramatically increase but IMP is very slightly effected or worsened by this decrease. For some particular combination of parameters values, proposed models are solved optimally and the real locations of opened hubs are plotted on the maps in Figures 1 and 2. The black dots represent hubs that are common for all instances while white dots correspond to additional hubs specific to the instance. It is apparent that high accessibility and total demand flowing in/out of a node are determinants for hub locations. Moreover, cost savings resulting from interhub transfers encourage the opening of several hubs.

5. Conclusion

In MHCFP, the aim is to find the optimal design for hub-and-spoke networks while considering hub opening and demand routing costs and coverage constraints. Flow demand associated with a specific origin-destination node pair must be routed by visiting at least one hub node. In this study, it is assumed that a hub covers a nonhub node if the distance between these two is less than a predefined value, while there is no limit on the distance between hubs.

Our formulations of MHCFP are linear programming problems with continuous and integer variables. The results obtained from our computational experiments reveal that the number of hubs in the network decreases depending on the increase in hub opening costs, as expected. In fact, the increase in hub opening costs also affect how hub nodes are located such that hubs hubs have a tendency to be placed exactly at or near at nodes having high demands.

We can also derive from our numerical experiments that not taking into account flow costs can be consequential in the network design. For example, the total cost of designing a network and making it operational depending on the solution of UMHCFP could be as low as 80% below that based on MHCP. For other models, this percent can be even larger. This is attributed to the interhub flow cost discount factor which models the cost reduction due to the consolidated shipments.

There are several future research directions for MHCFP.Exact or heuristic methods to solve larger versions of MHCFP can be the initial point of a stream of work. In another line of research, MHCFP can be considered in a stochastic environment, where all costs and flow demands and even the network structure can depends on scenarios and the objective can be to design a resilient network to changing conditions.

Acknowledgements

This work has been supported by the Scientific Research Projects Commission of Galatasaray University under grant number 17.402.007.

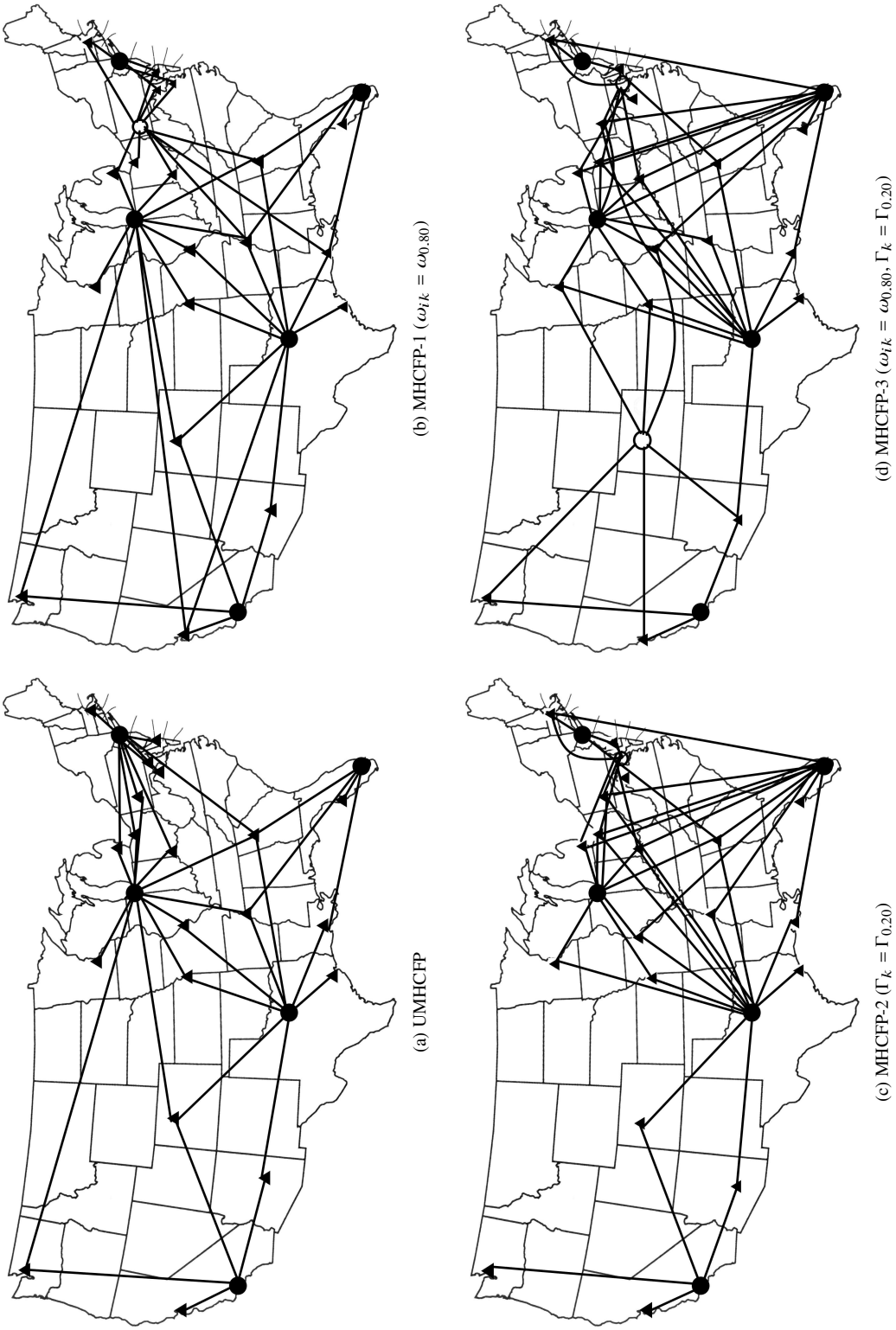


Figure 1: Optimum solution for models with CAB data set ($\alpha = 0.50, \Delta = 0.70$ and $f_i = 20,000$)

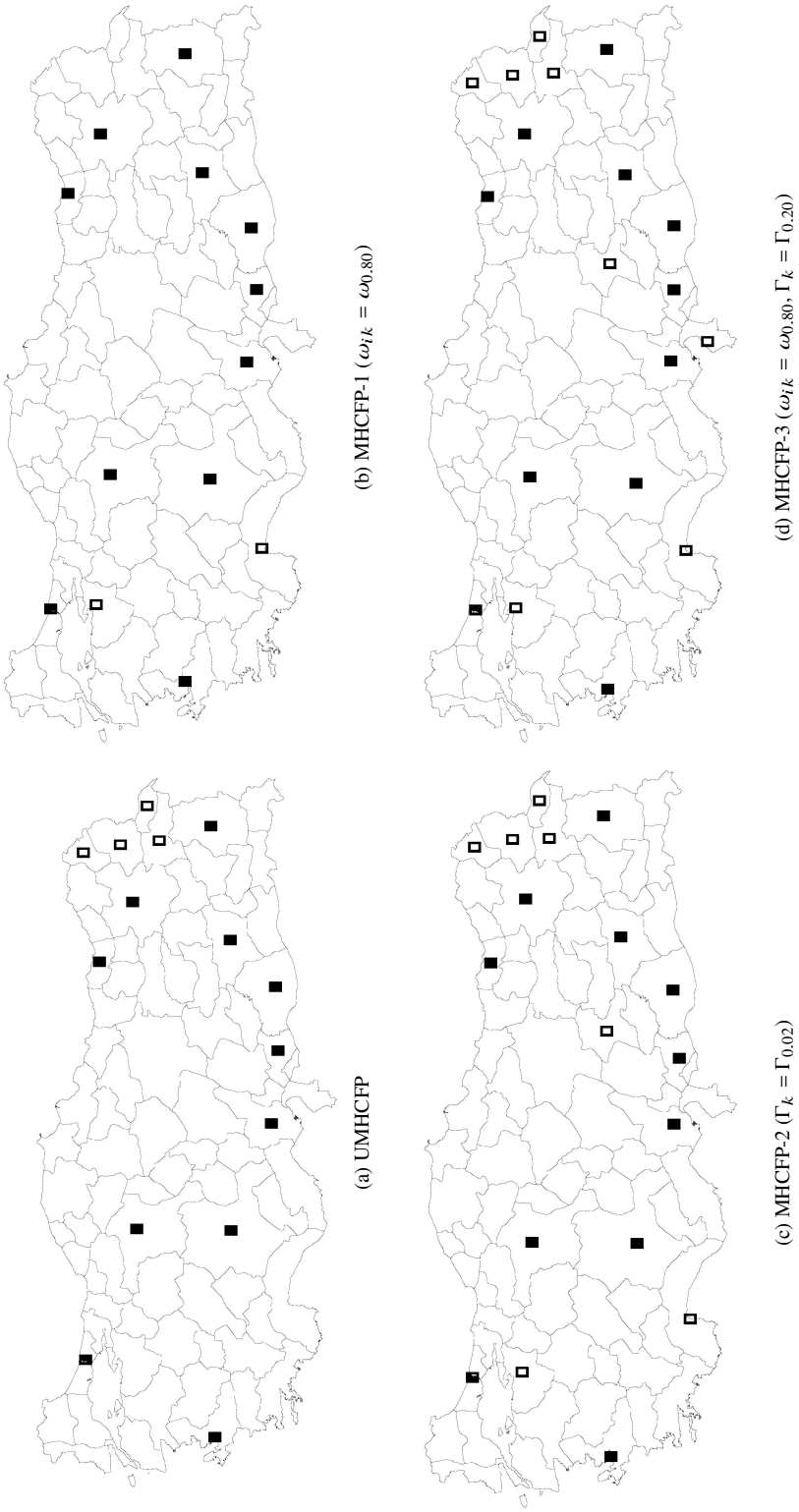


Figure 2: Optimum placement of hubs for models with TR data set ($\alpha = 0.60$ and $\Delta = 0.65$)

References

[1] Campbell J.F., O'Kelly M.E., Twenty-five years of hub location research, Transportation Science, 46, (2012), pp. 153-169.

[2] Campbell J.F., Integer programming formulations of discrete hub location problems, European Journal of Operational Research, 72, (1994), pp. 387-405.

[3] Farahani R.Z., Hekmatfar M., Arabani A.B., Nikbakhsh E., Hub location problems: A review of models, classification, solution techniques, and applications, Computers & Industrial Engineering, 64, (2013), pp. 1096-1109.

[4] Hakimi S., Optimum locations of switching centers and the absolute centers and medians of a graph, Operations Research, 12, (1964), pp. 450-459.

[5] Toh R., Higgins R., The impact of hub and spoke network centralization and route monopoly on domestic airline profitability, Transportation Journal, 24, (1985), pp. 16-27.

[6] O'Kelly M.E., The location of interacting hub facilities, Transportation Science, 20, (1986b), pp. 92-106.

[7] O'Kelly M.E., Activity levels at hub facilities in interacting networks, Geographical Analysis, 18, (1986), pp. 343-356.

[8] Bryan D., O'Kelly M., Hub-and-spoke networks in air transportation: An analytical review. Journal of Regional Science, 39, (1999), pp. 275-295.

[9] Campbell J., Ernst A., Krishnamoorthy, M., "Hub Location Problem" in Facility Location:Applications and Theory, Springer-Verlag, Berlin, 2002, pp. 373-407.

[10] Alumur S., Kara B., Network hub location problems: The state of the art, European Journal of Operational Research, 190, (2008), pp. 1-21.

[11] Contreras I., "Location Science" in Hub Location Problems, Springer International Pub- lishing, Switzerland, 2015, pp. 311-344.

[12] Yetis Kara B., Tansel B., The single-assignment hub covering problem: Models and linearizations, The Journal of the Operational Research Society, 54, (2003), pp. 59-64.

[13] Ernst A.T., Jiang H., Krishnamoorthy M., Baatar D., Reformulations and computational results for uncapacitated single and multiple allocation hub covering problems, in: Unpublished Report, CSIRO Mathematical and Information Sciences, 2005.

[14] Wagner B., Model formulations for hub covering problems, The Journal of the Operational Research Society, 59, (2008), pp. 932-938.

[15] Ernst A.T., Krishnamoorthy M., Efficient algorithms for the uncapacitated single allocation p-hub median problem, Location Science, 4, (1996), pp. 139-154.

[16] Weng K., Wang Y., "Evolutionary algorithms for multiple allocation hub set covering problem", in: 2008 IEEE International Conference on Networking, Sensing and Control, 2008. pp. 408-411. Conference Paper in Print Proceedings

[17] Qu B., Weng K., Path relinking approach for multiple allocation hub maximal covering problem, Computers & Mathematics with Applications, 57, (2009), pp. 1890-1894.

[18] Alumur Alev S., Yetis Kara B., A hub covering network design problem for cargo applications in Turkey, The Journal of the Operational Research Society, 60, (2009), pp. 1349-1359.

[19] Calik H., Alumur Alev S., Yetis Kara B., Karasan O.E., A tabu-search based heuristic for the hub covering problem over incomplete hub networks, Computers & Operations Research, 36, (2009), pp. 3088-3096.

[20] Lowe T.J., Sim T., The hub covering flow problem, The Journal of the Operational Research Society, 64, (2013), pp. 973-981.

[21] Alumur S., Kara B., Karasan O., Multimodal hub location and hub network design, Omega, 40, (2012), pp. 927-939.

[22] Campbell J.F., Location and allocation for distribution systems with transshipments and transportion economies of scale, Annals of Operations Research 40, (1992), pp. 77-99.

[23] Aykin T., Lagrangian relaxation based approaches to capacitated hub-and-spoke network design problem, European Journal of Operational Research, 79, (1994), pp.501-523.

[24] Bryan D., Extensions to the hub location problem: Formulations and numerical examples, Geographical Analysis, 30, (1998), pp. 315-330.

[25] O'Kelly M., Bryan D.,Hub location with flow economies of scale, Transportation Research Part B: Methodological, 32, (1998), pp. 605-616.

[26] Ernst A., Krishnamoorthy M., Solution algorithms for the capacitated single allocation hub location problem, Annals of Operations Research, 86, (1999), pp. 141-159.

[27] Ebery J., Krishnamoorthy M., Ernst A., Boland N., The capacitated multiple allocation hub location problem: Formulations and algorithms, European Journal of Operational Research, 120, (2000), pp. 614-631.

[28] Marín, A., Formulating and solving splittable capacitated multiple allocation hub location problems, Computers & Operations Research, 32, (2005), pp. 3093-3109.

[29] Sasaki M., Fukushima M., On the hub-and-spoke model with arc capacity constraints, Journal of the Operations Research Society of Japan, 46, (2003), pp. 409-428.

[30] Carello G., Della Croce F., Ghirardi M., Tadei R., Solving the hub location problem in telecommunication network design: A local search approach, Networks, 44, (2004), pp. 94-105.

[31] Yetis Kara B., Tansel B., The single-assignment hub covering problem: Models and linearizations, The Journal of the Operational Research Society, 54, (2003), pp. 59-64.

[32] Yaman H.,Star p-hub median problem with modular arc capacities, Computers & Operations Research, 35, (2008), pp. 3009-3019.

[33] Rodríguez-Martín I., Salazar-González J.J., Solving a capacitated hub location problem, European Journal of Operational Research, 184, (2008), pp. 468-479.

[34] da Graça Costa M., Captivo M.E., Clímaco J., Capacitated single allocation hub location problem-a bi-criteria approach, Computers & Operations Research, 35, (2008), pp. 3671-3695.

[35] Mohammadi M., Tavakkoli-Moghaddam R., Rostamib H., A multi-objective imperialist competitive algorithm for a capacitated hub covering location problem, International Journal of Industrial Engineering Computations, 2, (2011), pp. 671-688.

[36] Contreras I., Cordeau J.F., Laporte G., Exact solution of large-scale hub location problems with multiple capacity levels, Transportation Science, 46, (2012), pp. 439-459.

[37] Sedehzadeh S., Tavakkoli-Moghaddam R., Jolai F., A new multi-mode and multi-product hub covering problem: A priority m/m/c queue approach, International Journal of Industrial Mathematics, 7, (2015), pp. 139-148.

[38] Karimia H., Bashiri M., Nickel S., Capacitated single allocation p-hub covering problem in multi-modal network using tabu search, International Journal of Engineering, 29, (2016), pp. 797-808.

[39] Merakli M., Yaman H., A capacitated hub location problem under hose demand uncertainty, Computers and Operations Research, 88, (2017), pp. 58-70.

[40] Hoff A., Peiró J., Ángel C., Martí R., Heuristics for the capacitated modular hub location problem, Computers & Operations Research, 86, (2017), pp. 94-109.

[41] Demir I., Ergin F. C., Kiraz, B., A New Model for the Multi-Objective Multiple Allocation Hub Network Design and Routing Problem, IEEE Access, 7, (2019), pp. 90678-90689.

[42] Danach K., Gelareh S., Neamatian Monemi R., The capacitated single-allocation p-hub location routing problem: a Lagrangian relaxation and a hyper-heuristic approach, EURO Journal on Transportation and Logistics, 8, (2019), pp.597–631.

[43] Taherkhani G., Alumur S.A., Hosseini M., Benders Decomposition for the Profit Maximizing Capacitated Hub Location Problem with Multiple Demand Classes, Transportation Science, 54:6, (2020), pp 1446-1470.

[44] Butun C., Petrovic S., Muyldermans L., The Capacitated Directed Cycle Hub Location and Routing Problem Under Congestion, European Journal of Operational Research, (2020), *In Press*.

[45] Beasley J., Or-library. URL: http://people.brunel.ac.uk/mastjjb/jeb/info.html, 1990.

[46] Yetis Kara B., Turkish hub data set. URL:http://www.bilkent.edu.tr/ bkara/dataset.php, 2017.

[47] McCarl B.A., Meeraus A., van der Eijk P., Bussieck M., Dirkse S., Nelissen F., McCarl Expanded GAMS User Guide, GAMS Release 24.6. GAMS Development Corporation. Washington, DC, USA.

MANAS Journal of Engineering
ISSN 1694-7398 | e-ISSN 1694-7398

Volume 9, Issue 1, (2021) Pages 85-92
https://doi.org/10.51354/mjen.852312

DEEP LEARNING METHOD FOR HANDWRITING RECOGNITION

Ayşe Ayvacı[1,*], Abdullah Erdal Tümer [2,3]

[1]*Necmettin Erbakan University, Institute of Natural Sciences, Department of Computer Engineering, Konya/Turkey*
[2]*Necmettin Erbakan University, Faculty of engineering and architecture, Department of Computer Engineering, 42195, Konya/Turkey*
[3] *Kyrgyz – Turkish Manas University, Bishkek, Kyrgyzstan*

ABSTRACT

The advancement of technology nowadays resulted into documents, such as forms and petitions, being filled out in computer and digital environment. Yet in some cases, documents are still preserved in traditional style, on print. Due to its distinct proportions, however, its storage, sharing and filing has become a complication. The relocation of these written documents to digital environment is therefore of great significance. In this view, this study aims to explore methodologies of digitizing handwritten documents. In this study, the documents converted to image format were pre-processed using image processing methods. These operations include dividing lines of the document into image format, dividing into words which then divided into characters, and finally, a classification operation on the characters. As classification phase, one of the deep learning methods is the Convolution Neural Network method is used in image recognition. The model was trained using the EMNIST dataset, and in the character, dataset created from the documents at hand. The dataset created had a success rate of 87.81%. Characters classified as finishers are sequentially combined and the document is transferred to the computer afterwards.

ARTICLE INFO

Research article

Received: 3.01.2021
Accepted: 28.06.2021

Keywords:
Character recognition, convolutional neural network, deep learning, handwriting recognition, image processing

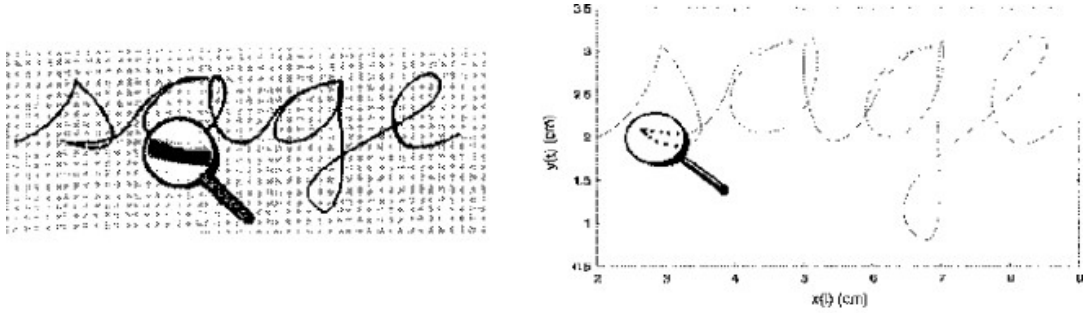
[*]Corresponding author

1. Introduction

Writing process is a complex activity consisting of cognitive, kinesthetic, and perceptual-motor components [1]. Writing is the written form of the words, feelings, and thoughts in the language. Handwriting is unique for each individual. It is a tool developed long ago to expand human knowledge and facilitate communication [2].

As technology advances, numerous documents are filled in the digital environment. Yet there are still transactions which are still being done on paper. The task of transferring these written documents to computer can be done by one or more people. In this circumstance, handwriting recognition systems have been developed to transfer written documents automatically to computer. Handwriting recognition is the process of defining and making sense out of letters, numbers, and symbols that are written by hand on papers, tablets, and smartphones, using computer systems.

There are two methods available in handwriting recognition systems - offline method and online method. In the offline method (Figure 1a), the writing is used as a picture. It contains the spatial brightness of the image [2]. On the other hand, the online method (Figure 1b) contains spatial-timer values of characters.



(a) (b)

Figure 1. *(a) Offline handwriting (b) Online handwriting (from [2])*

The character recognition rates of online methods are more successful than offline methods. Offline methods have 95%, 85%, and 78% recognition for 10, 100, 1000 words, and above respectively [2, 3]. In online methods, the success rate decreases as the number of words requested to be recognized increases. An accuracy rate of 80% is achieved at 21,000 words and above [2, 4].

The convenience of character recognition technology can be utilized in several business areas:

- Checks sent to banks are automatically recognized and the necessary account transactions are performed electronically;
- Bringing doctors' handwriting to an understandable format. In a study conducted on this subject, 500 prescription samples taken from doctors were identified handwriting recognition and keywords from the doctor's handwriting. With these specified words, it is understood what the desired drug is in the prescription report written by the doctor. Hidden Markov Model was used as the model in the project. Multilayer perceptron (MLP) was also used to increase performance [5]
- Recognition of addresses on letters and automatic parsing by postal code. In the project carried out by CEDAR for USPS (United States Postal Service), the aim is to correctly separate envelopes according to the address where letters will automatically go in post offices [6].

Handwriting characters are more difficult to distinguish than optical characters. Consequently, accuracy rates in current studies are less than 90%. Some of difficulties in handwriting recognition include:

- Having too many different fonts, numbers, and symbols
- Difference in writing style
- Cursive or interconnected writing of letters
- Pencil or paper used
- Difference of letters' shapes and sizes depending on individual's writing style and speed.

Despite these challenges, researchers have conducted many studies for more successful character recognition. For instance, Fanany [7] used a workflow algorithm and a machine learning model to recognize handwriting characters in a form document. Convolutional Neural Network (CNN) is used for feature extraction of characters and Support Vector Machines (SVM) is used for classification. The study presents 83.37% accuracy in ten different test form documents.

Mahapatra, Choudhury, Karsh [8], in another study, used K-Nearest Neighbor (KNN), SVM-based classifiers such as CNN, hybrid KNN-SVM and v-SVM on different handwriting datasets such as EMNIST, Devanagari handwriting character and Kannada-MNIST. The successes of the developed models have been compared. As a result, EMNIST, Devanagari Handwritten Character, and Kannada-MNIST achieved a success rate of 89.02%, 86.67%, and 95.3%, respectively.

Saha and Jaiswal [9] also tested the EMNIST and UCI Devanagari datasets with the CNN classification model. In this model, success rate of EMNIST dataset was 79.3% while it is 93% for the UCI Devanagari dataset.

Cohen, Afshar, Tapson, and Van Schaik [10], evaluated the EMNIST dataset with an OPIUM- based classifier. The model achieved success rate of 78.02% in the EMNIST balanced dataset.

Baykal, Aktaş and Yildiz [11], created an active offline verification system by selecting the signature characteristic, which is one of the biometric characteristics.

In the study, recognition of handwriting characters was performed offline. The current study, unlike the other studies in the literature, did not use a separate algorithm to extract features on images. Instead of a separate algorithm, a single

method was used to both extract and classify character traits using the CNN method. The CNN method was used for classification and feature extraction. The Balanced part of the EMNIST [12] was used to train the model. The dataset used in the study was drawn from the Kaggle data store (from https://www.kaggle.com/crawford/emnist). EMNIST Balanced dataset consists of 131,600 characters and 47 classes. The model trained with the EMNIST Balanced dataset, was also tested on an independently generated dataset.

The study consists of two stages: pre-processing and classification. The following sections consist of pre-processing, classification, experimental study and result consecutively.

2. Pre-Processing

Having characters of the same size and shape affects the success rate in recognition. Therefore, preprocessing was applied to the letters and numbers in the data and test set. Pre-machining is done in two ways:

i. *Pre-processing:* With this process, the text on the paper is separated into characters. The separated characters are then tested correctly. For this, following operations are performed:

- Letters on paper are loaded into the computer in picture format. The installation process is done with the help of a scanner or camera. An original image is given in Figure 2.

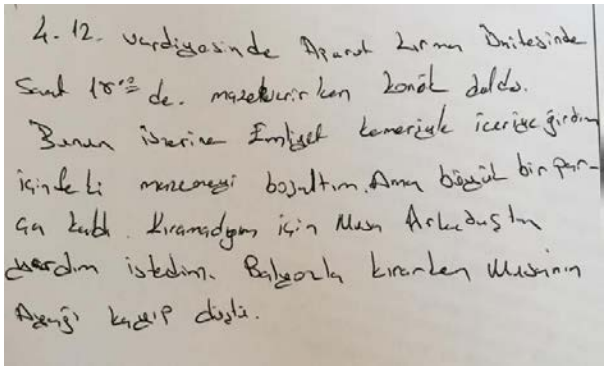


Figure 2. *Original image*

- The image is first converted into a gray-level image. Figure 3 is an example of an image that is converted into grayscale.

Figure 3. *Grayscale image*

The grayscale image is converted into a black-and-white image by applying the threshold. Afterwards, the image looks as in Figure 4.

Figure 4. *Black-and-white image*

Through the contour extraction process in the black-and-white image, the lines are determined by drawing a curve that connects all continuous points that have the same color and density along the borderlines. Fig. 5 also shows an image divided into lines.

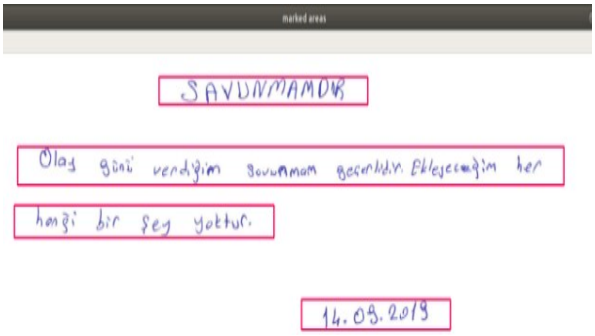


Figure 5. *Image divided into rows*

Application of re-contour extraction process of lines resulted into division of document into words. Figure 6 includes an image divided into sentences.

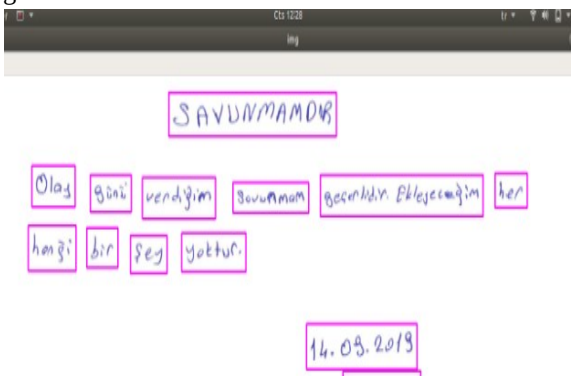


Figure 6. *Image divided into words*

- The reserved words over the contour extraction process are completed. Throughout this process, the boundaries of the character within the words are determined
- Over characters whose boundaries are set, the ROI (area of interest) operation is performed. This process is the process of obtaining a new image that has only character

boundaries according to the pixel values of the corresponding region. The resulting new images are 28x28 pixels. The new images obtained are characters in black and white format, as in Figure 7.

Figure 7. *Characters allocated to test*

ii. *Normalization*: After pre-processing, normalization is applied to the EMNIST dataset. Through this process, the images are brought to the same size and type.

3. Classification

In this study, a well-known deep learning algorithm CNN is used. CNN algorithms are applied in many different areas such as natural language processing (NLP) [13, 14], biomedical processing [15, 16], especially image [17, 18] and audio processing [19, 20]. It is a deep learning algorithm that has the best classification success, especially in the field of image processing. CNN processes the image in various layers. The layer structure of the CNN model is shown in Figure 8. These layers [21] are:

- Convolutional layer- Which is used to determine properties;
- Non-Linearity layer- Which introduces non-linearity to the system;
- Pooling (Down-sampling) layer - Which reduces the number of weights and controls compliance;
- Flattening layer – Which prepares data for the classical neural networks; and
- Fully-Connected layer – Which is the standard neural network used in classification.

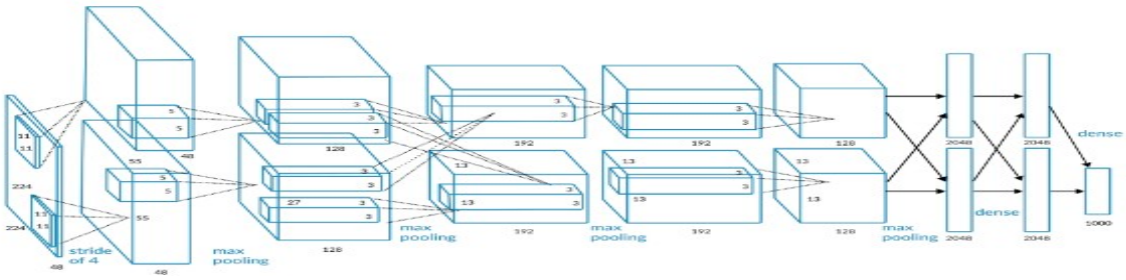


Figure 8. *CNN's layer structure (from www.missinglink.ai)*

The Convolution layer is used to detect the properties of the picture. This layer uses a low or high- level filter to extract properties. These filters are usually matrices of odd numbers. By moving the filter over the image and using matrix multiplication, its properties are detected. In Figure 9, the example of convolution filter and its application are shown. After filters, zero values are added in order for the image to not lose its original size.

The Non-linearity layer comes after convolution layers. Since all layers can be linear functions, the

Neural Network acts as a single perceptron, which means that the results are calculated linearly. In this layer, one of the activation functions is used. The best results about the speed of neural network training are the Rectifier (Relu) function, therefore this function was used. The Relu function is shown in Equation 1 below:

$$RELU\ Func. = -f(x) = max \tag{1}$$

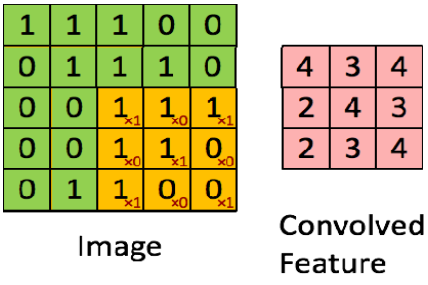


Figure 9. *Convolution filter (form www.medium.com/@tuncerergin)*

The Pooling layer is inserted between successive convolutional layers. This layer is designed to reduce the spatial size of the representation and the number of parameters and calculations within the network. The most used is max-pooling. In max-pooling, a filter is created first. This filter moves around the image and takes the largest

number in the area it occupies. Figure 10 shows an example of a max-pooling process.

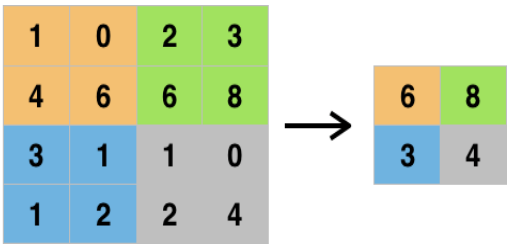


Figure 10. *Max-pooling process (from www.medium.com/@tuncerergin)*

The Fatting layer, the last layer, prepares the input data for the Fully Connected Layer. Neural networks take inputs from a one-dimensional array. The data in a network is in a state ofmatrices, that was converted into a one-dimensional array, in convolutional and pool layers. Figure 11 shows an example of the last layer translated into a one-dimensional array.

The Fully Connected Layers take data from the flattening layer and performs learning through the Neural network. During the training, the dropout method (shown in Figure 12) is used to prevent underfitting, unnecessary memorization, and overfitting proms. Decreasing the number of connections within the network improves training performance.

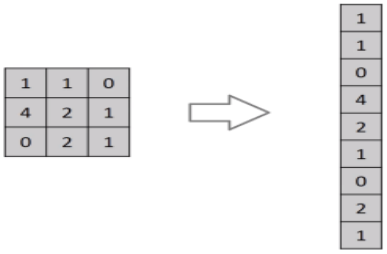


Figure 11. *Flatting layer (from www.medium.com/@tuncerergin)*

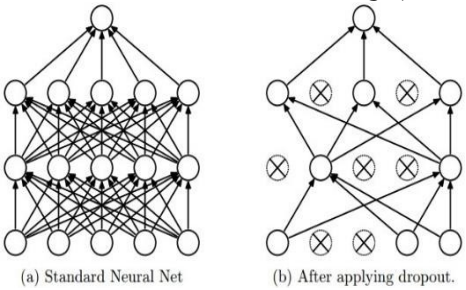


Figure 12. *Dropout layer (from www.medium.com/@tuncerergin)*

4. Experimental Work

The study, a computer with a working i5 processor with 8 GB of RAM is also written in Python

2.6. The OpenCV library was used for pre-processing. The Keras Library was used in deep learning. The study consisted of two phases. The first stage is the acquisition of characters, the creation of a CNN algorithm to classify the characters that we get in the second stage, and the training and testing of characters.

In the first stage, the letters written on paper are first thrown into the computer environment in PDF format. The document in PDF format is converted to JPG format, which is the image format. Images 2 using OpenCV library. the pre-processing steps described in the section were performed, and thus the articles written in the image were separated the first line by line, the words for each line were separated sequentially, each separated word was divided into characters, and the separation of characters in the image-shaped text was performed. The character is normalized and the character images are translated into black-and-white (binary system).

In the second stage, model training was first performed to classify the characters. The Balanced part of the EMNIST dataset was used as the dataset for the training of the model, and Convolutional neural networks were used as the deep learning architecture. To create the model; after creating layers of the CNN algorithm, the model was trained on the EMNIST dataset 2 times according to different epochs and batch size values. The trained models were tested using the detected letters in the documents using the image processing method. Figure 13 shows block diagram of the study.

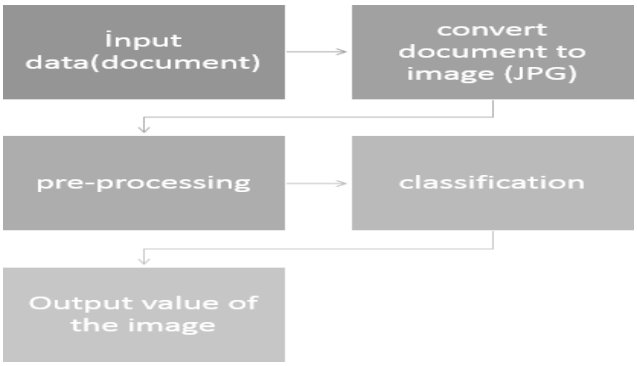


Figure 13. *Block diagram of the study*

This study used handwritten letters and numbers consisting of 112,799 training sets and 18,799 test sets in 28x28 pixels sizes found in the EMNIST Balanced dataset. The 10% of the training set is reserved for validation. In the training of the developed model, 4 CONV2D layers, 4 Max-Pooling layers, 1 Flatten layer, 1 Dropout layer and 2 Dense layers were used. Layers of the model created are shown in Figure 14 and Figure 15.

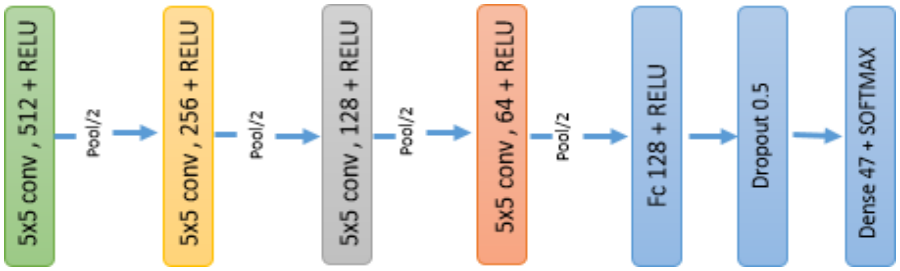


Figure 14. *CNN's block diagram used in the study*

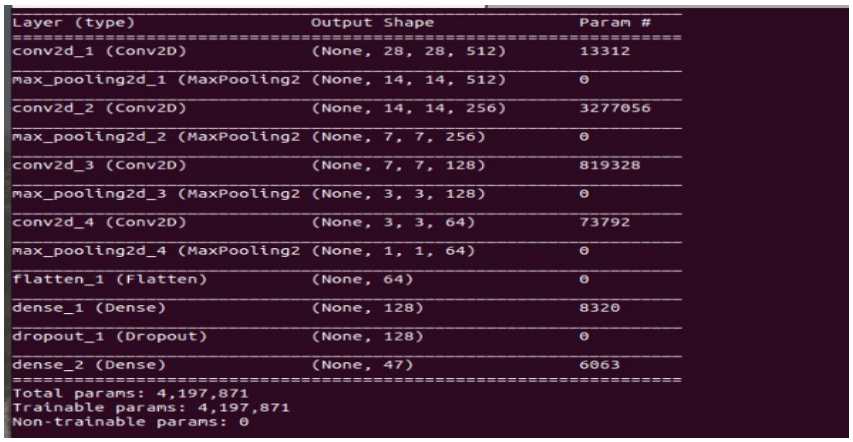


Figure 15. *Description of the layers of the model used in the study.*

The Model was trained twice, and the trainings lasted about 20 hours. The first training was performed with 5 epoch and 512 batch size, and the second training was performed with 10 epoch and 1024 batch size. Result of these two pieces of training showed that the second training with a rate of 87.81% success is the best score recorded. The values of the two pieces of training are also given in Table 1.

According to the model's 10 epoch and batch size 1024, its graphics are shown in Figure 16 and Figure 17.

Table 1. *Study results*

Epochs/ batch_size	Acc	Loss	Val_acc	Val_loss	Test_acc	Test_loss
5/512	%84.64	%37	%85.30	%35	%86	%36
10/1024	%87.81	%34	%87.61	%34	%87	%35

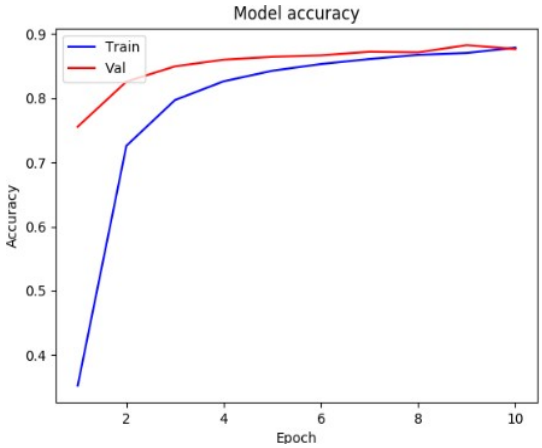


Figure 16. *Success graph of the model.*

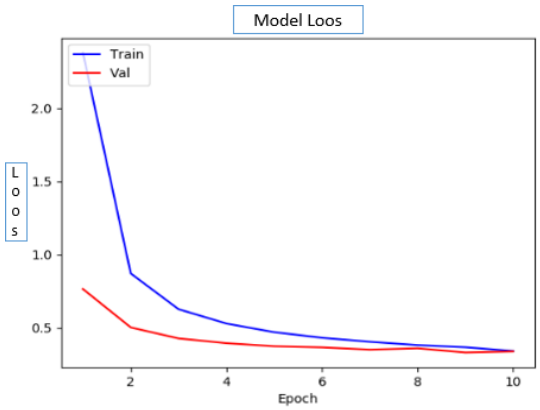


Figure 17. *Error graph of the model*

In this study, the independently created datasets were tested separately in both two trained models. The data tested has values of 28x28 pixels. Figure 18 shows the results of the characters that are obtained from handwritten documents in trained models.

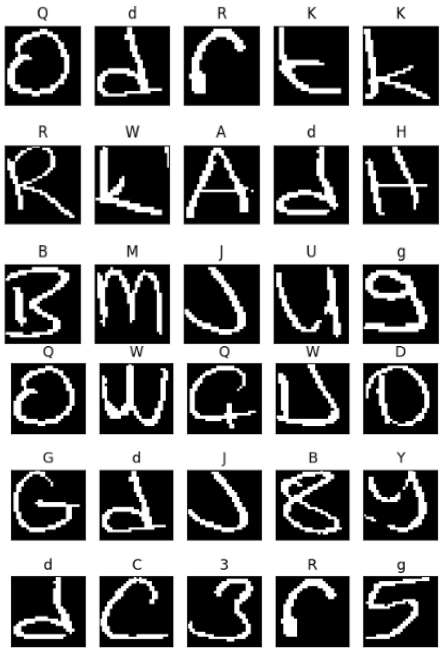


Figure 18. *The outputs of the characters tested in the model.*

5. Conclusions

This paper primarily discussed the recognition of characters from handwritten documents to computer environment. By means of encoding, these characters were recorded into computer media. For this purpose, the CNN model has been created throughout the operation. Pre-processing, slicing, and character recognition were then performed altogether as part of the procedure. At the end of the process, the obtained characters - 10 epoch and 1024 batch size trained models – both achieved a success rate of 87.81%.

The success rates of studies conducted using the Balanced part of the EMNIST dataset and different classifier algorithms are given in Table 2. Looking at Table 2, the method we used was more successful than other methods.

Table 2. *Success rates on the EMNIST dataset for CNN, Linear Classifier, and OPIUM- based Classifier*

Algorithm Name	**Success Rate**
OPIUM- based	78.02%
Linear Classifier	50.93%
CNN	87.81%

In reference to data in Table 1, it is foreseen that a more successful model can be obtained by increasing the values used in epoch and batch size. For this reason, future studies and its corresponding tests and trainings is recommended to carry out increase on the number of layers in order to improve the success rate of handwriting recognition, and the algorithm as well.

References

[1] B. Engel-Yeger, L. Nagauker-Yanuv, and S. Rosenblum, “Handwriting performance, self-reports, and perceived self-efficacy among children with dysgraphia,” *American Journal of Occupational*

Therapy, vol. 63, no. 2, pp. 182-192, 2009.

[2] R. Plamondon, and S. N. Srihari, "Online and off-line handwriting recognition: a comprehensive survey," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, no. 1, pp. 63-84, 2000.

[3] S. N. Srihari, "High-performance reading machines," *Proceedings of the IEEE*, vol. 80, no. 7, pp. 1120-1132, 1992.

[4] G. Seni, R. K. Srihari, and N. Nasrabadi, "Large vocabulary recognition of on-line handwritten cursive words," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 18, no. 7, pp. 757-762, 1996.

[5] P. P. Roy, A. K. Bhunia, A. Das *et al.*, "Keyword spotting in doctor's handwriting on medical prescriptions," *Expert Systems with Applications*, vol. 76, pp. 113-128, 2017.

[6] S. N. Srihari, S.-H. Cha, H. Arora *et al.*, "Handwriting identification: Research to study validity of individuality of handwriting and develop computer-assisted procedures for comparing handwriting," *Technical Report CEDAR-TR-01-1*, 2001.

[7] Fanany, M. I. Handwriting recognition on form document using convolutional neural network and support vector machines (CNN-SVM). In 2017 5th international conference on information and communication technology (ICoIC7) (pp. 1-6). IEEE. (2017, May).

[8] D. Mahapatra, C. Choudhury, and R. K. Karsh, " Mahapatra, D., Choudhury, C., & Karsh, R. K. Handwritten character recognition using knn and svm based classifier over feature vector from autoencoder. In International Conference on Machine Learning, Image Processing, Network Security and Data Sciences (pp. 304-317). Springer, Singapore. 2020, July.

[9] P. Saha, and A. Jaiswal, "Handwriting Recognition Using Active Contour," *Artificial Intelligence and Evolutionary Computations in Engineering Systems*, pp. 505-514: Springer, 2020.

[10] S. A. Gregory Cohen, Jonathan Tapson, and Andre van Schaik, "EMNIST: an extension of MNIST to handwritten letters," 2017.

[11] Baykal, B., Aktaş, T. Ö., & Yildiz, O. "Makine Öğrenmesi Yöntemlerİ İle Otomatik Çevrimdışı İmza Tanıma ve Doğrulama Sistemi". In 2017 International Artificial Intelligence and Data Processing Symposium (IDAP) (pp. 1-5). IEEE. 2017, September.

[12] NIST, "The EMNIST Dataset," 2017, April.

[13] Sun, B., Yang, L., Zhang, W., Dong, P., Young, C., Dong, J., & Lin, M. Demonstration of applications in computer vision and nlp on ultra power-efficient cnn domain specific accelerator with 9.3 tops/watt. In 2019 IEEE International Conference on Multimedia & Expo Workshops (ICMEW) (pp. 611-611). IEEE. (2019, July).

[14] Akhtyamova, L., Ignatov, A., & Cardiff, J. A Large-scale CNN ensemble for medication safety analysis. In International Conference on Applications of Natural Language to Information Systems (pp. 247-253). Springer, Cham. 2017, June.

[15] M. Cho, J. Ha, C. Park *et al.*, "Combinatorial feature embedding based on CNN and LSTM for biomedical named entity recognition," *Journal of Biomedical Informatics*, vol. 103, pp. 103381, 2020.

[16] Momeni, A., Thibault, M., & Gevaert, O. Dropout-enabled ensemble learning for multi-scale biomedical data. In International MICCAI Brainlesion Workshop (pp. 407-415). Springer, Cham. 2018, September.

[17] M. Amin-Naji, A. Aghagolzadeh, and M. Ezoji, "Ensemble of CNN for multi-focus image fusion,"*Information fusion*, vol. 51, pp. 201-214, 2019.

[18] C. Tian, Y. Xu, and W. Zuo, "Image denoising using deep CNN with batch renormalization," *Neural Networks*, vol. 121, pp. 461-473, 2020.

[19] Z. Mushtaq, S.-F. Su, and Q.-V. Tran, "Spectral images based environmental sound classification using CNN with meaningful data augmentation," *Applied Acoustics*, vol. 172, pp. 107581, 2020.

[20] Y. Su, K. Zhang, J. Wang *et al.*, "Environment sound classification using a two-stream CNN based on decision-level fusion," *Sensors*, vol. 19, no. 7, pp. 1733, 2019.

[21] T. Ergin, "Convolutional Neural Network (ConvNet yada CNN) nedir, nasıl çalışır? ," 2018,October

MANAS Journal of Engineering
ISSN 1694-7398 | e-ISSN 1694-7398

Volume 9, Issue 1, (2021) Pages 93-103
https://doi.org/10.51354/mjen.846484

On some inequalities for derivatives of algebraic polynomials in unbounded regions with angles

Cevahir Doğanay Gün

Gaziantep University, Gaziantep, Turkey, cevahirdoganaygun@gmail.com, ORCID: 0000-0003-3046-7667

ABSTRACT

In this work we study Bernstein-Walsh-type estimations for the derivative of an arbitrary algebraic polynomial in regions with interior zero and exterior non zero angles.

ARTICLE INFO

Research article

Received:
Accepted:

Keywords:
Algebraic polynomial,
quasicircle,
smooth curve,
inequalities.

*Corresponding author

1. Introduction

Let $\mathbb{C}$ denote the complex plane and $\overline{\mathbb{C}} := \mathbb{C} \cup \{\infty\}$; $G \subset \mathbb{C}$ be a bounded Jordan region with boundary $L := \partial G$ such that $0 \in G$; Let $\{z_j\}_{j=1}^{l}$ be the fixed system of distinct points on the curve L. We consider generalized Jacobi weight function $h(z)$ which is defined as follows:

$$h(z) := \prod_{j=1}^{l} |z - z_j|^{\gamma_j}, \quad z \in \mathbb{C}, \tag{1}$$

where $\gamma_j > -2$, for all $j = 1,2,\dots,l$.
Let $\wp_n$ denotes the class of all algebraic polynomials $P_n(z)$ of degree at most $n \in \mathbb{N}$.
Let $p > 0$. For the Jordan region G, we introduce:

$$\|P_n\|_p := \|P_n\|_{A_p(h,G)} := \left(\iint_G h(z)|P_n(z)|^p d\sigma_z\right)^{1/p}, 0 < p < \infty, \tag{2}$$

$$\|P_n\|_\infty := \|P_n\|_{A_\infty(1,G)} := \max_{z \in \overline{G}} |P_n(z)|, p = \infty,$$

and $A_p(1, G) \equiv A_p(G)$, where σ be the two-dimensional Lebesgue measure.
When L is rectifiable, for any $p > 0$, let

$$\|P_n\|_{\mathcal{L}_p(h,L)} := \left(\int_L h(z)|P_n(z)|^p |dz|\right)^{1/p} < \infty, \ 0 < p < \infty, \tag{3}$$

$$\|P_n\|_{\mathcal{L}_\infty(1,L)} := \max_{z \in L} |P_n(z)|, p = \infty,$$

and $\mathcal{L}_p(1, L) \equiv \mathcal{L}_p(L)$.
Let us set $\Omega := \overline{\mathbb{C}} \setminus \overline{G} = extL$; $\Delta(w, R) := \{w: |w| > R, R > 1\}, \Delta := \Delta(0,1)$ and let $w = \Phi(z)$ be the univalent conformal

mapping of Ω onto Δ such that $\Phi(\infty)=\infty$ and $\lim_{z\to\infty}\frac{\Phi(z)}{z}>0$; $\Psi:=\Phi^{-1}$. For $R>1$ we define $L_R:=\{z:|\Phi(z)|=R\}$, $G_R:=intL_R$, $\Omega_R:=extL_R$.
Well known Bernstein -Walsh Lemma [26] says that:

$$\|P_n\|_{C(\overline{G}_R)}\le R^n\|P_n\|_{C(\overline{G})}. \tag{4}$$

Analogous estimation with respect to the quasinorm (4) for $p>0$ was obtained in [19] for $h(z)\equiv 1$ (i.e., $\gamma_j=0$ for all $j=1,2,\dots,l$) and in [8, Lemma 2.4] for $h(z)\neq 1$, defined as in (1) as following:

$$\|P_n\|_{\mathcal{L}_p(h,L_R)}\le R^{n+\frac{1+\gamma^*}{p}}\|P_n\|_{\mathcal{L}_p(h,L)},\ \gamma^*=\max\{0;\gamma_j:1\le j\le l\}. \tag{5}$$

To give a similar estimation to (5) for the $A_p(h,G)$ −norm, first of all we will give the following definition.

Definition 1. [20, p.97], [23] The Jordan arc (or curve) L is called K −*quasiconformal* ($K\ge 1$), if there is a K −quasiconformal mapping f of the region D ⊃ L such that f(L) is a line segment (or circle).
Let $F(L)$ denote the set of all sense preserving plane homeomorphisms f of the region $D\supset L$ such that $f(L)$ is a line segment (or circle) and let

$$K_L:=\inf\{K(f):f\in F(L)\},$$

where $K(f)$ is the maximal dilatation of f. Then L is a quasiconformal curve, if $K_L<\infty$, and L is a K −quasiconformal curve, if $K_L\le K$.

A curve L is called a *quasiconformal*, if it is a K −*quasiconformal* for some $K>1$.
The Bernstein-Walsh type estimates for the norm (2), for the regions with quasiconformal boundary and weight function $h(z)$, defined in (1) with $\gamma_j>-2$, for all $p>0$ as follows

$$\|P_n\|_{A_p(h,G_R)}\le c_1R^{*\,n+\frac{1}{p}}\|P_n\|_{A_p(h,G)}, \tag{6}$$

was found in [3] (see, also [2]), where $R^*:=1+c_2(R-1)$, $c_2>0$ and $c_1:=c_1(G,p,c_2)>0$ constants, independent from n and R. It's well known that quasiconformal curves can be non-rectifiable (see, for example, [16], [20, p.104]).
Analogous estimation was studied for $A_p(1,G)$-norm, $p>0$, for arbitrary Jordan region in [4, Theorem1.1] and for any $P_n\in\wp_n$, $R_1=1+\frac{1}{n}$ and arbitrary R, $R>R_1$, was obtain

$$\|P_n\|_{A_p(G_R)}\le cR^{n+\frac{2}{p}}\|P_n\|_{A_p(G_{R_1})},$$

where $c=\left(\frac{2}{e^p-1}\right)^{\frac{1}{p}}\left[1+O(\frac{1}{n})\right]$, $n\to\infty$.
For a rectifiable quasiconformal curve L, N. Stylianopoulos [24] obtained the following estimate:

$$|P_n(z)|\le c\frac{\sqrt{n}}{d(z,L)}\|P_n\|_{A_2(G)}|\Phi(z)|^{n+1},\quad z\in\Omega, \tag{7}$$

where $d(z,L):=\inf\{|\zeta-z|:\zeta\in L\}$, a constant $c=c(L)>0$ depending only on L.
Analogous results of (7)-type for $|P_n(z)|$, different weight function h, unbounded region Ω were obtained in [17, p.418-428], [5], [6], [7], [8], [9], [10], [11], [15], [22] and others.
In this work, we study the pointwise estimations for the derivative $|P_n'(z)|$ in unbounded region Ω with zero angles as the following type

$$|P_n'(z)|\le c_2\eta_n(G,h,p,d(z,L),|\Phi(z)|)\|P_n\|_p\ ,z\in\Omega, \tag{8}$$

where $c_2=c_2(G,p)>0$ is a constant independent of n,Z and P_n, and $\eta_n(G,h,p,d(z,L),|\Phi(z)|)\to\infty$, $n\to\infty$, depending on the properties of the G, h and from the distance of point $z\in\Omega$ to the $\overline{G}$.

2. Definitions and main results

Throughout this paper, $c, c_0, c_1, c_2, \ldots$ are positive and $\varepsilon_0, \varepsilon_1, \varepsilon_2, \ldots$ are sufficiently small positive constants (generally, different in different relations), which depends on G in general and, on parameters inessential for the argument, otherwise, the dependence will be explicitly stated. For any $k \geq 0$ and $m > k$, notation $i = \overline{k,m}$ means $i = k, k+1, \ldots, m$.
Let $z = z(s),\ s \in [0, mesL]$ denote the natural representation of L .

Definition 2. We say that $L \in C_\theta,\ if\ L$ has a continuous tangent $\theta(z) := \theta(z(s))$ at every point $z(s)$. Then we write $G \in C_\theta \Leftrightarrow \partial G \in C_\theta$.

According to the "three-point" criterion [13, p.100], every piecewise smooth curve (without any cusps) is quasiconformal. Moreover, according to [23], we have the following:

Corollary 3. *If $G \in C_\theta$, then ∂G is $(1+\varepsilon)-$quasiconformal for arbitrary small $\varepsilon > 0$.*

Now we give the definitions of regions with a piecewise smooth curve, which we present our main result and some notation that will be used later in the text.

Definition 4. [5] We say that a Jordan region $G \in C_\theta(\lambda_1, \ldots, \lambda_l),\ 0 < \lambda_j \leq 2,\ j = \overline{1,l}$,*if* $L = \partial G$ consists of the union of finite smooth arcs $\{L_j\}_{j=1}^{l}$, such that they have exterior (with respect to $\overline{G}$) angles $\lambda_j \pi,\ 0 < \lambda_j \leq 2,$ at the corner points $\{z_j\}_{j=1}^{l} \in L,$ where two arcs meet.

Without loss of generality, we assume that these points on the curve $L = \partial G$ are located in the positive direction such that, G has exterior $\lambda_j \pi,\ 0 < \lambda_j < 2, j = \overline{0, l_1},$ angle at the points $\{z_j\}_{j=1}^{l_1}$, $l_1 \leq l,$ and interior zero angle (i.e. $\lambda_j = 2 -$interior cusps) at the points $\{z_j\}_{j=l_1+1}^{l}$.

It is clear from Definition 4, the each region $G \in C_\theta(\lambda_1, \ldots, \lambda_l),\ 0 < \lambda_j \leq 2,\ j = \overline{1,l},$ may have exterior nonzero $\lambda_j \pi,\ 0 < \lambda_j < 2,$ angles at the points $\{z_j\}_{j=1}^{l_1} \in L$, and interior zero angles ($\lambda_j = 2$) at the the points $\{z_j\}_{j=l_1+1}^{l} \in L$. If $l_1 = l = 0$, then the region G doesn't have such angles, and in this case we will write: $G \in C_\theta$; if $l_1 = l \geq 1$, then G has only $\lambda_i \pi,\ 0 < \lambda_i < 2,$ $i = \overline{1, l_1},$ exterior nonzero angles, and in this case we will write: $G \in C_\theta(\lambda_i)$; if $l_1 = 0$ and $l \geq 1$, then G has only interior zero angles, and in this case we write: $G \in C_\theta(2)$.

Throughout this work, we will assume that the points $\{z_j\}_{j=1}^{l} \in L$ defined in (1) and Definition 4 are identical and $w_j := \Phi(z_j)$. For simplicity of exposition and in order to avoid cumbersome calculations, without loss of generality, we will take $l_1 = 1, l = 2$. Then, after this assumption, in the future we will have region $G \in C_\theta(\lambda_1, 2),\ 0 < \lambda_1 < 2,$ such that at the point $z_1 \in L$ region G have exterior nonzero $\lambda_1 \pi,\ 0 < \lambda_1 < 2,$ and at the point $z_2 \in L$ - interior zero angle. Note that, the notation "$G \in C_\theta(\lambda_1, \lambda_2),$ $0 < \lambda_1, \lambda_2 < 2$" means that the region G has two exterior nonzero $\lambda_j \pi, 0 < \lambda_j < 2, j = 1,2$, angles at the point $z_j \in L$.
For $0 < \delta_j < \delta_0 := \frac{1}{4}\min\{|z_1 - z_2| : j = 1,2\}$, $\delta := \min\limits_{1 \leq j \leq 2} \delta_j$, let

$$\Omega(z_j,\ \delta_j) := \Omega \cap \{z : |z - z_j| \leq \delta_j\}; \qquad (9)$$
$$\Omega(\delta) := \cup_{j=1}^{2}\ \Omega(z_j,\ \delta), \widehat{\Omega} := \Omega \setminus \Omega(\delta).$$

In this work, we study problem of (8) type in regions with piecewise smooth boundary without exterior cusps and generalized Jacobi weight function $h(z)$, as defined in (1).
Now, we start to formulate the new results.

Theorem 5. *Let $p > 1; G \in C_\theta(\lambda_1, 2)$, for some $0 < \lambda_1 < 2$; $h(z)$ be defined as in (1). Then, for any $P_n \in \wp_n,\ n \in \mathbb{N},\ \gamma_j > -2$ and arbitrary small $\varepsilon > 0$*

$$|P_n'(z)| \le c_1 \left[\frac{|\Phi(z)|^{n+1}}{d(z,L_{R_1})} G_{n,1}(z) + \frac{|\Phi(z)|^{2(n+1)}}{d^{2/p}(z,L_{1+1/n})} B_{n,1}(z) E_{n,1}\right] \|P_n\|_p$$

holds, where $c_1 = c_1(G, \gamma_i, p, \varepsilon) > 0$;

$$G_{n,1}(z) := \begin{cases} n^{\frac{\gamma_1+2}{p}\tilde{\lambda}_1}, & \gamma_1 \ge \frac{2}{\tilde{\lambda}_1}(\gamma_2+2) - 2, \gamma_2 \ge \frac{\tilde{\lambda}_1}{2\lambda_1} - 2, & z \in \Omega(\delta), \\ n^{\frac{\gamma_2+2}{p}2}, & \frac{1}{\lambda_1} - 2 \le \gamma_1 < \frac{2}{\tilde{\lambda}_1}(\gamma_2+2) - 2, \gamma_2 \ge \frac{\tilde{\lambda}_1}{2\lambda_1} - 2, & z \in \Omega(\delta), \\ n^{\frac{\gamma_2+2}{p}2}, & \gamma_1 < \frac{1}{\lambda_1} - 2, \gamma_2 \ge \frac{\tilde{\lambda}_1}{2\lambda_1} - 2, & z \in \Omega(\delta), \\ 1, & \gamma_1 < \frac{1}{\lambda_1} - 2, \gamma_2 < -\frac{3}{2}, & z \in \Omega(\delta), \\ n^{\frac{2}{p}+\varepsilon}, & for\ all\ \lambda_1, \gamma_1, \gamma_2, & z \in \widehat{\Omega}(\delta), \end{cases}$$

$$E_{n;1} := \begin{cases} n^{\frac{\tilde{\gamma}\cdot\hat{\lambda}}{p}}, & if\ \tilde{\gamma}\cdot\hat{\lambda} \ge 1, \\ n^{\frac{1}{p}}, & if\ \tilde{\gamma}\cdot\hat{\lambda} < 1, \end{cases} \quad B_{n,1}(z) := \begin{cases} n^{\hat{\lambda}}, & z \in \Omega(\delta), \\ n^{1+\varepsilon}, & z \in \widehat{\Omega}(\delta), \end{cases};$$

$$\hat{\lambda} := \begin{cases} \max\{1;\lambda\} + \varepsilon, & if\ 0 < \lambda < 2, \\ 2, & if\ \lambda = 2, \end{cases}; \quad \tilde{\gamma} := \begin{cases} \tilde{\gamma}_1, & if\ 0 < \lambda < 2, \\ \tilde{\gamma}_2, & if\ \lambda = 2, \end{cases}$$

$$\tilde{\gamma}_i := \max\{0;\gamma_i\}, \qquad i = 1,2; \quad \tilde{\lambda}_1 := \max\{1;\lambda_1\} + \varepsilon.$$

Theorem 6. *Let* $p > 1$; $G \in C_\theta(\lambda_1, \lambda_2)$, *for some* $0 < \lambda_j < 2, j = 1,2$; $h(z)$ *be defined as in (1). Then, for any* $P_n \in \wp_n$, $n \in \mathbb{N}$, $\gamma_j > -2$ *and arbitrary small* $\varepsilon > 0$

$$|P_n'(z)| \le c_2 \|P_n\|_p \left[\frac{|\Phi(z)|^{n+1}}{d(z,L_{R_1})} G_{n,2}(z) + \frac{|\Phi(z)|^{2(n+1)}}{d^{2/p}(z,L_{1+1/n})} B_{n,2}(z) E_{n,2}\right]$$

holds, where $c_2 = c_2(G, \gamma_i, p, \varepsilon) > 0$;

$$G_{n,2}(z) := \begin{cases} n^{\frac{\gamma_1+2}{p}\tilde{\lambda}_1}, & \gamma_1 \ge \frac{1}{\lambda_1} - 2, \gamma_2 < \frac{1}{\lambda_2} - 2, & z \in \Omega(\delta), \\ n^{\frac{\gamma_2+2}{p}\tilde{\lambda}_2}, & \gamma_1 < \frac{1}{\lambda_1} - 2, \gamma_2 \ge \frac{1}{\lambda_2} - 2, & z \in \Omega(\delta), \\ 1, & \gamma_1 < \frac{1}{\lambda_1} - 2, \gamma_2 < \frac{1}{\lambda_2} - 2, & z \in \Omega(\delta), \\ n^{\frac{\gamma_1+2}{p}\tilde{\lambda}_1}, & \gamma_1 \ge \frac{\lambda_2}{\lambda_1}(\gamma_2+2) - 2, \gamma_2 \ge \frac{1}{\lambda_2} - 2, & z \in \Omega(\delta), \\ n^{\frac{\gamma_2+2}{p}\tilde{\lambda}_2}, & \frac{1}{\lambda_1} - 2 \le \gamma_1 < \frac{\lambda_2}{\lambda_1}(\gamma_2+2) - 2, \gamma_2 \ge \frac{1}{\lambda_2} - 2, & z \in \Omega(\delta), \\ n^{\frac{2}{p}+\varepsilon}, & for\ all\ \lambda_1, \gamma_1, & z \in \widehat{\Omega}(\delta), \end{cases}$$

$$E_{n,2} := \begin{cases} n^{\frac{\tilde{\gamma}\cdot\tilde{\lambda}}{p}}, & if\ \tilde{\gamma}\cdot\tilde{\lambda} \ge 1, \\ n^{\frac{1}{p}}, & if\ \tilde{\gamma}\cdot\tilde{\lambda} < 1, \end{cases}; \quad B_{n,2}(z) := \begin{cases} n^{\tilde{\lambda}}, & z \in \Omega(\delta), \\ n^{1+\varepsilon}, & z \in \widehat{\Omega}(\delta), \end{cases};$$

$$\hat{\lambda} := \begin{cases} \max\{1;\lambda\} + \varepsilon, & if\ 0 < \lambda < 2, \\ 2, & if\ \lambda = 2, \end{cases}; \quad \tilde{\gamma} := \begin{cases} \tilde{\gamma}_1, & if\ 0 < \lambda < 2, \\ \tilde{\gamma}_2, & if\ \lambda = 2, \end{cases}$$
$$\tilde{\gamma}_i := \max\{0;\gamma_i\}, \quad \tilde{\lambda}_i := \max\{1;\lambda_i\} + \varepsilon, i = 1,2.$$

Analogously, we also can give a theorem for the regions such as $G \in C_\theta(2,2)$.

3. Some auxiliary results

Lemma 1. [1] *Let L be a K −quasiconformal curve, $z_1 \in L$, $z_2, z_3 \in \Omega \cap \{z: |z - z_1| \prec d(z_1, L_{R_0})\}$; $w_j = \Phi(z_j)$, $(z_2, z_3 \in G \cap \{z: |z - z_1| \prec d(z_1, L_{R_0})\}; w_j = \varphi(z_j))$, $j = 1,2,3$. Then*

a) *The statements $|z_1 - z_2| \prec |z_1 - z_3|$ and $|w_1 - w_2| \prec |w_1 - w_3|$ are equivalent.*

So are $|z_1 - z_2| \asymp |z_1 - z_3|$ and $|w_1 - w_2| \asymp |w_1 - w_3|$.

b) *If $|z_1 - z_2| \prec |z_1 - z_3|$, then*

$$\left|\frac{w_1 - w_3}{w_1 - w_2}\right|^{K^2} \prec \left|\frac{z_1 - z_3}{z_1 - z_2}\right| \prec \left|\frac{w_1 - w_3}{w_1 - w_2}\right|^{K^{-2}},$$

where $\varepsilon < 1$, $c > 1$, $R_0 > 1$ are constants, depending on G.

Corollary 7. *Under the assumptions of Lemma 1, if $z_3 \in L_{R_0}$, then*

$$|w_1 - w_2|^{K^2} \prec |z_1 - z_2| \prec |w_1 - w_2|^{K^{-2}}.$$

Corollary 8. *If $L \in C_\theta$, then*

$$|w_1 - w_2|^{1+\varepsilon} \prec |z_1 - z_2| \prec |w_1 - w_2|^{1-\varepsilon},$$

for all $\varepsilon > 0$.

The following lemma is a consequence of the results given in [18], [21], [27] and of estimate for the $|\Psi'|$ (see, for example, [14, Th.2.8]):

$$|\Psi'(\tau)| \asymp \frac{d(\Psi(\tau), L)}{|\tau| - 1}. \tag{10}$$

Let $w_j := \Phi(z_j), \varphi_j := \arg w_j$. Without loss of generality, we will assume that $\varphi_l < 2\pi$. Additionally to the notations (9), for $\eta_j = \min\limits_{t \in \partial\Phi(\Omega(z_j, \delta_j))} |t - w_j| > 0$ and $\eta := \min\{\eta_j, j = \overline{1,l}\}$ let us set: $\Delta_j(\eta_j) := \{t: |t - w_j| \le \eta_j\} \subset \Phi(\Omega(z_j, \delta_j))$, $\Delta(\eta) := \cup_{j=1}^{l} \Delta_j(\eta)$, $\widehat{\Delta}_j = \Delta \backslash \Delta(\eta_j)$; $\widehat{\Delta}(\eta) := \Delta \backslash \Delta(\eta)$; $\Delta_1' := \Delta_1'(1)$, $\Delta_1'(\rho) := \left\{t = R \cdot e^{i\theta}: R \ge \rho > 1, \frac{\varphi_0 + \varphi_1}{2} \le \theta < \frac{\varphi_1 + \varphi_2}{2}\right\}$, $\Delta_j' := \Delta_j'(1)$, $\Delta_j'(\rho) := \left\{t = R \cdot e^{i\theta}: R \ge \rho > 1, \frac{\varphi_{j-1} + \varphi_j}{2} \le \theta < \frac{\varphi_j + \varphi_0}{2}\right\}$, $j = 2,3\ldots, l$, where $\varphi_0 = 2\pi - \varphi_l; \Omega_j := \Psi(\Delta_j')$, $L_{R_1}^j := L_{R_1} \cap \Omega_j$. Clearly, $\Omega = \cup_{j=1}^{l} \Omega_j$.

The following lemma is a consequence of the results given in [27] and [18].

Lemma 2. *Let $G \in C_\theta(\lambda, \ldots, \lambda_l)$, $0 < \lambda_j < 2$, $j = 1,2,\ldots, l$,. Then*

i) *for any $w \in \Delta_j$, $|w - w_j|^{\lambda_j + \varepsilon} \prec |\Psi(w) - \Psi(w_j)| \prec |w - w_j|^{\lambda_j - \varepsilon}$, $|w - w_j|^{\lambda_j - 1 + \varepsilon} \prec |\Psi'(w)| \prec |w - w_j|^{\lambda_j - 1 - \varepsilon}$,*

ii) *for any $w \in \overline{\Delta} \backslash \Delta_j$, $(|w| - 1)^{1+\varepsilon} \prec d(\Psi(w), L) \prec (|w| - 1)^{1-\varepsilon}$, $(|w| - 1)^{\varepsilon} \prec |\Psi'(w)| \prec (|w| - 1)^{-\varepsilon}$.*

Let $\{z_j\}_{j=1}^{l}$ be a fixed system of distinct points on curve L ordered in the positive direction and the weight function $h(z)$ be defined as in (1).

Lemma 3. [6] *Let L is a K −quasiconformal curve; $R = 1 + \frac{c}{n}$. Then, for any fixed $\varepsilon \in (0,1)$ there exist a level curve $L_{1+\varepsilon(R-1)}$ such that the following holds for any polynomial $P_n(z) \in \wp_n$, $n \in \mathbb{N}$:*

$$\|P_n\|_{L_p\left(\frac{h}{|\Phi'|}, L_{1+\varepsilon(R-1)}\right)} \prec n^{\frac{1}{p}} \|P_n\|_p, \ p > 0. \tag{11}$$

Lemma 4. [6] *Let L be a K −quasiconformal curve; $h(z)$ be defined as in (1). Then, for arbitrary $P_n(z) \in \wp_n$, any $R > 1$ and $n = 1,2,\ldots,$ we have*

$$\|P_n\|_{A_p(h,G_R)} \prec \tilde{R}^{n+\frac{1}{p}}\|P_n\|_{A_p(h,G)},\ p>0, \tag{12}$$

where $\tilde{R}=1+c(R-1)$ *and* c *is independent from* n *and* R.

Lemma 5. *Let* $G\in C_\theta(\lambda_1,\dots,\lambda_l),\ 0<\lambda_j\le 2, j=\overline{1,l}$. *Then, for arbitrary* $P_n(z)\in\wp_n$ *and any* $p>0$, *we have:*

$$\|P_n\|_{A_p(h,G_{1+c/n})} \prec \|P_n\|_{A_p(h,G)}. \tag{13}$$

4. Proof of Theorems

4.1. Proof of Theorems 5 and 6.

Proof. We will prove both theorems simultaneously. Suppose that $G\in C_\theta(\lambda;2)$ $(C_\theta(\lambda_1;\lambda_2))$ for some $0<\lambda<2$; $h(z)$ be defined as in (1). For $z\in\Omega$, we define:

$$T_n(z):=\frac{P_n(z)}{\Phi^{n+1}(z)}\ . \tag{14}$$

Then

$$T_n'(z)=\frac{P_n'(z)}{\Phi^{n+1}(z)}+P_n(z)\left(\frac{1}{\Phi^{n+1}(z)}\right)',\ z\in\Omega.$$

For any $R>1$ and $R_1:=1+\frac{R-1}{2}$, Cauchy integral representation for the region Ω_{R_1} gives

$$\begin{aligned} T_n'(z)&=-\frac{1}{2\pi i}\int_{L_{R_1}}T_n(\zeta)\frac{d\zeta}{(\zeta-z)^2}\\ &=-\frac{1}{2\pi i}\int_{L_{R_1}}\frac{P_n(\zeta)}{\Phi^{n+1}(\zeta)}\frac{d\zeta}{(\zeta-z)^2},\ z\in\Omega_{R_1}, \end{aligned} \tag{15}$$

and

$$\left(\frac{1}{\Phi^{n+1}(z)}\right)'=-\frac{1}{2\pi i}\int_{L_{R_1}}\frac{1}{\Phi^{n+1}(\zeta)}\frac{d\zeta}{(\zeta-z)^2}\ ,\ z\in\Omega_{R_1}.$$

Then, from (15), we get

$$P_n'(z)=\Phi^{n+1}(z)\left[T_n'(z)-P_n(z)\left(\frac{1}{\Phi^{n+1}(z)}\right)'\right]$$

$$=\Phi^{n+1}(z)\left[-\frac{1}{2\pi i}\int_{L_{R_1}}\frac{P_n(\zeta)}{\Phi^{n+1}(\zeta)}\frac{d\zeta}{(\zeta-z)^2}+\frac{P_n(z)}{2\pi i}\int_{L_{R_1}}\frac{1}{\Phi^{n+1}(\zeta)}\frac{d\zeta}{(\zeta-z)^2}\right],\ z\in\Omega_{R_1}.$$

Therefore,

$$|P_n'(z)|\le\frac{|\Phi^{n+1}(z)|}{2\pi}\left[\int_{L_{R_1}}\left|\frac{P_n(\zeta)}{\Phi^{n+1}(\zeta)}\right|\frac{|d\zeta|}{|\zeta-z|^2}+|P_n(z)|\int_{L_{R_1}}\left|\frac{1}{\Phi^{n+1}(\zeta)}\right|\frac{|d\zeta|}{|\zeta-z|^2}\right].$$

Since $|\Phi(\zeta)|>1$, for $\zeta\in L_{R_1}$, then, we have:

$$\begin{aligned} |P_n'(z)|&<\frac{|\Phi(z)|^{n+1}}{2\pi}\left[\int_{L_{R_1}}|P_n(\zeta)|\frac{|d\zeta|}{|\zeta-z|^2}\ +|P_n(z)|\int_{L_{R_1}}\frac{|d\zeta|}{|\zeta-z|^2}\right]\\ &\le\frac{|\Phi(z)|^{n+1}}{2\pi}\left[\frac{1}{d(z,L_{R_1})}\int_{L_{R_1}}|P_n(\zeta)|\frac{|d\zeta|}{|\zeta-z|}\ +|P_n(z)|\int_{L_{R_1}}\frac{|d\zeta|}{|\zeta-z|^2}\right]. \end{aligned} \tag{16}$$

Denote by

$$A_n(z):=\int_{L_{R_1}}|P_n(\zeta)|\frac{|d\zeta|}{|\zeta-z|};\ B_n(z):=\int_{L_{R_1}}\frac{|d\zeta|}{|\zeta-z|^2}, \tag{17}$$

and will be estimate these integrals separately.
To estimate $A_n(z)$, first of all replacing the variable $\tau = \Phi(\zeta)$ and multiplying the numerator and denominator of the integrant by $\prod_{j=1}^{2} \left|\Psi(\tau) - \Psi(w_j)\right|^{\frac{\gamma_j}{p}} |\Psi'(\tau)|^{\frac{2}{p}}$ and applying the Hölder inequality, we obtain:

$$A_n(z) = \int_{L_{R_1}} |P_n(\zeta)| \frac{|d\zeta|}{|\zeta - z|} \tag{18}$$

$$= \sum_{i=1}^{3} \int_{F_{R_1}^i} \frac{\prod_{j=1}^{2} |\Psi(\tau)-\Psi(w_j)|^{\frac{\gamma_j}{p}} \left|P_n(\Psi(\tau))(\Psi'(\tau))^{\frac{2}{p}}\right| |\Psi'(\tau)|^{1-\frac{2}{p}}}{\prod_{j=1}^{2} |\Psi(\tau)-\Psi(w_j)|^{\frac{\gamma_j}{p}} |\Psi(\tau)-\Psi(w)|} |d\tau|$$

$$\leq \sum_{i=1}^{3} \left(\int_{F_{R_1}^i} \prod_{j=1}^{2} \left|\Psi(\tau) - \Psi(w_j)\right|^{\gamma_j} |P_n(\Psi(\tau))|^p |\Psi'(\tau)|^2 |d\tau| \right)^{\frac{1}{p}}$$

$$\times \left(\int_{F_{R_1}^i} \left(\frac{|\Psi'(\tau)|^{1-\frac{2}{p}}}{\prod_{j=1}^{2} |\Psi(\tau)-\Psi(w_j)|^{\frac{\gamma_j}{p}} |\Psi(\tau)-\Psi(w)|} \right)^q |d\tau| \right)^{\frac{1}{q}}$$

$$=: \sum_{i=1}^{3} A_n^i(z),$$

where $F_{R_1}^j := \Phi(L_{R_1}^1) = \Delta_j' \cap \{\tau : |\tau| = R_1\}, j = 1,2;\ F_{R_1}^3 := \Phi(L_{R_1}^1) \backslash (F_{R_1}^1 \cup F_{R_1}^{21})$ and

$$A_n^i(z) := \left(\int_{F_{R_1}^i} \left|f_{n,p}(\tau)\right|^p |d\tau| \right)^{\frac{1}{p}} \left(\int_{F_{R_1}^i} \frac{|\Psi'(\tau)|^{2-q}}{\prod_{j=1}^{2} \left|\Psi(\tau) - \Psi(w_j)\right|^{\gamma_j(q-1)} |\Psi(\tau) - \Psi(w)|^q} |d\tau| \right)^{\frac{1}{q}}$$

$$=: J_{n,1}^i \cdot J_{n,2}^i(z),$$

$$f_{n,p}(\tau) := h^{\frac{1}{p}}(\Psi(\tau)) P_n(\Psi(\tau)) (\Psi'(\tau))^{\frac{2}{p}},\ |\tau| = R_1.$$

Applying to Lemmas 3, 4 and 5, we get:

$$J_{n,1}^i \prec n^{\frac{1}{p}} \|P_n\|_p,\ i = 1,2,3. \tag{19}$$

For the estimation of the integral $J_{n,2}^i(z)$, for $i = 1,2,3$, and $j = 1,2$, we set:

$$E_{R_1}^{11}(w_j) = \left\{\tau : \tau \in F_{R_1}^j,\ \left|\tau - w_j\right| < c_j(R_1 - 1)\right\},$$
$$E_{R_1}^{12}(w_j) := \left\{\tau : \tau \in F_{R_1}^j, c_j(R_1 - 1) \leq \left|\tau - w_j\right| < \eta\right\},$$
$$E_{R_1}^{13}(w_j) := \left\{\tau : \tau \in \Phi(L_{R_1}^j), \left|\tau - w_j\right| \geq \eta\right\},$$

where $0 < c_j < \eta$ is chosen so that $\left\{\tau : \left|\tau - w_j\right| < c_j(R_1 - 1)\right\} \cap \Delta \neq \varnothing$ and $\Phi(L_{R_1}^j) = \bigcup_{k=1}^{3} E_{R_1}^{1k}(w_j)$. Taking into consideration these notations, (19) can be written as:

$$\sum_{i=1}^{3} J_{n,2}^i(z) =: J_2(z) = \sum_{i=1}^{3} \sum_{j=1}^{2} J_2(E_{R_1}^{1i}(w_j), z) \tag{20}$$
$$=: \sum_{i=1}^{3} \sum_{j=1}^{2} J_{2,j}^i(z)$$

and, consequently,

$$A_n(z) \prec n^{\frac{1}{p}} \|P_n\|_p \cdot \sum_{j=1}^{2} \sum_{i=1}^{3} J_{2,j}^i(z) =: \sum_{j=1}^{2} \sum_{i=1}^{3} A_{n,i}^j(z), \tag{21}$$

where

$$A_{n,i}^j(z) := n^{\frac{1}{p}} \|P_n\|_p \cdot J_{2,j}^i(z),\ i = 1,2,3; j = 1,2. \tag{22}$$

$$\left(J_{2,j}^{i}(z)\right)^{q} := \int_{E_{R_1}^{1i}(w_j)} \frac{|\Psi'(\tau)|^{2-q}|d\tau|}{\prod_{j=1}^{2}|\Psi(\tau)-\Psi(w_j)|^{\gamma_j(q-1)}|\Psi(\tau)-\Psi(w)|^{q}}$$

$$\asymp \sum_{j=1}^{2} \int_{E_{R_1}^{1i}(w_j)} \frac{|\Psi'(\tau)|^{2-q}|d\tau|}{|\Psi(\tau)-\Psi(w_j)|^{\gamma_j(q-1)}|\Psi(\tau)-\Psi(w)|^{q}}, \ i = 1,2,3,$$

since the points w_1 and w_2 are isolated.

Therefore, we need to estimate the quantity (21). In case of $j = 1$, for any $p > 1, 0 < \lambda_1 < 2, \gamma > -2$, and for all sufficiently small $\varepsilon > 0$, in [12] is proved following estimate:

$$\sum_{i=1}^{3} A_{n,i}^{1}(z) \prec \|P_n\|_p \cdot \begin{cases} n^{\frac{\gamma_1+2}{p}\tilde{\lambda}_1}, & 0 < \lambda_1 < 2, \\ & \gamma_1 \geq \frac{1}{\lambda_1} - 2, \quad z \in \Omega(\delta), \\ n^{\frac{2}{p}+\varepsilon}, & \text{for all } \lambda_1, \gamma_1, \quad z \in \widehat{\Omega}(\delta), \\ 1, & \text{otherwise}, \end{cases} \tag{23}$$

where $\tilde{\lambda}_1 := \max\{1; \lambda_1\} + \varepsilon$.
Similarly to the case $j = 1$, for the case $j = 2$, we obtain:

$$\sum_{i=1}^{3} A_{n,i}^{2}(z) \prec \|P_n\|_p \cdot \begin{cases} n^{\frac{\gamma_2+2}{p}2+\varepsilon}, & \gamma_2 \geq -\frac{3}{2}, \quad z \in \Omega(\delta), \\ n^{\frac{2}{p}+\varepsilon}, & \gamma_1 > -2, \quad z \in \widehat{\Omega}(\delta), \\ 1, & \text{otherwise} \end{cases} \tag{24}$$

Combining (23) and (24), for the region $G \in C_\theta(\lambda_1, 2)$, any $p > 1, \gamma_1 > -2, 0 < \lambda_1 < 2$, and for all sufficiently small $\varepsilon > 0$, we obtain:

$$A_n(z) = \sum_{k=1}^{3} A_{n,k}^{1} \prec \|P_n\|_p \times \tag{25}$$

$$\times \begin{cases} n^{\frac{\gamma_1+2}{p}\tilde{\lambda}_1}, & \gamma_1 \geq \frac{1}{\lambda_1} - 2, \quad z \in \Omega(\delta), \\ n^{\frac{2}{p}+\varepsilon}, & \text{for all } \lambda_1, \gamma_1, \quad z \in \widehat{\Omega}(\delta), \\ 1, & \text{otherwise} \end{cases} + \begin{cases} n^{\frac{\gamma_2+2}{p}2+\varepsilon}, & \gamma_2 \geq -\frac{3}{2}, \quad z \in \Omega(\delta), \\ n^{\frac{2}{p}+\varepsilon}, & \gamma_1 > -2, \quad z \in \widehat{\Omega}(\delta), \\ 1, & \text{otherwise} \end{cases}$$

$$\prec \|P_n\|_p \cdot \begin{cases} n^{\frac{\gamma_1+2}{p}\tilde{\lambda}_1}, & \gamma_1 \geq \frac{2}{\tilde{\lambda}_1}(\gamma_2+2) - 2, \gamma_2 \geq \frac{\tilde{\lambda}_1}{2\lambda_1} - 2, \quad z \in \Omega(\delta), \\ n^{\frac{\gamma_2+2}{p}2}, & \frac{1}{\lambda_1} - 2 \leq \gamma_1 < \frac{2}{\tilde{\lambda}_1}(\gamma_2+2) - 2, \gamma_2 \geq \frac{\tilde{\lambda}_1}{2\lambda_1} - 2, \quad z \in \Omega(\delta), \\ n^{\frac{\gamma_2+2}{p}2}, & \gamma_1 < \frac{1}{\lambda_1} - 2, \gamma_2 \geq \frac{\tilde{\lambda}_1}{2\lambda_1} - 2, \quad z \in \Omega(\delta), \\ 1, & \gamma_1 < \frac{1}{\lambda_1} - 2, \gamma_2 < -\frac{3}{2}, \quad z \in \Omega(\delta), \\ n^{\frac{2}{p}+\varepsilon}, & \text{for all } \lambda_1, \gamma_1, \gamma_2, \quad z \in \widehat{\Omega}(\delta), \end{cases}$$

$$=: \|P_n\|_p \cdot G_{n,1}(z)$$

If the angle at point z_2 is equals $\lambda_2\pi$ such that $0 < \lambda_1, \lambda_2 < 2,$, then, analogously to (25), for the region $G \in C_\theta(\lambda_1, \lambda_2)$ all $0 < \lambda_1, \lambda_2 < 2$,we have:

$$A_n(z) = \sum_{k=1}^{3} A_{n,k}^{1} \prec \|P_n\|_p \times \tag{26}$$

$$\begin{cases} n^{\frac{\gamma_1+2}{p}\ \tilde{\lambda}_1}, & \gamma_1 \geq \frac{1}{\lambda_1}-2, \quad z\in\Omega(\delta), \\ n^{\frac{2}{p}+\varepsilon}, & \text{for all } \lambda_1,\gamma_1, \quad z\in\widehat{\Omega}(\delta), \\ 1, & \text{otherwise} \end{cases} + \begin{cases} n^{\frac{\gamma_2+2}{p}\ \tilde{\lambda}_2}, & \gamma_2 \geq \frac{1}{\lambda_2}-2, \quad z\in\Omega(\delta), \\ n^{\frac{2}{p}+\varepsilon}, & \text{for all } \lambda_2,\gamma_2, \quad z\in\widehat{\Omega}(\delta), \\ 1, & \text{otherwise} \end{cases}$$

$$\prec \|P_n\|_p \times \begin{cases} n^{\frac{\gamma_1+2}{p}\ \tilde{\lambda}_1}, & \gamma_1 \geq \frac{1}{\lambda_1}-2, \gamma_2 < \frac{1}{\lambda_2}-2, & z\in\Omega(\delta), \\ n^{\frac{\gamma_2+2}{p}\ \tilde{\lambda}_2}, & \gamma_1 < \frac{1}{\lambda_1}-2, \gamma_2 \geq \frac{1}{\lambda_2}-2, & z\in\Omega(\delta), \\ 1, & \gamma_1 < \frac{1}{\lambda_1}-2, \gamma_2 < \frac{1}{\lambda_2}-2, & z\in\Omega(\delta), \\ n^{\frac{\gamma_1+2}{p}\ \tilde{\lambda}_1}, & \gamma_1 \geq \frac{\lambda_2}{\lambda_1}(\gamma_2+2)-2, \gamma_2 \geq \frac{1}{\lambda_2}-2, & z\in\Omega(\delta), \\ n^{\frac{\gamma_2+2}{p}\ \tilde{\lambda}_2}, & \frac{1}{\lambda_1}-2 \leq \gamma_1 < \frac{\lambda_2}{\lambda_1}(\gamma_2+2)-2, \gamma_2 \geq \frac{1}{\lambda_2}-2, & z\in\Omega(\delta), \\ n^{\frac{2}{p}+\varepsilon}, & \text{for all } \lambda_1,\gamma_1, & z\in\widehat{\Omega}(\delta), \end{cases}$$

$$=: \|P_n\|_p \times G_{n,2}(z).$$

Now, let us estimate $B_n(z)$. Let $G \in C_\theta(\lambda_1, 2)$. By replacing the variable $\tau = \Phi(\zeta)$ and according to (10) and Lemma 2, we obtain:

$$B_n(z) = \int_{L_{R_1}} \frac{|d\zeta|}{|\zeta - z|^2} = \int_{|\tau|=R_1} \frac{|\Psi'(\tau)||d\tau|}{|\Psi(\tau)-\Psi(w)|^2} |\Psi'(\tau)| \asymp \frac{d(\Psi(\tau),L)}{|\tau|-1}. \tag{27}$$

$$= \int_{\{|\tau|=R_1\}\cap\Delta_1} \frac{|\tau-w_1|^{\lambda_1-1-\varepsilon}|d\tau|}{|\tau-w|^{2(\lambda_1-\varepsilon)}} + \int_{\{|\tau|=R_1\}\cap\Delta_2} \frac{d(\Psi(\tau),L)|d\tau|}{(|\tau|-1)|\Psi(\tau)-\Psi(w)|^2}$$

$$+ \int_{\{|\tau|=R_1\}\cap\ (\hat{\Delta}_1\cup\hat{\Delta}_2)} \frac{(|\tau|-1)^{-\varepsilon}|d\tau|}{|\tau-w|^{2(1-\varepsilon)}}$$

$$=: B_n^1(z) + B_n^2(z) + B_n^3(z).$$

Let us set:
$F_1 := \{\{|\tau| = R_1\} \cap \Delta_1 : |\tau - w_1| \geq |\tau - w|\}$, $F_2 := \{\{|\tau| = R_1\} \cap \Delta_2 : |\tau - w_2| \geq |\tau - w|\}$,
$F_3 := (\{|\tau| = R_1\}) \backslash (F_1 \cup F_2)$.
Under this notations we have:

$$B_n^1(z) = \int_{F_1} \frac{|\tau-w_1|^{\lambda_1-1-\varepsilon}|d\tau|}{|\tau-w|^{2(\lambda_1+\varepsilon)}} + \int_{F_2} \frac{|\tau-w_1|^{\lambda_1-1-\varepsilon}|d\tau|}{|\tau-w|^{2(\lambda_1+\varepsilon)}}$$

$$\prec \begin{cases} \left(\frac{1}{n}\right)^{\lambda_1-1-\varepsilon} \int_{F_1} \frac{|d\tau|}{|\tau-w|^{2(\lambda_1-\varepsilon)}} + \int_{F_2} \frac{|d\tau|}{|\tau-w|^{\lambda_1+1+\varepsilon}}, & \text{if } \lambda_1 \geq 1, \\ \int_{F_1} \frac{|d\tau|}{|\tau-w|^{2(\lambda_1+\varepsilon)-\lambda_1+1-\varepsilon}} + \left(\frac{1}{n}\right)^{\lambda_1-1-\varepsilon} \int_{F_2} \frac{|d\tau|}{|\tau-w|^{2(\lambda_1+\varepsilon)}}, & \text{if } \lambda_1 < 1, \end{cases}$$

$$\prec \begin{cases} n^{\lambda_1+\varepsilon}, & \text{if } \lambda_1 \geq 1, \\ n^{\lambda_1+\varepsilon}, & \text{if } \lambda_1 < 1, \end{cases} \forall \varepsilon > 0;$$

$$B_n^2(z) = \int_{\{|\tau|=R_1\}\cap(\hat{\Delta}_1\cup\hat{\Delta}_2)} \frac{d(\Psi(\tau),L)|d\tau|}{(|\tau|-1)|\Psi(\tau)-\Psi(w)|^2} \prec \begin{cases} n^2, & z\in\Omega(\delta), \\ n^{1+\varepsilon}, & z\in\widehat{\Omega}(\delta), \end{cases} \forall \varepsilon > 0.$$

$$B_n^3(z) = \int_{\{|\tau|=R_1\}\cap\ \hat{\Delta}_1} \frac{(|\tau|-1)^{-\varepsilon}|d\tau|}{|\tau-w|^{2(1-\varepsilon)}} \prec n^{1+\varepsilon}, \forall \varepsilon > 0.$$

So, from (27), we have:

$$B_n(z) \prec B_{n,1}(z) := \begin{cases} n^{\hat{\lambda}}, & z\in\Omega(\delta), \\ n^{1+\varepsilon}, & z\in\widehat{\Omega}(\delta), \end{cases} \forall \varepsilon > 0. \tag{28}$$

Similarly, for the region $G \in C_\theta(\lambda_1, \lambda_2)$, we obtain:

$$B_n(z) \prec B_{n,2}(z) := \begin{cases} n^{\tilde{\lambda}}, & z \in \Omega(\delta), \\ n^{1+\varepsilon}, & z \in \widehat{\Omega}(\delta), \end{cases}, \forall \varepsilon > 0. \tag{29}$$

Now, combining (16), (17), (25), (28) and (29) for the region , any $p > 1$, $\gamma_1 > -2, 0 < \lambda_1 < 2$, and for all sufficiently small $\varepsilon > 0$,we obtain:

$$|P_n'(z)| \prec |\Phi(z)|^{n+1} \left[\frac{1}{d(z, L_{R_1})} \int_{L_{R_1}} |P_n(\zeta)| \frac{|d\zeta|}{|\zeta - z|} + |P_n(z)| \int_{L_{R_1}} \frac{|d\zeta|}{|\zeta - z|^2} \right]$$
$$\prec |\Phi(z)|^{n+1} \|P_n\|_p \left[\frac{1}{d(z, L_{R_1})} G_{n,1}(z) + |P_n(z)| B_{n,1}(z) \right], \text{if } G \in C_\theta(\lambda_1, 2),$$

$$|P_n'(z)| \prec |\Phi(z)|^{n+1} \|P_n\|_p \left[\frac{1}{d(z, L_{R_1})} G_{n,2}(z) + |P_n(z)| B_{n,2}(z) \right], \text{if } G \in C_\theta(\lambda_1, \lambda_2),$$

Now, using estimates for $|P_n(z)|$ ([25, Theorem 1 and Corollary 1]) for the cases $G \in C_\theta(\lambda_1, 2)$ and $G \in C_\theta(\lambda_1, \lambda_2)$, we get:

$$|P_n'(z)| \prec \|P_n\|_p \left[\frac{|\Phi(z)|^{n+1}}{d(z, L_{R_1})} G_{n,1}(z) + \frac{|\Phi(z)|^{2(n+1)}}{d^{\frac{2}{p}}\left(z, L_{1+\frac{1}{n}}\right)} E_{n,1} B_{n,1}(z) \right], \text{if } G \in C_\theta(\lambda_1, 2),$$

and

$$|P_n'(z)| \prec \|P_n\|_p \left[\frac{|\Phi(z)|^{n+1}}{d(z, L_{R_1})} G_{n,2}(z) + \frac{|\Phi(z)|^{2(n+1)}}{d^{\frac{2}{p}}\left(z, L_{1+\frac{1}{n}}\right)} E_{n,2} B_{n,2}(z) \right], \text{if } G \in C_\theta(\lambda_1, \lambda_2),$$

where

$$E_{n;1} := \begin{cases} n^{\frac{\tilde{\gamma} \cdot \hat{\lambda}}{p}}, & \text{if } \tilde{\gamma} \cdot \hat{\lambda} \geq 1, \\ n^{\frac{1}{p}}, & \text{if } \tilde{\gamma} \cdot \hat{\lambda} < 1, \end{cases}; \qquad E_{n,2} := \begin{cases} n^{\frac{\tilde{\gamma} \cdot \tilde{\lambda}}{p}}, & \text{if } \tilde{\gamma} \cdot \tilde{\lambda} \geq 1, \\ n^{\frac{1}{p}}, & \text{if } \tilde{\gamma} \cdot \tilde{\lambda} < 1, \end{cases}$$

$$\hat{\lambda} := \begin{cases} \max\{1; \lambda\} + \varepsilon, & \text{if } 0 < \lambda < 2, \\ 2, & \text{if } \lambda = 2, \end{cases}; \quad \tilde{\gamma} := \begin{cases} \tilde{\gamma}_1, & \text{if } 0 < \lambda < 2, \\ \tilde{\gamma}_2, & \text{if } \lambda = 2, \end{cases}$$
$$\tilde{\gamma}_i := \max\{0; \gamma_i\}, \ i = 1,2; \ \tilde{\lambda} := \max\{1; \lambda_1, \lambda_2\} + \varepsilon.$$

Therefore, we complete the proof of Theorems 5 and 6.

References

[1] Abdullayev F. G., Andrievskii V. V., "On the orthogonal polynomials in the domains with K-quasiconformal boundary", Izv. Akad. Nauk Azerb. SSR., Ser. FTM. 1, (1983), 3-7.

[2] Abdullayev F. G., "On the interference of the weight boundary contour for orthogonal polynomials over the region", J. of Comp. Anal. and Appl., 6 (1), (2004), 31-42.

[3] Abdullayev F. G., Aral D., "On the Bernstein-Walsh type Lemmas in regions of the complex plane", Ukr. Math. J., Vol. 63 (3), (2011), 337-350.

[4] Abdullayev F. G., Özkartepe P., "An analogue of the Bernstein-Walsh lemma in Jordan regions of the complex plane", Journal Ineq. and Appl., 2013:570 (2013), 7 pp.

[5] Abdullayev F. G., Gün C. D., "On the behavior of the algebraic polynomials in regions with piecewise smooth boundary without cusps", Ann. Polon. Math., 111, (2014), 39-58.

[6] Abdullayev F. G., Özkartepe P., "On the behavior of the algebraic polynomial in unbounded regions with piecewise Dini-smooth boundary", Ukr. Math. J. , 66 (5), (2015), 645-665.

[7] Abdullayev F. G., Gün C. D., Özkartepe P., "Inequalities for algebraic polynomials in regions with exterior cusps", J. Nonlinear Funct. Anal. 3, (2015), 1-32.

[8] Abdullayev F. G., Özkartepe P., "On the growth of algebraic polynomials in the whole complex plane", J. Korean Math. Soc. 52 (4), (2015), 699-725.

[9] Abdullayev F. G., Özkartepe P., "Uniform and pointwise polynomial inequalities in regions with cusps in the weighted Lebesgue space", Jaen Journal on Approximation, 7 (2), (2015), 231-261.

[10] Abdullayev F. G., Özkartepe P., "Polynomial inequalities in Lavrentiev regions with interior and exterior zero angles in the weighted Lebesgue space", Publications de l'Institut Mathématique (Beograd), 100 (114), (2016), 209-227.

[11] Abdullayev F. G., Özkartepe N.P., "Uniform and pointwise Bernstein-Walsh-type inequalities on a quasidisk in the complex plane", Bull. Belg. Math. Soc., 23 (2), (2016), 285-310.

[12] Abdullayev F. G., Gün C. D., "Bernstein-Walsh -type inequalities for derivatives of algebraic polynomials", 2020. (to appear)

[13] Ahlfors L., "Lectures on quasiconformal mappings", Princeton, NJ: Van Nostrand, 1966.

[14] Andrievskii V. V., Belyi V. I., Dzyadyk V. K., "Conformal invariants in cocstructive theory of functions of complex plane", Atlanta, World Federatin Publ. Com., 1995.

[15] Andrievskii V. V., "Weighted polynomial inequalities in the complex plane", J. Approx.Theory, 164 (9), (2012), 1165-1183.

[16] Belinskii P. P., "General properties of quasiconformal mappings", Nauka, Sib. otd., Novosibirsk, 1974. [in Russian]

[17] Dzjadyk V. K., " Introduction to the theory of uniform approximation of function by polynomials", Nauka, Moskow, 1977.

[18] Gaier D., "On the convergence of the Bieberbach polynomials in regions with corners", Constructive Approximation, 4, (1988), 289-305.

[19] Hille E., Szegö G., Tamarkin J. D., "On some generalization of a theorem of A.Markoff", Duke Math. J., 3 (1937), 729-739.

[20] Lehto O., Virtanen K.I., "Quasiconformal mapping in the plane", Springer Verlag, Berlin, 1973.

[21] Mergelyan S. N., "Some questions of constructive functions theory", Proc. of the Steklov Institute of Mathematics, Vol. XXXVII, 1951, 1-92. [in Russian]

[22] Özkartepe P., "Pointwise Bernstein-Walsh-type inequalities in regions with piecewise Dini-smooth boundary", MJEN, 5 (3), (2017), 35-47.

[23] Rickman S., "Characterisation of quasiconformal arcs", Ann. Acad. Sci. Fenn., Ser. A, Mathematica., 395, (1966), 30 pp.

[24] Stylianopoulos N., "Strong asymptotics for Bergman polynomials over domains with corners and applications", Const. Approx., 38 (1), (2012), 59-100.

[25] Tunc T., Şimşek D., Oruç E., "Pointwise Bernstein-Walsh-type inequalities in regions with interior zero angles in the Bergman space", Trans. of NAS of Azerbaijan Ser. of Phys.-Tech. and Math. Sciences, 37 (1), (2017), 169-182.

[26] Walsh J. L., "Interpolation and approximation by rational functions in the complex domain", AMS, 1960.

[27] Warschawski S. E., "Über das Randverhalten der Ableitung der Abbildungsfunktion bei konformer Abbildung", Math.Z., 35, (1932), 321-456.