



SAUCIS

Sakarya University Journal of Computer and Information Sciences

e-ISSN 2636-8129

VOLUME 3

NUMBER 2

AUGUST 2020

Burcu Demirelli Okkalioglu

// Improving Prediction Performance of Dynamic Neighbor Selection // in User-Based Collaborative Filtering

89

Onur Cihan Distributed Solution of Road Lighting Problem Over Multi-Agent Networks

99

Fatma Betul Akyol, Oguz Altun Detection of Cervix Cancer from Pap-smear Images

112

Kiyas Kayaalp, Sedat Metlek Classification of Robust and Rotten Apples by Deep Learning Algorithm

121

Mehmet Cem Catalbas An Investigation into the Relationship between Curse of Dimensionality and Dunning-Kruger Effect

131

Negin Manshouri, Mesut Melek (Masoud Maleki), Temel Kayikcioglu The Efficacy of Frontal and Temporal Lobes in Detecting 2D&3D Video Transition Using EEG

149

Goksu Zekiye Ozen, Rayimbek Sultanov, Yunus Ozen, Zahide Yilmaz Gunes A Convolutional Neural Network Using Raw EEG Signal Obtained from Single Channel for Automatic Sleep Staging



SAUCIS

Sakarya University Journal of Computer and Information Sciences
Volume: 3 – Issue No: 2 (August 2020)
<http://saucis.sakarya.edu.tr/issue/56299>

Editor in Chief

Nejat Yumusak, Sakarya University, Turkey, nyumusak@sakarya.edu.tr

Associate Editors

Muhammed Fatih Adak, Sakarya University, Turkey, fatihadak@sakarya.edu.tr

Mustafa Akpınar, Sakarya University, Turkey, akpınar@sakarya.edu.tr

Veysel Harun Sahin, Sakarya University, Turkey, vsahin@sakarya.edu.tr

Editorial Assistants - Secretary

Muhammed Kotan, Sakarya University, Turkey, mkotan@sakarya.edu.tr

Sumeyye Kaynak, Sakarya University, Turkey, sumeyye@sakarya.edu.tr

Editorial Board

Ahmet Ozmen, Sakarya University, Turkey, ozmen@sakarya.edu.tr

Ayhan Istanbulu, Balıkesir University, Turkey, iayhan@balikesir.edu.tr

Aysegül Alaybeyoğlu, İzmir Katip Celebi University, Turkey, alaybeyoglu@gmail.com

Bahadır Karasulu, Canakkale Onsekiz Mart University, bahadirkarasulu@comu.edu.tr

Celal Ceken, Sakarya University, Turkey, celalceken@sakarya.edu.tr

Cemal Kose, Karadeniz Technical University, Turkey, ckose@ktu.edu.tr

Cemil Oz, Sakarya University, Turkey, coz@sakarya.edu.tr

Cihan Karakuzu, Bilecik Seyh Edebali University, cihan.karakuzu@bilecik.edu.tr

Fahri Vatanserver, Bursa Uludag University, fahriv@uludag.edu.tr

Hamdi Melih Saraoglu, Dumlupınar University, Turkey, hmelih.saraoglu@dpu.edu.tr

Ibrahim Turkoglu, Firat University, Turkey, iturkoglu@firat.edu.tr

Ilyas Cankaya, Ankara Yildirim Beyazit University, Turkey, icankaya@ybu.edu.tr

Levent Alhan, Sakarya University, Turkey, leventalhan@sakarya.edu.tr

Kamal Z Zamli, Malaysia Pahang University, Malaysia, kamalz@ump.edu.my

Mehmet Can Sen, Ankara Yildirim Beyazit University, mehmetcansen@outlook.com

Muhammed Fatih Adak, Sakarya University, Turkey, fatihadak@sakarya.edu.tr



SAUJCIS

Editorial Board (Cont.)

Murat Kayri, Batman University, Turkey, murat.kayri@batman.edu.tr

Mustafa Akpınar, Sakarya University, Turkey, akpınar@sakarya.edu.tr

Nilufer Yurtay, Sakarya University, Turkey, nyurtay@sakarya.edu.tr

Nuri Yılmaz, Texas A&M University, US, nuri.yilmazer@tamuk.edu

Okan Erkaymaz, Zonguldak Bulent Ecevit University, oerkaymaz@beun.edu.tr

Orhan Er, Bozok University, Turkey, orhan.er@bozok.edu.tr

Omer Hulusi Dede, Sakarya University of Applied Sciences, ohdede@sakarya.edu.tr

Resul Das, Firat University, Turkey, rdas@firat.edu.tr

Resul Kara, Duzce University, Turkey, resulcara@duzce.edu.tr

Veysel Harun Sahin, Sakarya University, Turkey, vsahin@sakarya.edu.tr



SAUCIS

Sakarya University Journal of Computer and Information Sciences
Volume: 3 – Issue No: 2 (August 2020)
<http://saucis.sakarya.edu.tr/issue/56299>

Contents

Author(s), Paper Title	Pages
<i>Burcu Demirelli Okkalioglu,</i> Improving Prediction Performance of Dynamic Neighbor Selection in User-Based Collaborative Filtering	74-88
<i>Onur Cihan,</i> Distributed Solution of Road Lighting Problem Over Multi-Agent Networks	89-98
<i>Fatma Betul Akyol, Oguz Altun,</i> Detection of Cervix Cancer from Pap-smear Images	99-111
<i>Kiyas Kayaalp, Sedat Metlek,</i> Classification of Robust and Rotten Apples by Deep Learning Algorithm	112-120
<i>Mehmet Cem Catalbas,</i> An Investigation into the Relationship between Curse of Dimensionality and Dunning-Kruger Effect	121-130
<i>Negin Manshouri, Mesut Melek (Masoud Maleki), Temel Kayikcioglu,</i> The Efficacy of Frontal and Temporal Lobes in Detecting 2D&3D Video Transition Using EEG	131-148
<i>Goksu Zekiye Ozen, Rayimbek Sultanov, Yunus Ozen, Zahide Yilmaz Gunes,</i> A Convolutional Neural Network Using Raw EEG Signal Obtained from Single Channel for Automatic Sleep Staging	149-158

Improving Prediction Performance of Dynamic Neighbor Selection in User-Based Collaborative Filtering

 Burcu Demirelli Okkalioglu¹

¹Corresponding Author; Yalova University, Faculty of Engineering, Computer Engineering Department; bdokkalioglu@yalova.edu.tr; https://orcid.org/0000-0003-2867-4667; +902268155346

Received 05 April 2020; Revised 09 June 2020; Accepted 17 July 2020; Published online 28 August 2020

Abstract

Recommender systems have become more and more popular in online environments in recent years. Although different approaches are introduced to build a powerful recommender system, collaborative filtering is one of the most used approaches in the recommender systems. Yet, researchers still introduce new methods to improve prediction performances in collaborative filtering. k nearest neighbor algorithm is one of the most dominant and prevalent one in collaborative filtering. The underlying approach behind it is to select a predefined k neighbors for an active user among all users. In the traditional algorithm, the value of k is constant and is determined before the prediction process. Recently, scholars proposed to use dynamic k neighbor selection for each user. Inspired from this work, we propose to improve prediction performance, accuracy and coverage, of collaborative filtering systems under k nearest neighbor approach. We first propose that users who rate the target item should become nominees for dynamic k neighbor selection instead of all possible users whose similarities can be calculated. The similarity calculation is the most crucial point of the k nearest neighbor algorithm. Furthermore, we also propose to use the significance-weighting approach in addition to the traditional Pearson correlation coefficient when identifying the best dynamic k neighbors for each user. The experimental results on the two well-known datasets show that the prediction accuracy and coverage improve in the dynamic k neighbor selection method by selecting neighbors among users who rated the target item and introducing the significance-weighting factor into the neighbor selection phase to find more eligible neighbors.

Keywords: dynamic k , significance-weighting, collaborative filtering, accuracy, coverage

Kullanıcı Tabanlı İşbirlikçi Filtrelemede Dinamik Komşu Seçiminin Tahmin Performansını Artırma

Öz

Öneri sistemleri son yıllarda çevrimiçi ortamlarda giderek daha popüler hale geldi. Güçlü bir öneri sistemi oluşturmak için farklı yaklaşımlar kullanılmasına rağmen, işbirlikçi filtreleme öneri sistemlerinde en çok kullanılan yaklaşımlardan biridir. Buna rağmen araştırmacılar, işbirlikçi filtrelemede tahmin performanslarını iyileştirmek için hala yeni yöntemler sunuyorlar. k en yakın komşu algoritması, işbirlikçi filtrelemede en dominant ve yaygın algoritmalarından biridir. Bu algoritmanın arkasındaki temel yaklaşım, aktif kullanıcı için tüm kullanıcılar arasından önceden tanımlanmış k tane komşu seçmektir. Geleneksel algoritmada k değeri sabittir ve tahmin işleminden önce belirlenir. Son zamanlarda, araştırmacılar her kullanıcı için dinamik k komşu seçimini kullanmayı önermişlerdir. Bu çalışmadan esinlenerek, en yakın komşu yaklaşımı altında işbirlikçi filtreleme sistemlerinin tahmin performansını, doğruluğunu ve kapsamını, iyileştirmeyi öneriyoruz. İlk olarak, benzerlikleri hesaplanabilen olası tüm kullanıcılar yerine hedef ürünü oylayan kullanıcıların dinamik k komşu seçimi için aday olmaları önerilir. Benzerlik hesaplaması, en yakın komşuluk algoritmasının en önemli noktasıdır. Ayrıca, her kullanıcı için en iyi dinamik k komşularını belirlerken geleneksel Pearson korelasyon katsayısına ek olarak önem-ağırlıklandırma yaklaşımını kullanmayı öneriyoruz. Bilinen iki veri kümesindeki deneysel sonuçlar, daha kalifiye komşular seçmek için komşuları hedef ürünü oylayan kullanıcılar arasından seçmenin ve komşu seçme aşamasına önem-ağırlıklandırma yöntemi eklemenin dinamik k komşu seçimi metodunun tahmin performansını iyileştirdiğini göstermiştir.

Anahtar Kelimeler: dinamik k , önem-ağırlıklandırma, işbirlikçi filtreleme, doğruluk, kapsam

1. Introduction

As the World Wide Web rapidly spread, people have begun to spend lots of time using it. People can buy anything they want in seconds or they can watch a movie without going to a cinema. The main problem is to find the right thing people need to buy or want to watch over the web. Recommender system (RS) techniques offer users to solve the information overload problem by forming a system that takes users' preferences and makes a prediction users will likely prefer. RSs have become popular in recent years [1]. They have been started to be used in many different sectors from product recommendations on e-commerce web sites to the music recommendation. The majority of e-commerce sites such as Netflix, Amazon, Spotify, and eBay use RSs to provide customers with personalized service of their favorite products [2]-[4]. The purpose of these companies is to speed up their customers' decision-making process with the help of RS.

There are mainly three types of RSs which are content-based RS, collaborative filtering (CF) and hybrid RS [5]. In content-based RSs [6], [7], keywords are used to describe items. Algorithms used in content-based RSs try to predict similar items for users that users like in the past. If the user did not like or buy an item containing a specific keyword, there might be no chance to recommend this type of item to the user, so the variety of suggestions is reduced, and this is undesirable. CF is the most known and popular algorithms in RS [8]-[10]. It is based on the assumption that users with similar preferences in the past will make the same choices in the future. Basically, CF systems collect users' ratings about items by explicitly or indirectly and compute the similarity between users or items and finally produce a recommendation. One of the most important advantages is that it does not need content of items in opposition to content-based RS. CF can correctly recommend complex items without requiring to know the contents of them. The hybrid RS has emerged to eliminate the disadvantages of existing RSs and to create a more efficient recommender system [11], [12].

CF algorithms can be divided into two main groups which are neighborhood-based and model-based. Although neighborhood-based approaches use a whole user-item matrix to produce a recommendation [13]-[15], model-based approaches use the ratings in the user-item matrix to learn a model [16]-[19]. Neighborhood-based approaches are also grouped in user-based and item-based [20]. In user-based CF, the similarities between users are calculated whereas item-based CF calculates similarities between items to make a prediction.

Many approaches are introduced in neighborhood-based CF in literature. However, k nearest neighbor (k -nn) algorithm is the most famous and dominant one in user-based and item-based CF [13], [21], [22]. The algorithm begins by calculating similarities between users or items. The algorithm selects the most similar neighbors among all users based on predetermined k value. The algorithm ends with a prediction by using ratings of the k best neighbors. It is important to choose the k value correctly for the performance of the algorithm. On the other hand, a significant issue is that the k is a fixed value. A traditional k -nn algorithm does not allow users to choose their own k value. However, allowing users to select a dynamic k value increases the accuracy of the prediction [23]. Similar to traditional k -nn approaches, an active user, who is looking for a prediction, selects the best k neighbors among all users whose similarities can be computed in the dynamic k neighbors algorithm [23]. This may lead a poor coverage compared to accuracy [23]. In addition to accuracy, coverage is a prominent factor, as well. Therefore, instead of choosing the best k neighbors among all possible users, an active user should select the best k neighbors within users who rated the target item to increase the coverage when dynamic k neighbor selection is utilized. Otherwise, there may be some neighbors in the best k neighbors who have not rated the target item but join the prediction process, such an occurrence reduces the coverage result.

There is another prominent issue that needs to be addressed beyond the coverage. When the correlation between users is calculated, a classical Pearson correlation coefficient (PCC) is used for the similarity calculation between users. It considers commonly rated items of two users even if there is only one common item. However, PCC does not weigh up the number of commonly rated items. When the correlation is calculated as 1.0 between two users, it means that they match perfectly even if these two users have only one co-rated item. However, this result may not be enough to tell that two users are similar due to the fact that there is only one co-rated item. Therefore, as the number of co-rated items

between users decreases, the similarity values between users should be decreased. Herlocker et al. [14] propose to add a correlation significance-weighting factor to PCC. With the help of the weighted factor, the similarity value of users with fewer common items is reduced. We propose to use modified PCC in conjunction with the dynamic neighbor selection approach in our work.

The main contributions of this paper are summarized as follows:

- We study how to improve coverage and accuracy results of dynamic k neighbors for each user while selecting neighbors among only users who rated the target item.
- We examine how the number of common ratings between two users besides dynamic k neighbors approach affects the prediction accuracy and coverage.
- We finally scrutinize how much contribution we achieve over existing methods using two well-known movie-recommender datasets by performing several experiments

The rest of the paper is organized as follows. In Section 2, the related work is given. Section 3 presents detailed information about neighborhood-based CF algorithms and a correlation significance-weighting factor approach. In Section 4, we describe our proposed method. In Section 5, several experiments are performed to evaluate the proposed approaches. Finally, Section 6 represents our conclusion and describes possible future research directions.

2. Related Work

Neighborhood-based schemes are the most popular algorithms used in CF systems, which are also called as memory-based CF [14], [24]-[26]. Neighborhood-based algorithms are divided into two parts in terms of the implementation, either user-based CF [24] or item-based CF [26]. User-based algorithms reveal a correlation between users, whereas item-based algorithms scrutinize relationships between items. The first comprehensive study on the design of the neighborhood-based CF is performed by Herlocker et. al [14]. The researchers tested various parameters to improve the quality of predictions. According to their study, different similarity weight functions have been tested and the PCC outperforms the others. In addition to selecting and calculating similarities, determining similar neighbors is also important in neighborhood-based algorithms. Although the most k similar neighbors are selected from descendingly sorted similarities weights [24], threshold-based approaches are also used to select similar neighbors. Kim and Yang [27] introduce a threshold-based approach to identify the best neighbors. If the similarity weight value of users is greater than the predefined value, those users are selected as appropriate neighbors.

Herlocker et. al [14] also point to a different issue. They claim that the number of commonly rated items between users is a significant factor while calculating similarities. Considering the structure of RSs, most of the items are unrated. Due to sparsity, it is difficult to find users with many co-rated items, which may lead to inaccurate predictions. PCC only calculates similarities between users regardless of how many items are commonly rated by two users. Therefore, it is not a good way to select the best neighbors just by PCC. The study in [14] introduced a new term called “a correlation significance-weighting factor” to decrease the effect of the similarity weight of a user when there are a few co-rated items between users. They claim that the more co-rated items two users vote, the more reliable the calculated similarity weight between them would be. Ma et al. [28] also use PCC-based significance weighting instead of classical PCC. The researchers modified the approach which was proposed by Herlocker et. al [14]. Polatidis and Georgiadis [29] proposed to use the enhanced PCC by dividing the similarity process into multi-levels. Their algorithm contains four levels instead of one as in [14]. The number of co-rated items is divided into four parts. The researchers define four positive numbers for each level to add the similarity weights. As the number of co-rated items increases between two users, the value of similarity weight is increased with the predefined positive number. Moreover, the researchers add a threshold value and compare this value with the similarity weight. If the similarity weight of users is greater than the threshold value and the number of co-rated items between users is appropriate for any level, the similarity weight is updated with the value of that level.

There are various studies that propose new approaches for neighborhood selections to increase the accuracy of the predictions. Zeybek and Kaleli [23] propose a dynamic neighbor selection instead of a constant number of neighbors. They claim that choosing the different number of k values for each user gives better results than the traditional neighbor selection method. Their algorithm starts by determining the best k value offline for each user and stores those values at a table to use at the online prediction process. Their result shows that using dynamic k values outperforms the traditional method. Motivated by this work, we examine how to improve the accuracy results using dynamic k values. Our study differs from the work in [23]. The researchers select the most k similar neighbors regardless of whether the target item is rated or not by those users. In our work, in the offline phase, we dynamically select the most similar k neighbors for each user among users who rated the target item. We also add “a correlation significance-weighting factor” to our study. Our aim is to improve the prediction accuracy and coverage of the dynamic k neighbor selection method by selecting neighbors among users who rated the target item and introducing the significance-weighting factor into the neighbor selection phase to find more eligible neighbors.

3. Neighborhood-based CF Algorithms

CF systems are based on the fact that users who had similar tastes in the past will likely have similar tastes in the future as well. The first step to form an effective CF system is to select similar users for an active user. There are mainly two approaches to find similar users, which is also called neighbors or neighbor users. These approaches are defined as correlation-based similarity and clustering. In this paper, the correlation-based approach is used to select neighbor users.

Neighborhood-based CF algorithms are commonly used in RS. The underlying approach of neighborhood-based CF is to find similar users when an active user asks a prediction for a target item (q). The algorithm starts by computing similarities between an active user and all users in the system. There are numerous methods to compute similarities between the active user and other users. The most popular and prevalent one is PCC [30]. Then, the algorithm selects the most similar k users as neighbors of the active user. The algorithm ends by producing a prediction for the active user on q . Figure 1 illustrates a general view of the neighborhood-based CF algorithm.

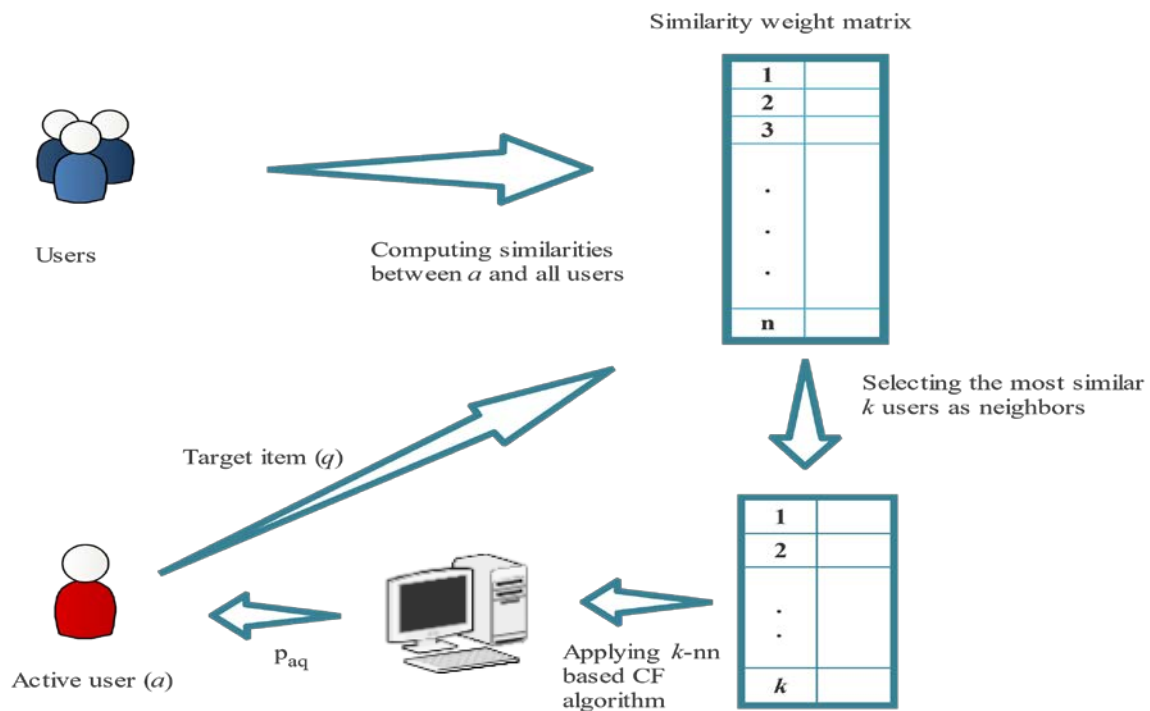


Figure 1 A General View of the Neighborhood-based CF Algorithm

3.1 Similarity Calculation

PCC is commonly used in user-based CF systems. PCC calculates similarities between an active user and other users as in Equation 1,

$$sim(\mathbf{a}, \mathbf{u}) = \frac{\sum_{p \in P} (r_{a,p} - \bar{r}_a)(r_{u,p} - \bar{r}_u)}{\sqrt{\sum_{p \in P} (r_{a,p} - \bar{r}_a)^2} \sqrt{\sum_{p \in P} (r_{u,p} - \bar{r}_u)^2}} \quad (1)$$

where P demonstrates the set of items rated both an active user, a , and a user, u , $r_{a,p}$ and $r_{u,p}$ represent ratings for item p rated by a and u , respectively. \bar{r}_a and \bar{r}_u are the mean of P items' ratings of a and u . According to Equation 1, similarities are calculated in the range between -1.0 and 1.0. 1.0 illustrates that two users are perfectly matched. It means they are the most similar. -1.0 shows the greatest degree of dissimilarity of two users.

3.2 Selecting k Neighbors

After calculating the similarities between users, the most important step is to select appropriate neighbors in order to produce accurate predictions. There are mainly two neighbor selection methods. The first one sorts calculated similarities and then selects top k users as the most similar neighbors [24], [25]. According to the second one named correlation weight threshold, a threshold value is defined, and all calculated similarities are compared to that value. If the similarity value is greater than the predefined value, that user is selected as a neighbor of an active user [27]. In this paper, we use the first method and select the most similar k users as neighbors.

3.3 Providing Predictions

After calculating similarities between the active user and all users using Equation 1 and selecting the most k similar users for the active user, the prediction $p_{a,q}$ is calculated as seen in Equation 2,

$$p_{a,q} = \bar{r}_a + \frac{\sum_{u=1}^k (r_{u,q} - \bar{r}_u) sim(a, u)}{\sum_{u=1}^k sim(a, u)} \quad (2)$$

where \bar{r}_a and \bar{r}_u are mean ratings for user a and user u respectively. k represents k -nearest-neighbors of a . $r_{u,q}$ is the rating of user u on item q and $sim(\mathbf{a}, \mathbf{u})$ illustrates the similarity value calculated using PCC between user a and user u .

3.4 A Correlation Significance-Weighting Factor

As mentioned before PCC is the most common method to measure similarities between users and is able to achieve higher performance than other similarity functions [30]. However, PCC does not take the number of co-rated items between users into account. Table 1 shows an example of the user-item matrix. The matrix contains six users and ten items. The first row of the table indicates the active user's ratings. The active user asks for a prediction for item 10.

Table 1 An Example of User-Item Matrix

	i_1	i_2	i_3	i_4	i_5	i_6	i_7	i_8	i_9	i_{10}
a	5		3	4			4			?
u_1	4		1	2		2	3			3
u_2	3	3	3	4			3			5
u_3	2		3	2			4		5	4
u_4	1		5	5	1		2	2		1
u_5		2		4		2		2	4	5

Table 2 shows the similarity weights between the active user and all users in Table 1. According to Table 2, user 5 is the most similar neighbor of the active user. However, they have one common item and it is hard to tell that they have similar preferences.

Table 2 Similarity Values Between the Active User and Other Users

$sim(a, u_1)$	$sim(a, u_2)$	$sim(a, u_3)$	$sim(a, u_4)$	$sim(a, u_5)$
0,9487	0	-0,4264	-0,7921	1

Herlocker et al. [14] proposed an approach that adds a correlation significance-weighting factor to classical PCC. A recommender system is usually sparse, and it is difficult to find users rated more ratings common with the active user. The more items are rated by the user and the active user, the more they have likely similar preferences. Herlocker et al. [14] proposed to use Equation 3, which adds a correlation significance-weighting factor to the similarity calculation.

$$sim'(a, u) = \left\{ \begin{array}{ll} \frac{|I_a \cap I_u|}{SW} * sim(a, u), & \text{if } |I_a \cap I_u| < SW \\ sim(a, u), & \text{otherwise} \end{array} \right\} \quad (3)$$

In Equation 3, $|I_a \cap I_u|$ shows the number of co-rated items, which are rated by the active user and user u . SW is a threshold value, which is a constant value. If the value of $|I_a \cap I_u|$ is less than SW , the similarity weight is decreased. Otherwise, the similarity weight remains the same. Note that the more co-rated items between users there are, the more reliable results would be.

4. Proposed Method

It is crucial that how many users are selected as the best neighbors in the k -nn based CF algorithms. Most of the algorithms apply the fixed k value. It means that the system selects the fixed k best users among all users whenever an active user attends in the system and wants a prediction for an item (q). Zeybek and Kaleli [23] proposed to use dynamic k value for all users to improve accuracy. The researchers show that accuracy results get better while the coverage decreases. Our method improves the overall accuracy and coverage of the dynamic k neighbor selection method. After calculating similarities between an active user and all users, similarities are sorted in the traditional k -nn based CF algorithm, and top k users are selected. However, we propose that the best k users should be selected among users who rated the target item, q , because some of the selected neighbors may not have rated the target item. Even in the worst case, none of them may not have rated the target item. Such an occurrence is closely associated with a decrease in coverage. Therefore, we first propose to select top k neighbors among users who have rated the target item when calculating a dynamic k value for each user offline.

We also claim that the number of co-rated items is prominent to get a higher prediction performance. As seen in Table 2, the similarity value between a and user 5 is 1.0. It is the greatest similarity value. Is it enough to tell that these two users have the same taste? Unfortunately, Table 1 shows that they have only one common item. A few number of co-rated items between a and any user do not reflect the real correlation. The more co-rated items two users have, the more reliable the correlation is. Herlocker et al. [14] proposed to use a correlation significance-weighting factor in order to improve prediction accuracy. According to their experimental results, the significance-weighting factor which is selected 50 gives the best results among other values. We also set the significance-weighting factor to 50 in our experiment. If the number of co-rated items between users is less than the predefined weighting factor, their similarity value was decreased using Equation 3. Otherwise, as seen from Equation 3 the similarity value is equal to classical PCC.

We combine two approaches, significance-weighting factor in the correlation calculation and an update to the neighbor selection, in our work as the second proposed method. After similarities are calculated using Equation 1, we recalculate similarities using Equation 3 based on commonly rated items

(significance-weighting). The first step is to identify k value for each user in dynamically. Note that top k neighbors are selected among only users who rated the target item. This is the training phase. The second, testing phase, is to produce predictions using the dynamic k -value selected in the first case.

5. Experiments

Throughout our experiments, we test if the accuracy and the coverage results get better than the existing methods [14], [23]. We perform different experiments to test our intuition and demonstrate how dynamic k values and the significance-weighting factor contribute to the results.

5.1. Datasets and Evaluation Metrics

A well-known recommender system data sets, MovieLens and FilmTrust are used in the experiments. MovieLens is a movie recommendation website (www.grouplens.org) and each rating is based on a 5-star scale. The standard 100K MovieLens data set, including 943 users and 1862 items, is used. FilmTrust is also a movie recommender dataset that consists of 1508 users and 2071 items. The rating values used in dataset are 0.5, 1, 1.5, 2, 2.5, 3, 3.5 and 4. Table 3 shows general information about datasets such as the number of users, items, and ratings and rating range.

Table 3 Datasets

	users	items	ratings	rating scale
MovieLens	943	1682	100000	1-5
FilmTrust	1508	2071	35497	0.5-4

Mean Absolute Error (MAE) is used to measure the accuracy of the recommender system as given in Equation 4, where p_i is the prediction value, o_i is the real rating value, and R is the number of provided predictions. Since this metric measures the error between the predicted and original rating, lower MAE results indicate a better prediction.

$$MAE = \frac{1}{R} \sum_{i=1}^R |p_i - o_i| \quad (4)$$

We also measure the coverage using Equation 5. Coverage is the percentage of the items that the recommender system can produce a prediction.

$$coverage = \frac{\# \text{ of provided predictions}}{\# \text{ of all queried items for a prediction}} \quad (5)$$

5.2. Methodology

Our experiments are divided into two parts: offline and online phase. In all experiments, an active user is selected using the leave-one-out cross-validation method (AllBut1). Each time, one active user is selected from the whole dataset as a test data and the rest of users form our training dataset. After selecting an active user, a prediction is produced for each item (AllBut1) to identify the best k value for each active user offline. All k values are stored in a table for the stage of the prediction process.

In the online part, we randomly select five items among rated items by the active user as in [23] and estimate predictions for those items using dynamic k values and predefined significance weight value. We repeat this experiment 100 times to obtain average prediction results.

5.3. Experimental Results and Discussion

As explained before, each user should pick his value of k dynamically. Possible k values are set 5, 10, 15, 20, 25, 30, 35, 40, 45 and 50. After calculating dynamic k values for each user, these values are

stored in the table. When the prediction process is executed, k value for the active user is looked up from the table and top k neighbors among users who have rated the target item are selected.

Before the detailed experimental results, we want to demonstrate how dynamic k values and MAE change based on the proposed methods. Table 4 shows randomly selected ten users and their calculated dynamic k values for FilmTrust dataset. Table 4 contains four columns. The first two columns show the results for only dynamic k value and its MAE and the last two columns illustrate the results for the proposed method. The number of co-rated items, which is indeed SW is set 50. “NaN” shows user 4 has no neighbors or co-rated items with any users and his k value remains 50 as a default value. In such a case, the k value is not important because the prediction cannot be provided for such users.

Table 4 Dynamic k Value for FilmTrust Users Based on the Proposed Method

	the best k	MAE	the best k when SW =50	MAE
u_1	35	0.40045	15	0.32749
u_2	10	0.22344	5	0.26139
u_3	20	0.52459	5	0.35026
u_4	50	NaN	50	NaN
u_5	50	0.69064	10	0.63377
u_6	25	1.40790	25	1.33525
u_7	15	1.07709	50	1.11342
u_8	35	0.47680	20	0.44062
u_9	10	0.43242	25	0.30979
u_{10}	45	1.39214	5	1.26631

We conduct a new experiment in order to measure the effect of nominating users who rated the target item as possible neighbors in dynamic k neighbor selection. Recall that dynamic k neighbor selection nominates all users whose correlation/similarity can be calculated [23]. Each user in the dataset is selected as an active user and the rest form our training data. All rated items for each active user are estimated and MAE is calculated. Figure 2 shows the experimental results and the experiment compares four different approaches. The first and third use a fixed k value while the second and fourth algorithms use dynamic k value. First, we would like to show how MAE results change when neighbors are selected among users who have rated the target item. When k is fixed at 50, Figure 2 (a) shows that MAE is 0.8858 when neighbors are selected from all users while MAE is 0.7565 when neighbors are selected from users who rated the target item for MovieLens dataset. Even only choosing a fixed k neighbors among users who have rated the target item increases the accuracy dramatically. Furthermore, we also compare dynamic k neighbor selection in [23] with our first proposed method. MAE of our first proposed method is 0.7379 while MAE in [23] is 0.7961. Figure 2 illustrates that selecting dynamic k neighbors among users who have rated the target item for each active user gets the best accuracy by outperforming the previous two baseline methods and dynamic k neighbor selection in [23]. One can see in Figure 2 (b) that the accuracy results for FilmTrust dataset confirm our findings.

Besides accuracy, the coverage is also important in RS. Therefore, we also calculate the coverage values for baseline methods and the first proposed method. As expected, the coverage in our method of nominating users who rate the target item for the fixed k neighbor selection is better than nominating all users regardless of their rating on the target item. Furthermore, in dynamic k neighbor selection, our proposed method gives the best coverage. As mentioned before, we inspired from an earlier work by Zeybek and Kaleli [23]. The most important difference in our study is to choose the value of k from the users who have voted for the target item. When we compare our work with them, our coverage values increase although their coverage decreases. As seen from Figure 3, the coverage for MovieLens is 0.91 and the coverage for FilmTrust is 0.87. However, Zeybek and Kaleli [23] have shown the value of coverage in the range of 0.54 – 0.71 for MovieLens dataset in their work.

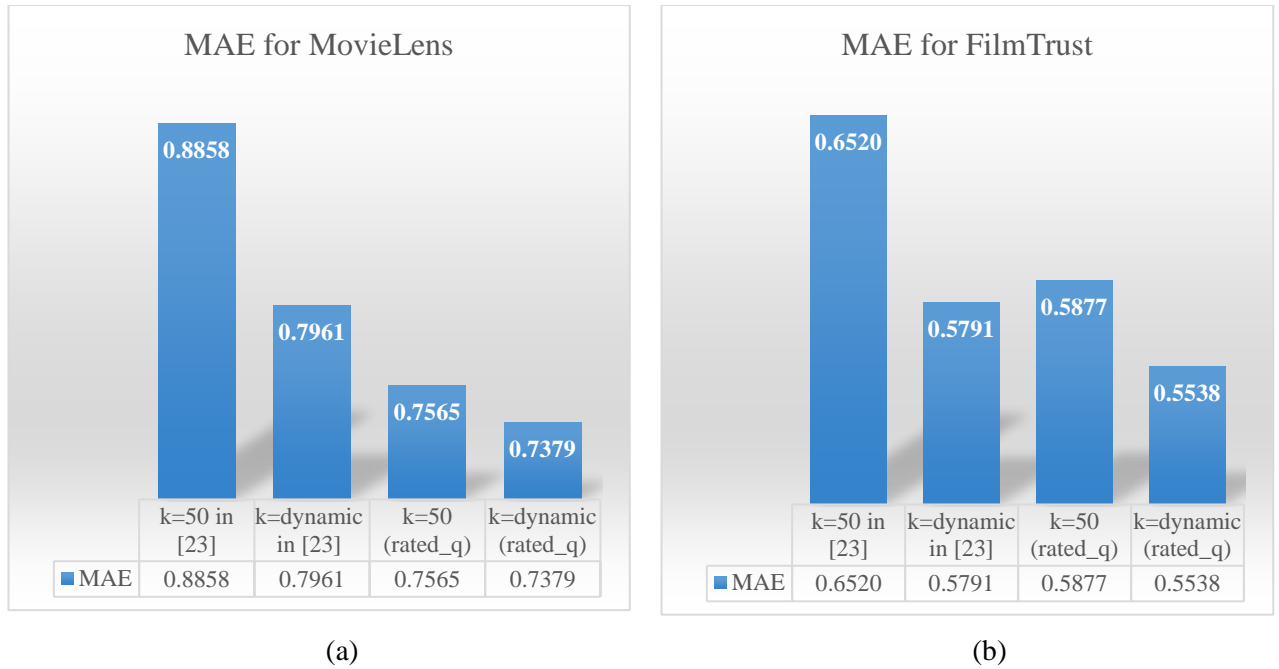


Figure 2 Accuracy Results Based on Selecting k (a) Accuracy Results for MovieLens; (b) Accuracy Results for FilmTrust

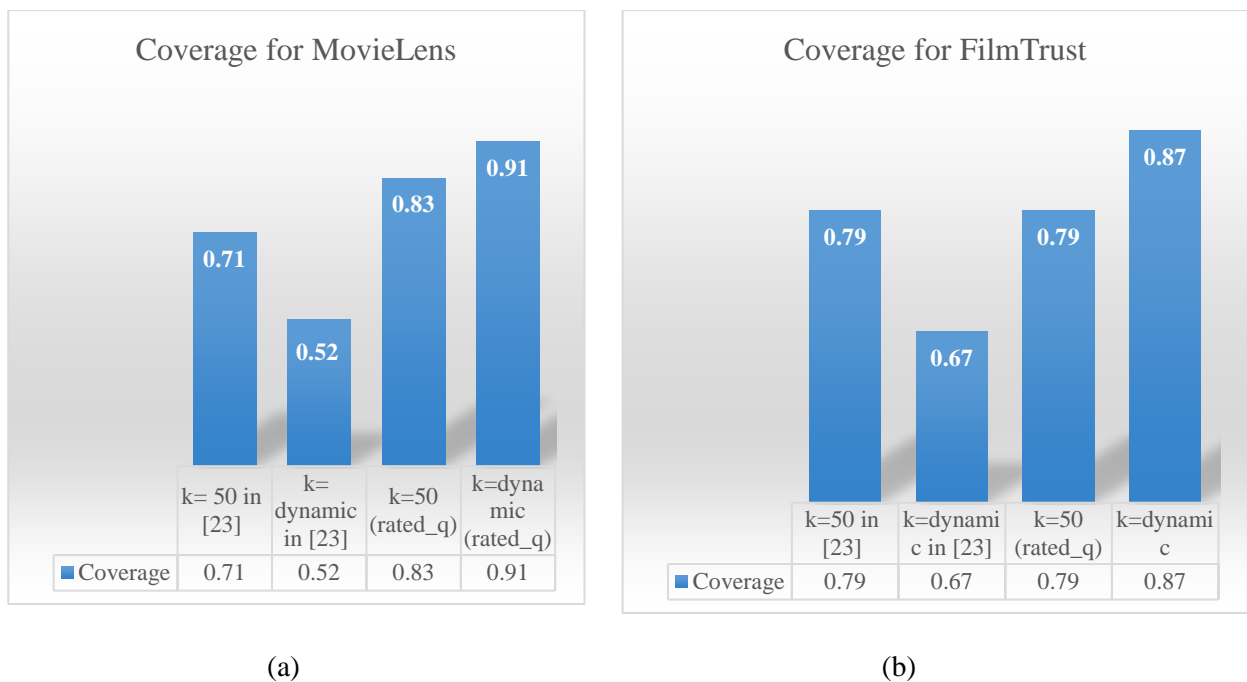


Figure 3 Coverage Results Based on Selecting k (a) Coverage Results for MovieLens; (b) Coverage Results for FilmTrust

We also test our second proposed approach. The number of co-rated items between users is crucial as much as higher similarity weights. As discussed throughout the paper, higher similarities between users may not always indicate that these two users have common tastes. Recall that the significance-weighting factor tries to balance the similarity of two users by considering the number of commonly rated items they have. The correlation significance-weighting factor is set 50 as in the work [14]. Figure 4 contains two sub-figures which are MAE results for MovieLens and FilmTrust. In this case, our second proposed method is compared with baseline methods. The first result shows the classical PCC with users who have rated the target item while the second result depicts the modified PCC with the significance-weight factor. The third result shows our proposed method's result and allows the system to decrease similarity

weights between users if the co-rated items are smaller than the threshold and choose dynamic k values. Figure 4 illustrates that using dynamic k value for each user with the significance-weight factor increases predictions of accuracy.

When we compare the results of Figure 4 with Figure 2, MAE for MovieLens of the baseline method, the first proposed method, and the second proposed method is 0.7565, 0.7379 and 0.7166, respectively. A lower MAE means a higher accuracy. FilmTrust dataset shows similar results. MAE for FilmTrust of the baseline method, the first proposed method, and the second proposed method is 0.5877, 0.5538 and 0.5310. The accuracy for the second proposed method gives the best.

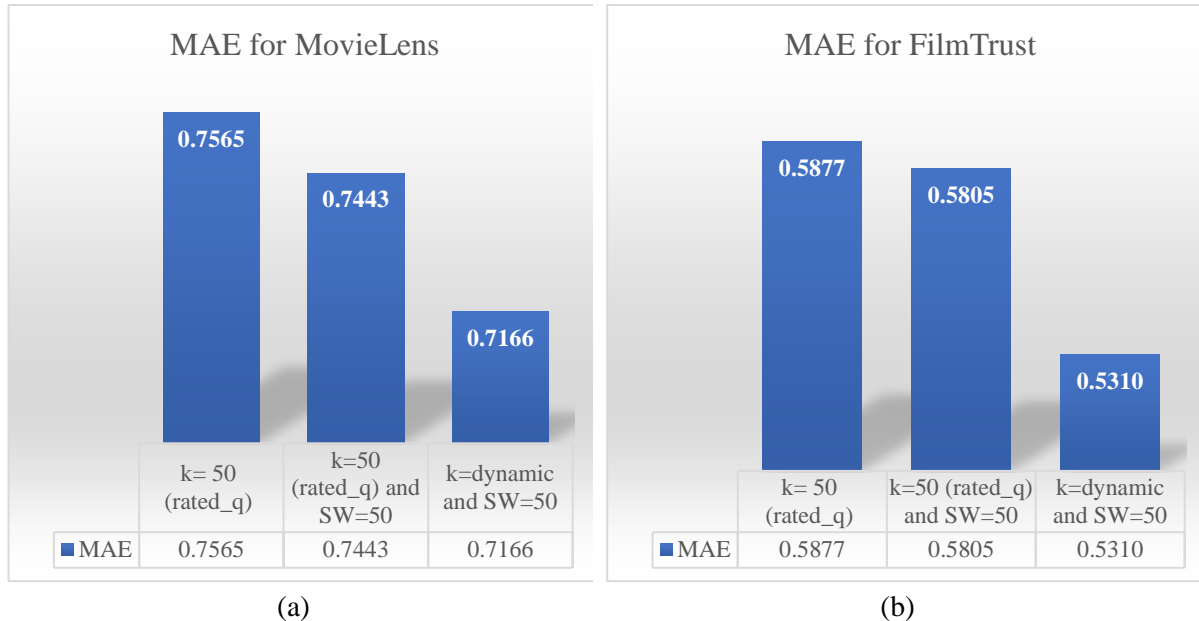


Figure 4 Accuracy Results Based on Selecting k with a Correlation Significance-weighting Factor (a) Accuracy Results for MovieLens; (b) Accuracy Results for FilmTrust

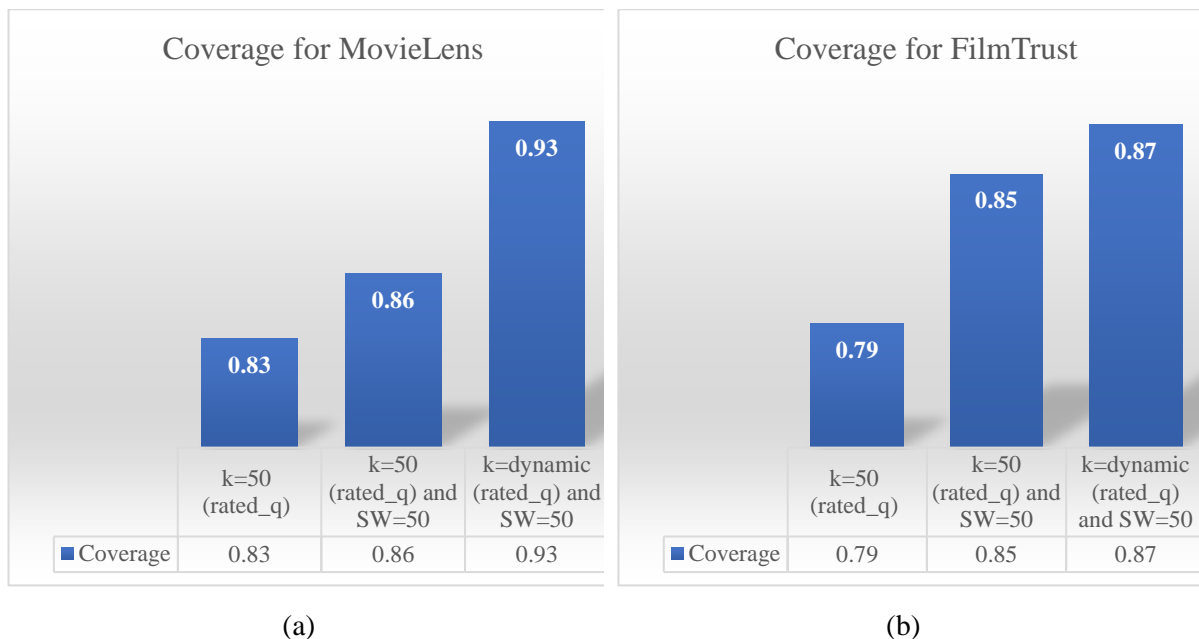


Figure 5 Coverage Results Based on Selecting k with a Correlation Significance-weighting Factor (a) Coverage Results for MovieLens; (b) Coverage Results for FilmTrust

After MAE results are given for the second proposed method in Figure 4, coverages for both datasets are measured. As seen from Figure 5, our second proposed method increases the coverage. The coverage of the classical PCC is 0.8308 while the coverage of the proposed method is 0.9247 for MovieLens.

Likewise, the coverage of the traditional PCC is 0.7896 while the coverage of the proposed method is 0.8673 for FilmTrust. Approximately 8 to 10 percent of increase has been achieved with the significance-weighting factor and the dynamic k neighbor selection.

A correlation significance-weighting factor is set 50 in all experiments. However, we also test to see if the results change with different significance-weighting factor values. The values of significance-weighting factors are set 20, 40, 50, 60 and 80. Note that dynamic k values for each user are selected in each experiment based on the predefined significance-weighting value offline. The results are shown in Table 5 and Table 6. When the results are examined, they are very close to each other. Table 5 illustrates MAE and coverage results of MovieLens dataset. In previous experiments, we set the significance-weighting factor as 50. However, someone may want to choose the significance-weighting factor as 40 according to this experiment.

Table 5 MovieLens Results Based on Various Significance-weighting (SW)

	SW=20	SW=40	SW=50	SW=60	SW=80
MAE	0.7201	0.7159	0.7166	0.7162	0.7150
Coverage	0.9236	0.9251	0.9247	0.9242	0.9237

The results in Table 6 demonstrate similar trends with Table 5. A correlation significance-weighting factor may be set as 60 according to this experiment. However, the improvement is minimal. The reason why the results are so close is that different k neighbors for each user are chosen for each value (SW).

Table 6 FilmTrust Results Based on Various Significance-weighting (SW)

	SW=20	SW=40	SW=50	SW=60	SW=80
MAE	0.5358	0.5313	0.5310	0.5303	0.5298
Coverage	0.8676	0.8676	0.8673	0.8674	0.8668

We also perform t-tests to show whether our proposed approaches results are statistically significant. In our experiments, null hypothesis regards that the prediction performances of the proposed approaches are the same as the baseline method whereas the alternative hypothesis considers that the prediction performances of the proposed approaches are better than the baseline method. Table 7 proves that the proposed approaches are statistically significant at the 99% level of confidence interval (all p values < 0.01) and the alternative hypothesis is accepted. In other words, the proposed approaches improve significantly the prediction performance based on the baseline method.

Table 7 Comparision of t-tests Between Baseline Method and Proposed Approaches

	MovieLens	FilmTrust
Baseline method ($k = 50$ in [23]) vs. $k = \text{dynamic (rated_q)}$	$t = 43.061$ $p = 4.4131e-225$	$t = 30.880$ $p = 1.6734e-154$
Baseline method ($k = 50$ in [23]) vs. $k = \text{dynamic (rated_q) and SW=50}$	$t = 42.093$ $p = 5.8468e-219$	$t = 30.764$ $p = 1.2178e-153$

6. Conclusions and Future Work

Neighborhood-based approaches are used commonly in collaborative filtering systems. Researchers find out new ways to improve the accuracy results of neighborhood-based approaches. Some of them propose new best neighbor selection algorithms although some introduce new similarity metrics. k -nn based CF algorithm is one of the most common neighbor selection algorithms in collaborative filtering. The important point of the traditional k -nn based algorithm is to select a predefined number of users as the best neighbors of the active user. A recent best neighbor selection algorithm offers to select dynamic k

neighbors for each active user instead of a constant k value. However, there is a critical issue when identifying the best neighbors: some of the selected neighbors should not have rated the target item. This might cause a decrease in coverage. During the neighbor selection process, we propose that dynamic k neighbors should be selected within users who have rated the target item. This criterion is important to identify the best neighbors. We also propose to use modified PCC instead of classical PCC. In the classical PCC, co-rated items between users participate in the similarity calculation. As the number of co-rated items between users increases, the related correlation metric is likely to be more reliable. This can be achieved by introducing significance weighting in similarity calculation. We combine this approach with the dynamic neighbor selection approach.

As seen from MovieLens results, the coverage of the method which selects the best k neighbors regardless of users who rated the target item is 0.71, whereas the coverage of the method which selects the best k neighbors among the users who rated the target item is 0.83 in traditional fixed k neighbor selection. Besides, when the dynamic k neighbors are selected among users who rated the target item, the coverage value is increased to 0.91. The coverage of the proposed method with the significance-weighting even reaches 0.93. Likewise, the coverage results of FilmTrust dataset increase approximately from 0.79 to 0.87. Approximately 8 to 10 percent of increase has been achieved with the significance-weighting factor and dynamic k value.

Mean absolute error results also decrease with the proposed methods. Mean absolute error of the baseline method, which does not consider whether the target item is rated or not, is 0.8858 for MovieLens whereas mean absolute error of the second proposed method, which considers if the target item is rated and utilizes significance weighting, is 0.7166. FilmTrust dataset shows similar results and mean absolute errors decrease with the proposed methods. The results show that it is prominent that selected neighbors should be within users who have rated the target item.

We plan to offer a new user similarity metric to improve our proposed method. There are methods that consider co-rated items between two users. These methods can be combined with existing user similarity metrics to improve prediction performance.

References

- [1] J. Bobadilla, F. Ortega, A. Hernando and A. Gutiérrez, "Recommender systems survey", *Knowledge-Based Systems*, vol. 46, pp. 109-132, 2013. Available: 10.1016/j.knosys.2013.03.012.
- [2] G. Linden, B. Smith and J. York, "Amazon.com recommendations: item-to-item collaborative filtering", *IEEE Internet Computing*, vol. 7, no. 1, pp. 76-80, 2003. Available: 10.1109/mic.2003.1167344.
- [3] C. Gomez-Urbe and N. Hunt, "The Netflix Recommender System: Algorithms, Business Value, and Innovation", *ACM Transactions on Management Information Systems*, vol. 6, no. 4, pp. 1-19, 2016. Available: 10.1145/2843948.
- [4] J. Pérez-Marcos and V. López Batista, "Recommender System Based on Collaborative Filtering for Spotify's Users," in Trends in Cyber-Physical Multi-Agent Systems. *The PAAMS Collection - 15th International Conference, PAAMS 2017, Cham, 2018*, pp. 214–220, doi: 10.1007/978-3-319-61578-3_22.
- [5] P. B. Thorat, R. M. Goudar, and S. Barve, "Survey on Collaborative Filtering, Content-based Filtering and Hybrid Recommendation System," *IJCA*, vol. 110, no. 4, pp. 31–36, Jan. 2015, doi: 10.5120/19308-0760.

- [6] P. Messina, V. Dominguez, D. Parra, C. Trattner, and A. Soto, "Content-based artwork recommendation: integrating painting metadata with neural and manually-engineered visual features," *User Model User-Adap Inter*, vol. 29, no. 2, pp. 251–290, Apr. 2019, doi: 10.1007/s11257-018-9206-9.
- [7] P. Lops, D. Jannach, C. Musto, T. Bogers, and M. Koolen, "Trends in content-based recommendation: Preface to the special issue on Recommender systems based on rich item descriptions," *User Model User-Adap Inter*, vol. 29, no. 2, pp. 239–249, Apr. 2019, doi: 10.1007/s11257-019-09231-w.
- [8] X. Su and T. M. Khoshgoftaar, "A Survey of Collaborative Filtering Techniques," *Advances in Artificial Intelligence*, vol. 2009, pp. 1–19, 2009, doi: 10.1155/2009/421425.
- [9] Y. Shi, M. Larson, and A. Hanjalic, "Collaborative Filtering beyond the User-Item Matrix: A Survey of the State of the Art and Future Challenges," *ACM Comput. Surv.*, vol. 47, no. 1, pp. 1–45, Jul. 2014, doi: 10.1145/2556270.
- [10] D. Kluver, M. D. Ekstrand, and J. A. Konstan, "Rating-Based Collaborative Filtering: Algorithms and Evaluation," in *Social Information Access*, vol. 10100, P. Brusilovsky and D. He, Eds. Cham: Springer International Publishing, 2018, pp. 344–390.
- [11] A. B. Barragáns-Martínez, E. Costa-Montenegro, J. C. Burguillo, M. Rey-López, F. A. Mikic-Fonte, and A. Peleteiro, "A hybrid content-based and item-based collaborative filtering approach to recommend TV programs enhanced with singular value decomposition," *Information Sciences*, vol. 180, no. 22, pp. 4290–4311, Nov. 2010, doi: 10.1016/j.ins.2010.07.024.
- [12] T. K. Paradarami, N. D. Bastian, and J. L. Wightman, "A hybrid recommender system using artificial neural networks," *Expert Systems with Applications*, vol. 83, pp. 300–313, Oct. 2017, doi: 10.1016/j.eswa.2017.04.046.
- [13] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions," *IEEE Trans. Knowl. Data Eng.*, vol. 17, no. 6, pp. 734–749, Jun. 2005, doi: 10.1109/TKDE.2005.99.
- [14] J. Herlocker, J. A. Konstan, and J. Riedl, "An Empirical Analysis of Design Choices in Neighborhood-Based Collaborative Filtering Algorithms," *Information Retrieval*, vol. 5, no. 4, pp. 287–310, Oct. 2002, doi: 10.1023/A:1020443909834.
- [15] C. Kaleli, "An entropy-based neighbor selection approach for collaborative filtering," *Knowledge-Based Systems*, vol. 56, pp. 273–280, Jan. 2014, doi: 10.1016/j.knsys.2013.11.020.
- [16] Y. Koren, R. Bell, and C. Volinsky, "Matrix Factorization Techniques for Recommender Systems," *Computer*, vol. 42, no. 8, pp. 30–37, Aug. 2009, doi: 10.1109/MC.2009.263.
- [17] L. M. de Campos, J. M. Fernández-Luna, J. F. Huete, and M. A. Rueda-Morales, "Combining content-based and collaborative recommendations: A hybrid approach based on Bayesian networks," *International Journal of Approximate Reasoning*, vol. 51, no. 7, pp. 785–799, Sep. 2010, doi: 10.1016/j.ijar.2010.04.001.

- [18] X. He, L. Liao, H. Zhang, L. Nie, X. Hu, and T.-S. Chua, "Neural Collaborative Filtering," in *Proceedings of the 26th International Conference on World Wide Web - WWW '17*, Perth, Australia, 2017, pp. 173–182, doi: 10.1145/3038912.3052569.
- [19] J. Li *et al.*, "Category Preferred Canopy-K-means based Collaborative Filtering algorithm," *Future Generation Computer Systems*, vol. 93, pp. 1046–1054, Apr. 2019, doi: 10.1016/j.future.2018.04.025.
- [20] X. Ning, C. Desrosiers, and G. Karypis, "A Comprehensive Survey of Neighborhood-Based Recommendation Methods," in *Recommender Systems Handbook*, F. Ricci, L. Rokach, and B. Shapira, Eds. Boston, MA: Springer US, 2015, pp. 37–76.
- [21] Y. Park, S. Park, W. Jung, and S. Lee, "Reversed CF: A fast collaborative filtering algorithm using a k-nearest neighbor graph," *Expert Systems with Applications*, vol. 42, no. 8, pp. 4022–4028, May 2015, doi: 10.1016/j.eswa.2015.01.001.
- [22] D.-K. Chae, S.-C. Lee, S.-Y. Lee, and S.-W. Kim, "On identifying k -nearest neighbors in neighborhood models for efficient and effective collaborative filtering," *Neurocomputing*, vol. 278, pp. 134–143, Feb. 2018, doi: 10.1016/j.neucom.2017.06.081.
- [23] H. Zeybek and C. Kaleli, "Dynamic k Neighbor Selection for Collaborative Filtering," *Anadolu university journal of science and technology A - Applied Sciences and Engineering*, pp. 1–1, Mar. 2018, doi: 10.18038/aubtda.346407.
- [24] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, "GroupLens: an open architecture for collaborative filtering of netnews," in *Proceedings of the 1994 ACM conference on Computer supported cooperative work - CSCW '94*, Chapel Hill, North Carolina, United States, 1994, pp. 175–186, doi: 10.1145/192844.192905.
- [25] J. A. Konstan, B. N. Miller, D. Maltz, J. L. Herlocker, L. R. Gordon, and J. Riedl, "GroupLens: applying collaborative filtering to Usenet news," *Commun. ACM*, vol. 40, no. 3, pp. 77–87, Mar. 1997, doi: 10.1145/245108.245126
- [26] B. Sarwar, G. Karypis, J. Konstan, and J. Reidl, "Item-based collaborative filtering recommendation algorithms," in *Proceedings of the tenth international conference on World Wide Web - WWW '01*, Hong Kong, Hong Kong, 2001, pp. 285–295, doi: 10.1145/371920.372071.
- [27] T.-H. Kim and S.-B. Yang, "An Effective Threshold-Based Neighbor Selection in Collaborative Filtering," in *Advances in Information Retrieval*, vol. 4425, G. Amati, C. Carpineto, and G. Romano, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 712–715.
- [28] H. Ma, I. King, and M. R. Lyu, "Effective missing data prediction for collaborative filtering," in *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval - SIGIR '07*, Amsterdam, The Netherlands, 2007, p. 39, doi: 10.1145/1277741.1277751.
- [29] N. Polatidis and C. K. Georgiadis, "A multi-level collaborative filtering method that improves recommendations," *Expert Systems with Applications*, vol. 48, pp. 100–110, Apr. 2016, doi: 10.1016/j.eswa.2015.11.023

- [30] J. S. Breese, D. Heckerman, and C. Kadie, "Empirical analysis of predictive algorithms for collaborative filtering," in *Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence*, Madison, Wisconsin, Jul. 1998, pp. 43–52.

Distributed Solution of Road Lighting Problem Over Multi-Agent Networks

 Onur CİHAN

Marmara University, Faculty of Engineering, Department of Electrical and Electronics Engineering;
onur.cihan@marmara.edu.tr; <https://orcid.org/0000-0002-5729-2417>; +90 216 777 3673

Received 7 April 2020; Revised 15 May 2020; Accepted 21 May 2020; Published online 28 August 2020

Abstract

In this study, we consider the solution of the road lighting problem by distributed algorithms over multi-agent networks where the objective is to determine the powers of the lamps that provide the desired road lighting level for a given road profile. The road is modeled as multiple road sections each with a length of 50 meters where a lighting pole is located in the middle of each section. Under given assumptions, the illumination levels of the road sections are expressed as linear functions of the powers of the lamps. When the processing units in the lighting poles can communicate wirelessly with the neighboring processing units and make simple calculations, it is shown that the power levels of the lamps that provide the desired lighting level for each road section can be calculated in a distributed manner. Finally, the model and the proposed solution has been verified by a numerical example.

Keywords: multi-agent systems, distributed algorithms, distributed control, road illumination

Yol Aydınlatma Probleminin Çok Etmenli Ağlarda Dağıtık Çözümü

Öz

Bu çalışmada, verilen bir yol profili için yol aydınlatma düzeyinin istenilen değerlerde olmasını sağlayan lamba güçlerini belirleme probleminin çok etmenli sistemler üzerinde tanımlanan bir algoritma ile dağıtık çözümü ele alınmıştır. Söz konusu yol, ellişer metrelik uzunluğa sahip bölümler halinde modellenmiş ve her bir bölümün ortasında bir aydınlatma direği konumlandırılmıştır. Yapılan varsayımlar altında yol bölümlerinin aydınlanma düzeyleri, lambaların güçlerinin bir doğrusal fonksiyonu olduğu ifade edilmiştir. Aydınlatma direklerinde bulunan işlem birimlerinin kendilerine yakın olan direklerdeki işlem birimleriyle haberleşebildiği ve basit hesaplamalar yapabildiği durumda, yolun istenilen bir aydınlık seviyesine sahip olabilmesi için gerekli olan lamba gücü seviyelerinin, dağıtık olarak hesaplanabildiği gösterilmiş ve sayısal sonuçlar ile modelin ve çözümün geçerliği doğrulanmıştır.

Anahtar Kelimeler: çok etmenli sistemler, dağıtık algoritmalar, dağıtık kontrol, yol aydınlatması

1. Giriş

Yolun yeterli ve doğru bir şekilde aydınlatılması, güvenli bir gece sürüşü için büyük öneme sahiptir. İstatistiklere göre doğru bir şekilde aydınlatılmamış bir yolda meydana gelen kaza oranı, gündüz meydana gelen kaza oranının üç katıdır [1]. Bunun yanı sıra, yolda seyir halinde bulunan araçların sürücülerinin konforlu bir sürüş deneyimi yaşayabilmeleri için, yolun aydınlık seviyesinin belirli bir değer aralığında ve parlıltı dağılımının homojene olabildiğince yakın olması istenmektedir [2]. Yol aydınlatmasının iyi olmadığı yerlerde yapılacak doğru bir aydınlatma ile trafik kazaları kayda değer oranda azaltılabilmektedir. Bu amaçla otoyol aydınlatma lambalarının konumlandırılmaları ve bu lambaların güçlerinin doğru şekilde ayarlanması büyük önem arz etmektedir.

Yol aydınlatma problemi, belirli varsayımlar altında bir doğrusal denklem olarak ifade edilebilmekte ve istenilen bir aydınlık seviyesini sağlayan lamba gücü değerleri bu denklemin çözümü olarak hesaplanabilmektedir. Ancak yolun belirli kısımlarının farklı düzeyde aydınlatılması istendiğinde, bu yeni denklem sisteminin çözülerek uygun lamba güçlerinin tekrar hesaplanması gerekmektedir. Bu

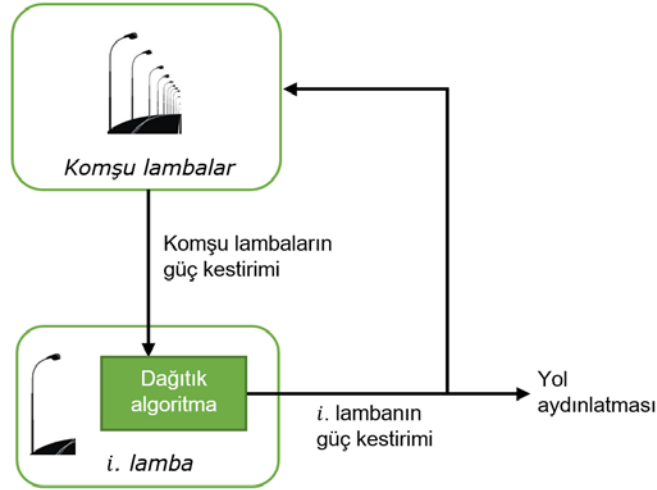
çalışmada, işlem birimleri ile güçleri kontrol edilebilen lambalara sahip bir otoyolun, istenilen aydınlık seviyesini sağlayacak şekilde, lamba güçlerini dağıtık olarak hesaplama mekanizması incelenmiştir. Bu amaçla, doğrusal denklem sistemlerinin çözümü için yeni bir yaklaşım olan ve çok etmenli sistemlerde kullanılan bir dağıtık algoritmadan yararlanılmıştır.

Literatürde doğrusal denklem sistemlerinin çözümü, araştırmacıların üzerinde çok uzun zamandır çalışılmakta olduğu ve birçok farklı mühendislik uygulamasında karşılaşılan önemli bir problemdir. Bu amaçla Jacobi yöntemi, Gauss-Seidel yöntemi, Kaczmarz yöntemi gibi birçok klasik algoritma başarı ile geliştirilmiştir [3]. Bu algoritmaların ortak özelliği, denklem sisteminin tamamının bilindiği varsayımına sahip olmalarıdır ve bu nedenle merkezleştirilmiş algoritmalar olarak adlandırılırlar. Ancak kısmi türevli diferansiyel denklemler [4], hesaplamalı akışkanlar dinamiği [5], elektromanyetik hesaplamaları [6], güç sistemleri tahminlemesi [7] ve arama motorları için geliştirilen pagerank algoritmaları [8] gibi birçok pratik uygulamada denklem sistemindeki bilinmeyen sayısı çok fazladır ve merkezleştirilmiş algoritmaların bu denklem sistemlerini çözmesi pratik değildir. Bunun yanında, doğrusal denklem sistemini oluşturan denklemler birbirinden çok farklı fiziksel konumlarda ortaya çıkabileceği için merkezleştirilmiş bir algoritmanın kullanılabilmesi için tüm denklemlerin merkezi işlemci tarafından tek bir yerde toplanması gerekmektedir.

Merkezleştirilmiş algoritmaların aksine, dağıtık algoritmalar çok fazla değişken sayısına sahip doğrusal denklem sistemlerinin çözümü için iyi bir alternatif olması nedeniyle son 10 yılda araştırmacılar tarafından çalışılmaktadır [7, 9-22]. Bu dağıtık algoritmalar, denklem sisteminin yalnızca bir kısmını içeren birden fazla sayıda alt sisteme ayırmakta ve her alt sistemin çok etmenli sistemleri oluşturan bir etmen tarafından ele alındığı ve komşu etmenlerle bilgi paylaşımı yaparak tüm denklem sisteminin çözümünü bulmalarını sağlamaktadır. Mou ve ark. tarafından, etmenlerin komşularıyla kendi alt sistemleri için ürettikleri çözümü paylaştıkları, komşularının çözümlerini de kullanarak çözümlerini güncelledikleri ve nihayetinde denklem sisteminin tamamına çözüm buldukları bir dağıtık algoritma önerilmiştir [9]. Bu algoritmanın başarısının ardından, ayrık zamanda [10-16] ve sürekli zamanda [17-21] benzer bilgi paylaşım kısıtları altında birçok algoritma geliştirilmiştir. Geliştirilen bu algoritmaların temel çalışma mantığı, her bir etmenin denklem sisteminin bir alt kümesine çözüm üretirken, komşu etmenlerin denklemleri ile çözümlerinde odayışım (konsensüs) sağlamaya çalışmalarıdır [22]. Denklem sisteminin katsayılar matrisinin seyrek matris olması durumunda etmenler arası bilgi paylaşımının daha etkin bir şekilde yapıldığı bir dağıtık algoritma [10]'da verilmiştir. Algoritmaların başlangıç çözümünün, denklem alt sistemini sağlayan bir çözüm olmak zorunda olmadığı, keyfi bir başlangıç vektörü seçimi ile denklem sisteminin çözülmesini sağlayan bir algoritma ise [11]'de önerilmiştir.

Literatürde önerilen algoritmaların birçoğu, doğrusal denklem sisteminin tek çözümü olduğu durumda çözümü bulmayı amaçlamaktadırlar. Ancak bir denklem sisteminin sonsuz çözümünün olması veya çözümsüz olması durumu da söz konusudur. Denklem sisteminin sonsuz çözümünün olduğu durumda en küçük Öklid normuna sahip olan çözümün bulunabilmesini sağlayan algoritmalar [11,12]'de geliştirilmişken, denklem sisteminin çözümü olmadığı durumda etmenlerin en küçük karesel hatayı sağlayan çözümü bulmasını sağlayan algoritmalar [13, 17]'de verilmiştir.

Bu makalede, verilen herhangi bir otoyol profili için istenilen aydınlık düzeyini sağlayacak olan lamba güçlerinin belirlenmesi problemi -literatürde yer alan merkezleştirilmiş hesaplama yöntemlerinden farklı olarak- çok etmenli ağlar üzerinden doğrusal denklem sistemlerinin çözümü için tasarlanmış dağıtık bir algoritma kullanılarak çözülmüştür. Önerilen sistemin genel işleyişi Şekil 1'de gösterilmiştir. Aydınlatma direklerinde kablosuz iletişim yeteneğine sahip olan akıllı lambaların her biri, iletişim menziline bulunan komşu lambalardan gelen lamba gücü kestirimlerini de kullanarak dağıtık bir algoritma yardımıyla yolda istenilen aydınlık seviyesine ulaşmak için gerekli lamba güçleri hesabını günceller ve komşu lambalarla paylaşır. Bu güncellemeler sonucunda, tüm akıllı lambalar istenilen aydınlık seviyesine ulaşmak için gerekli olan tüm lamba güçlerini belirler ve lamba gücünü ayarlayarak yolu istenildiği şekilde aydınlatabilir.



Şekil 1 Akıllı yol aydınlatma sisteminin işleyişi

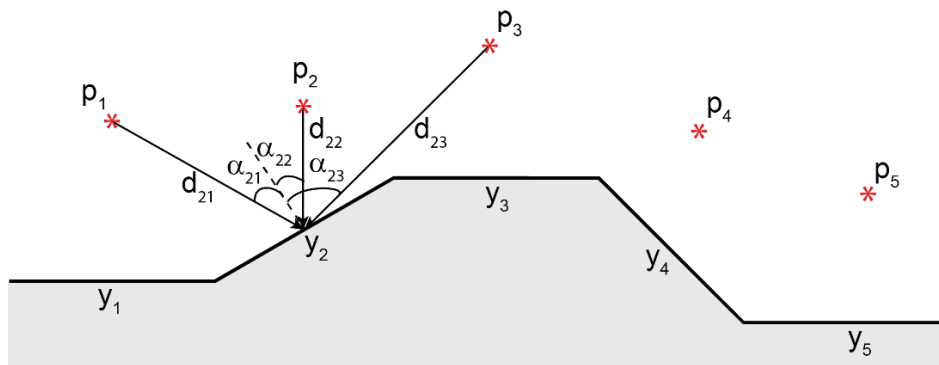
Makalenin geri kalanı şu şekilde düzenlenmiştir. Bölüm 2’de çok etmenli bir sistemde tanımlanmış akıllı yol aydınlatma modeli verilmiş ve bu problemin dağıtık olarak çözülebilmesi için kullanılacak bir algoritma tanıtılmıştır. Bu algoritmanın yol aydınlatma probleminin çözümündeki başarımı Bölüm 3’te detayları verilen benzetim çalışmaları ile gösterilmiştir. Bölüm 4’te ise, makalede elde edilen sonuçlar yorumlanarak gelecek çalışmalardan bahsedilmiştir.

2. Akıllı Yol Aydınlatma Modeli

Bu bölümde, birden fazla lamba ile aydınlatılan bir otoyola ait yol bölümlerinin aydınlık seviyelerinin lamba güçlerinin doğrusal bir fonksiyonu olduğu gösterilerek bu problemin çözülebilmesi için çok etmenli sistemler üzerinde tanımlanan dağıtık bir algoritmanın detayları verilmiştir.

2.1 Akıllı Yol Aydınlatma Sisteminin Çok Etmenli Ağ Modeli

Bu bölümde, bir yola ait farklı eğimlere sahip yol bölümlerinin birden fazla lamba tarafından aydınlatılmasının doğrusal bir denklem sistemi şeklinde modellenmesinden bahsedilmiştir. Bu amaçla, örnek bir yol aydınlatma sistemi Şekil 2’de gösterilmiştir. Şekilde yol bölümlerinin aydınlık seviyeleri y_i ($i = 1, \dots, n$) ve lamba güçleri p_j ($j = 1, \dots, n$) olarak temsil edilmekte olup; α_{ij} ifadesi j . lambadan çıkan ışığın i . yol bölümünün normali ile yaptığı açığı belirtmektedir. d_{ij} ise j . lamba ile i . yol bölümünün merkezi arasındaki mesafedir.



Şekil 2 Bir yol bölümünün birden fazla lamba ile aydınlatılması

Bu tanımlamalar kullanılarak, j . lambanın i . yol bölümünün aydınlık seviyesine katkısı

$$a_{ij} = \begin{cases} \frac{p_j \cos(\alpha_{ij})}{d_{ij}^2} & , \text{ eğer } j. \text{ lamba } i. \text{ yol bölümünü doğrudan aydınlatabiliyorsa} \\ 0 & , \text{ diğer durumlarda} \end{cases} \quad (1)$$

olarak ifade edilebilir. Birden fazla lamba tarafından aydınlatılan i . yol bölümünün aydınlık seviyesi

$$y_i = \sum_{j=1}^n a_{ij} p_j \quad (2)$$

denklemleri ile ifade edilebilir. Yolların toplam aydınlık seviyeleri $y = [y_1, y_2, \dots, y_n]^T$ ve lambaların güçleri $p = [p_1, p_2, \dots, p_n]^T$ vektörleri tanımlanarak,

$$y = Ap \quad (3)$$

doğrusal denklemi elde edilir. Bu denklemde yer alan $A = [a_{ij}]$, elemanları negatif olmayan ve $n \times n$ boyutlarında bir matristir. Yol aydınlatma probleminin çözülebilmesi için, $y = Ap$ denklemini sağlayan

$$p^* = A^{-1}y \quad (4)$$

vektörünün hesaplanması gerekmektedir.

Yorum 1: Denklem 4'te yer alan p^* çözümünün hesaplanabilmesi için son 10 yıla kadar literatürdeki algoritmaların temel varsayımı A matrisinin tüm satırlarının bir merkezi işlemci tarafından biliniyor olmasıdır. Ancak yol aydınlatma probleminde de görüleceği gibi, bazı mühendislik uygulamalarında A matrisinin bazı satırları ve y vektörünün aynı satırlara karşılık gelen elemanları fiziksel olarak birbirinden çok uzak konumlarda ortaya çıkabilmektedir. Bu nedenle her bir yol bölümünün istenilen aydınlatma düzeyine sahip olabilmesi için gerekli olan lamba güçlerini belirleyebilme problemi, dağıtık olarak çok etmenli sistemler üzerinde tanımlanarak çözülebilir.

2.2 Dağıtık Çözüm Algoritması

Bölüm 1'de bahsedildiği üzere doğrusal denklem sistemlerinin çözümü konusunda son yıllarda çeşitli algoritmalar geliştirilmiştir. Bu algoritmalar, sürekli ya da ayrık zamanda tanımlı olmasına; başlangıç çözümünün yerel denklem sistemini sağlamak zorunda olup olmamasına ve çözüm olmadığı durumda en küçük kareler çözümünü elde edebiliyor olup olmamasına göre farklı sınıflara ayrılmaktadır. Bölüm 2.1'de açıklanan yol aydınlatma probleminin çok etmenli sistemler yardımıyla çözümü için gerekli tanımlar ve dağıtık algoritma bu bölümde verilmiştir.

Etmenler arası iletişimin iki yönlü olarak sağlandığı ve $n > 1$ adet etmeden oluşan birçok etmenli sistemi ele alalım. $V = \{1, 2, \dots, n\}$ köşeler kümesi ve $E \subset V \times V$ kenarlar kümesi olmak üzere, çok etmenli sistem $G = (V, E)$ çizgesi ile temsil edilebilir. Eğer i . etmeden j . etmene bir bilgi akışı mevcut ise bu durumda $(i, j) \in E$ 'dir ve i . etmen ile j . etmen komşu etmenlerdir. Bir i etmeninin komşular kümesi $N_i = \{j : (j, i) \in E\}$ olarak tanımlanır ve i . etmeninin komşu sayısı $|N_i|$ olarak ifade edilebilir. Eğer ağda bulunan her i ve j etmeni arasında doğrudan veya dolaylı bir iletişim mevcut ise, bu ağ çizgesine kuvvetli bağlı çizge adı verilir. Her bir aydınlatma lambasının gücünün bir işlem birimi tarafından kontrol edildiği ve işlem birimleri arasında kablolu iletişimin mevcut olduğu durumda, aydınlatma lambaları etmen olarak tanımlanarak yol aydınlatma problemi çok etmenli sistemde tanımlanabilmektedir.

Ağı temsil eden çizgenin kuvvetli bağlı olduğu ve ağdaki etmenlerin bildikleri $y = Ap$ denkleminin alt kümelerinin birleşiminin $y = Ap$ denkleminde eşit olduğu durumda, keyfi başlangıç vektörlerinden başlayarak doğrusal denklem sisteminin eşsiz çözümüne ulaşabildikleri dağıtık bir algoritma Wang ve ark. tarafından

$$p_i(k+1) = P_{null(A_i)} \sum_{j \in N_i} w_{ij} p_j(k) + A_i^T (A_i A_i^T)^{-1} y_i \quad (i = 1, \dots, n) \quad (5)$$

olarak önerilmiştir [12]. Burada $p_i(k)$ değeri k . adımda i . etmenin çözüm vektörünü temsil ederken, w_{ij} parametresi j . etmenden i . etmene gelen çözüm vektörünün çarpıldığı katsayıdır ve

$$w_{ij} \begin{cases} > 0 & , \text{ eğer } (j, i) \in \mathcal{E} \text{ ise} \\ = 0 & , \text{ diğer durumlarda} \end{cases} \quad (6)$$

olarak tanımlanmaktadır. Bir i etmeninin komşu çözümlerini ağırlıklandırmak için kullandığı katsayılar

$$\sum_{j \in N_i} w_{ij} = 1 \quad (i = 1, \dots, n) \quad (7)$$

denklemini sağlanacak şekilde seçilmelidir. A_i matrisi, i . etmen tarafından bilinen $y_i = A_i p_i$ denkleminin katsayılar matrisidir. $P_{null(A_i)}$ matrisi ise A_i 'nin sıfır uzayına iz düşüm matrisidir ve

$$P_{null(A_i)} = I - A_i^T (A_i A_i^T)^{-1} A_i \quad (i = 1, \dots, n) \quad (8)$$

denklemini ile hesaplanmaktadır [22]. Denklem 5'te verilen dağıtık algoritmanın sözde kod yapısı Tablo 1'de verilmiştir.

Tablo 1 Kullanılan dağıtık algoritmanın sözde kod yapısı

```

for  $i = 1$  to  $n$ 
   $P_{null(A_i)} \leftarrow I - A_i^T (A_i A_i^T)^{-1} A_i$ 
   $p_i(0) \leftarrow n$  elemanlı rastgele bir vektör
  for each  $j \in N_i$ 
     $w_{ij} \leftarrow 1/|N_i|$ 
  end for
  for each  $j \in \{1, \dots, n\} \setminus N_i$ 
     $w_{ij} \leftarrow 0$ 
  end for
end for
for  $k = 0$  to  $k_{max}$ 
  for  $i = 1$  to  $n$ 
     $toplamlHata \leftarrow 0$ 
    for each  $j \in N_i$ 
       $p_j(k) \leftarrow$  komşu lamba  $j$ 'nin  $k$ . adımdaki çözümü
       $toplamlHata \leftarrow toplamlHata + w_{ij} p_j(k)$ 
    end for
     $p_i(k+1) \leftarrow P_{null(A_i)} \cdot toplamlHata + A_i^T (A_i A_i^T)^{-1} y_i$ 
  end for
end for

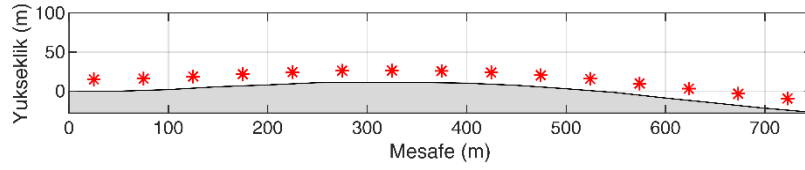
```

Yorum 2: Etmenlerin Denklem 5'te verilen algoritmayı kullanabilmesi için öncelikle $P_{null(A_i)}$ matrisini hesaplamaları gerekmektedir. Yol aydınlatma probleminin doğası gereği, $y = Ap$ denklem sisteminin her bir denklemini farklı bir fiziksel konumda ortaya çıkmaktadır ve etmenlerin bildikleri $y_i = A_i p_i$ denklemlerinde bulunan A_i matrisleri tek satırdan oluşmaktadır. Bu nedenle $A_i A_i^T$ sayıl bir değerdir ve $P_{null(A_i)}$ 'nin hesabı için yüksek işlem gücü gerekmemektedir.

3. Benzetim Çalışmaları

Bu bölümde, Bölüm 2.1'de detayları açıklanan yol aydınlatma probleminin Bölüm 2.2'deki algoritma ile çözümünü gösterebilmek için *GNU Octave* benzetim ortamında yapılan çalışmalardan bahsedilmiştir [23]. Bu amaçla, 750 metre uzunluğunda ve (otoyol aydınlatma standartlarına uygun olarak) ellişer metre aralıklarla aydınlatılmak istenen bir otoyol ele alınmıştır. Benzetim çalışmaları için kullanılacak

olan yol bölümleri doğrusal olarak modellenmiş olup, eğimleri bir önceki yol eğiminden $\pm\%3$ 'ten fazla değişmeyecek şekilde ve yol eğimi hiçbir zaman $\pm\%6$ 'yı geçmeyecek şekilde rastgele oluşturulmuştur. Yol profili Şekil 3'te ve yol bölümlerine ait eğim değerleri Tablo 2'de verilmiştir.



Şekil 3 Oluşturulan örnek yol profili

Tablo 2 Benzetim çalışmalarında kullanılan yol bölümlerinin eğim değerleri

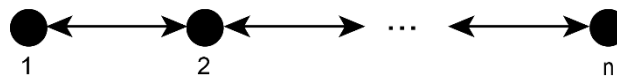
Yol bölümü	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Eğim (%)	0	0,53	-1,11	-1,80	-1,31	-2,79	-4,04	-3,34	-4,74	-2,80	0,09	1,47	0,54	1,04	-1,31

Her yol bölümü, yolun ortasına yerleştirilmiş ve standartlara uygun olarak 15 metre yükseklikteki bir aydınlatma direğinde bulunan bir lamba ile aydınlatılmak istendiği durumda, lambalar ve yol bölümlerinin orta noktaları arasındaki mesafelerin değerleri Tablo 3'teki gibi olmaktadır.

Tablo 3 Lambalar ile yol bölümlerinin orta noktaları arasındaki mesafeler (m)

Lamba	Yol bölümü														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	15,00	52,48	101,65	151,38	201,22	251,05	300,84	350,63	400,35	449,96	499,54	549,02	598,49	648,07	697,75
2	51,91	15,00	53,00	101,97	151,58	201,28	250,98	300,73	350,41	399,99	449,55	499,03	548,51	598,10	647,79
3	100,52	51,38	15,00	53,06	101,95	151,44	201,05	250,75	300,42	350,00	399,57	449,08	498,59	548,21	597,94
4	150,02	100,24	51,32	15,00	52,95	101,73	151,14	200,78	250,42	300,01	349,60	399,14	448,70	498,37	548,13
5	199,81	149,89	100,27	51,44	15,00	52,65	101,36	150,84	200,43	250,01	299,63	349,20	398,80	448,51	498,32
6	249,72	199,77	149,99	100,46	51,72	15,00	52,26	101,07	150,50	200,03	249,65	299,25	348,89	398,63	448,47
7	299,71	249,75	199,93	150,26	100,82	52,15	15,00	52,06	100,75	150,09	199,65	249,25	298,91	348,68	398,54
8	349,71	299,75	249,92	200,19	150,57	101,16	52,34	15,00	51,66	100,28	149,69	199,25	248,92	298,71	348,59
9	399,69	349,73	299,91	250,17	200,51	150,93	101,45	52,73	15,00	51,15	99,88	149,30	198,93	248,72	298,61
10	449,61	399,67	349,87	300,16	250,50	200,89	151,27	101,89	53,21	15,00	50,86	99,53	149,00	198,75	248,62
11	499,49	449,57	399,80	350,13	300,50	250,89	201,23	151,69	102,32	53,51	15,00	50,48	99,23	148,83	198,66
12	549,30	499,40	449,68	400,04	350,46	300,87	251,22	201,64	152,10	102,64	53,85	15,00	50,25	99,15	148,82
13	599,09	549,21	499,53	449,94	400,40	350,85	301,22	251,63	202,03	152,41	102,95	54,08	15,00	50,30	99,24
14	648,92	599,06	549,42	499,88	450,37	400,84	351,22	301,62	251,99	202,28	152,61	103,05	54,03	15,00	50,43
15	698,80	648,97	599,35	549,83	500,35	450,84	401,21	351,60	301,94	252,18	202,41	152,64	102,96	53,91	15,00

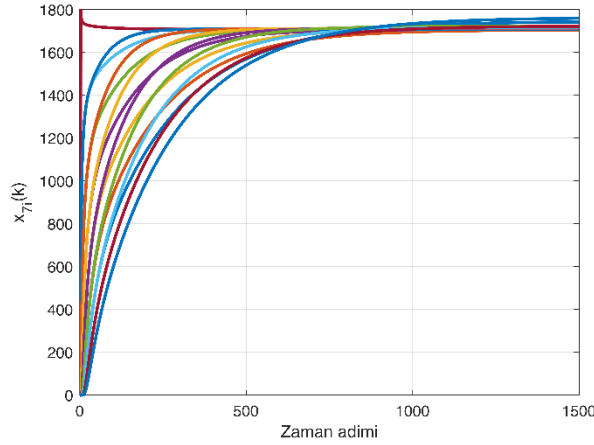
Lamba direklerine yerleştirilmiş olan, basit hesaplamalar yapabilen ve haberleşme yeteneğine sahip işlem birimleri vasıtasıyla, tüm yol bölümlerinin homojen olarak aydınlatılması amaçlanmaktadır. Benzetim çalışmalarında kablosuz haberleşme menzillerinin 50 ile 100 metre arasında olduğu varsayılmakta olup, her bir etmen yalnızca kendisine komşu olan etmenler ile haberleşebilmektedir. Bu durumda çok etmenli ağ, Şekil 4'te verilen çizge ile temsil edilebilir.



Şekil 4 Etmenlerin menzillerinin 50 ile 100 metre arası olduğu durumda etmenler arası iletişimi temsil eden çizge

Yol bölümlerinin tamamının 8 lux seviyesinde aydınlatılması istendiği durum için $y = [8, \dots, 8]^T$ olarak seçilmiştir. Lambaların yol bölümlerini doğrudan aydınlatıp aydınlatmadığı benzetim ortamında

hesaplanarak A_i matrisleri oluşturulmuş ve etmenlerin başlangıç değerleri $p_i(1) = [0, \dots, 0]^T$ olarak (yanmayan lamba) ve ağırlık katsayıları $w_{ij} = 1/|N_i|$ şeklinde seçilerek Denklem 5'te verilen algoritma kullanılarak dağıtık olarak çözüm elde edilmiştir. Örnek bir etmenin (7. etmen) çözümünün zamana bağlı değişimi Şekil 5'teki gibi olmaktadır. Şekilden de görüleceği gibi, çözüm fonksiyonu denklem sisteminin eşsiz çözümü olan ve Tablo 4'te verilen değerlere asimptotik olarak yakınsamıştır.



Şekil 5 Yedinci etmenin (lambanın), istenilen aydınlatma seviyelerine ulaşılması için gerekli lamba gücü seviyelerini dağıtık olarak hesabı

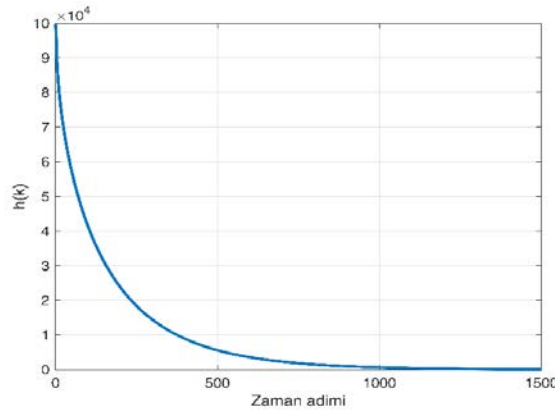
Tablo 4 Benzetim çalışmalarında incelenen yol aydınlatma probleminin eşsiz çözümü

Yol bölümü	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Lamba gücü	1746	1706	1709	1707	1710	1709	1708	1710	1712	1713	1714	1722	1719	1723	1764

Tüm etmenlerin denklem sistemini çözdüğünü göstermek amacıyla, etmenlerin çözümlerindeki hataların normlarının toplamını ifade eden

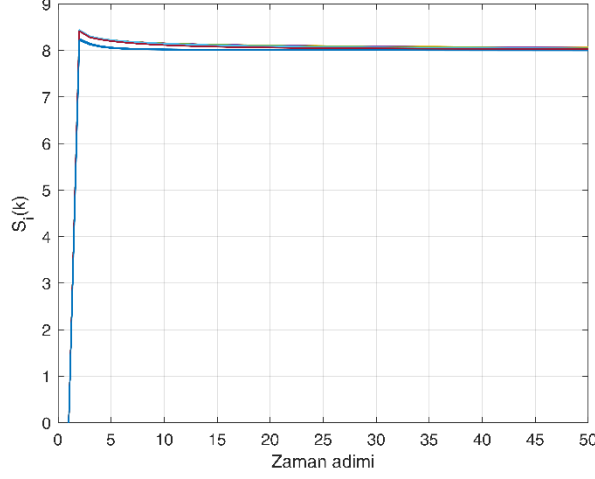
$$h(k) = \sum_{i=1}^n \|p_i(k) - p^*\| \quad (9)$$

fonksiyonunun zamanla değişimi Şekil 6'da verilmiştir. $h(k)$ 'nin asimptotik olarak sıfıra yakınsaması tüm etmenlerin denklem sistemini çözdüğünü, yani istenilen yol aydınlık seviyelerine ulaşabilmek için her bir lambanın sahip olması gereken güç değerini belirleyebildiğini göstermektedir. Her bir etmen, bağlı olduğu lambanın gücünü bu çözümden elde ettiği değere göre ayarlandığında homojen yol aydınlatması sağlanabilmektedir.



Şekil 6 Etmenlerin gerekli lamba güçleri seviyesini hesabındaki toplam hata normlarının ($h(k)$) zamanla değişimi

Etmenler denklem sistemini yinelemeli olarak çözerken, her zaman adımında elde ettikleri çözümleri kullanarak lamba güçlerini ayarladığında, yol bölümlerinin aydınlık seviyesinin zamana bağlı değişimi Şekil 7’de gösterildiği gibi olmaktadır ($S_i(k)$, i . yol bölümünün k . zaman adımındaki aydınlık seviyesini temsil etmektedir).



Şekil 7 Etmenler elde ettikleri çözüm vektörlerini kullanarak lamba güçlerini ayarladıklarında, yol aydınlık seviyelerinin zamana bağlı değişimi

Yorum 3: Şekil 7’den de görüleceği gibi istenilen aydınlık seviyesine ulaşmak için gerekli süre, Şekil 6’da gösterilen denklem sisteminin tamamının çözümü için gerekli süreden çok daha azdır. Bunun nedeni, i . etmenin denklem sistemini çözerken elde ettiği çözüm vektörünün i . elemanının (i . lambanın gücünün) gerçek çözümdeki değerine diğer elemanlardan çok daha hızlı yakınsamasıdır.

4. Sonuçlar ve Gelecek Çalışmalar

Otoyolun ellışer metre aralıklarla aydınlatıldığı ve yol eğiminin ellışer metre aralıklarla değiştiği varsayımı altında, yolun homojen olarak aydınlatılabilmesi için gerekli olan lamba güçlerinin belirlenmesi problemi doğrusal bir denklem sistemi şeklinde modellenmiştir. Her bir lambanın, bağlı oldukları ve kablosuz iletişim yeteneğine sahip basit işlem birimleri vasıtası ile güç ayarı yapabildiği sistem için, işlem birimlerinin kendilerine yakın olan diğer işlem birimleri ile bilgi paylaşımı yaparak tüm denklem sistemini (dolayısı ile gerekli olan tüm lamba güçlerini) dağıttık bir algoritma kullanarak çözebildikleri gösterilmiştir. Bulunan bu çözümler işlem birimleri tarafından kullanılarak lamba güçleri ayarlandığında, otoyolda istenilen aydınlatma seviyesine ulaşmanın mümkün olduğu yapılan benzetim çalışmaları ile doğrulanmıştır. Bulunan lamba gücü değerlerinin her lamba için farklı çıkması, homojen yol aydınlatması için lambaların farklı güçlerde aydınlatma sağlaması gerektiğini göstermektedir. Gelecek çalışmalarda, otoyollarda daha yaygın olarak kullanılan refüjden çift konsollu aydınlatma sistemi için problem modelinin geliştirilerek, uygulamaya yönelik pratik sonuçlar elde edilmesi planlanmaktadır.

Teşekkür

Bu çalışma TÜBİTAK Proje No: 117E204 tarafından desteklenmiştir.



Referanslar

- [1] H. Jin, S. Jin, L. Chen, S. Cen and K. Yuan, "Research on the Lighting Performance of LED Street Lights With Different Color Temperatures," *IEEE Photonics Journal*, vol. 7, no. 6, pp. 1-9, 2015.

- [2] C. Sun, X. Lee, I. Moreno, C. Lee, Y. Yu, T. Yang and T. Chung, "Design of LED Street Lighting Adapted for Free-Form Roads," *IEEE Photonics Journal*, vol. 9, no. 1, pp. 1-13, 2017.
- [3] W. Hackbusch, *Iterative Solution of Large Sparse Systems of Equations*, New York, USA:Springer-Verlag, 1994.
- [4] M. Krstic and A. Smyshlyaev, *Boundary Control of PDEs: A Course on Backstepping Designs*, Philadelphia, USA:SIAM Publishing, 2008.
- [5] J. Anderson, *Computational Fluid Dynamics: The Basics With Applications*, New York, USA:Mc-Graw-Hill Education, 1995.
- [6] B. Carpentieri, I. Duff, L. Giraud and M. Magolu monga Made, "Sparse Symmetric Preconditioners for Dense Linear Systems in Electromagnetism," *Numerical Linear Algebra with Applications*, vol. 11, pp. 753–771, 2004.
- [7] F. Pasqualetti, R. Carli and F. Bullo, "Distributed Estimation via Iterative Projections With Application to Power Network Monitoring," *Automatica*, vol. 48, no. 5, pp. 747-758, 2012.
- [8] D. Silvestre, J. Hespanha and C. Silvestre, "A Pagerank Algorithm Based on Asynchronous Gauss–Seidel Iterations," *Proc. - 2018 American Control Conference*, pp. 484–489, 2018.
- [9] S. Mou, J. Liu and A.S. Morse, "A Distributed Algorithm for Solving a Linear Algebraic Equation," *IEEE Transactions on Automatic Control*, vol. 60, no. 11, pp. 2863–2878, 2015.
- [10] J. Liu, S. Mou and A.S. Morse, "Asynchronous Distributed Algorithms for Solving Linear Algebraic Equations," *IEEE Transactions on Automatic Control*, vol. 63, no. 2, pp. 372–385, 2018.
- [11] J. Liu, A.S. Morse, A. Nedic and T. Başar, "Exponential Convergence of a Distributed Algorithm for Solving Linear Algebraic Equations," *Automatica*, vol. 83, pp. 37–46, 2017.
- [12] X. Wang, J. Zhou, S. Mou and M.J. Corless, "A Distributed Linear Equation Solver for Least Square Solutions," *Proc. - 56th IEEE Conference on Decision and Control*, pp. 5955–5960, 2017.
- [13] J. Liu, X. Gao and T. Başar, "A Communication-Efficient Distributed Algorithm for Solving Linear Algebraic Equations," *Proc. - 7th International Conference on Network Games, Control and Optimization*, pp. 62–69, 2014.
- [14] P. Wang, W. Ren and Z. Duan, "Distributed Algorithm to Solve a System of Linear Equations With Unique or Multiple Solutions From Arbitrary Initializations," *IEEE Transactions on Control of Network Systems*, vol. 6, no. 1, pp. 82-93, 2018.
- [15] X. Wang, S. Mou and D. Sun, "Improvement of a Distributed Algorithm for Solving Linear Equations," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 4, pp. 3113–3117, 2017.
- [16] S. Mou, Z. Lin, L. Wang, D. Fullmer and A.S. Morse, "A Distributed Algorithm for Efficiently

- Solving Linear Equations and its Applications," *Systems & Control Letters*, vol. 91, pp. 21–27, 2016.
- [17] B.D.O. Anderson, S. Mou, A.S. Morse and U. Helmke, "Decentralized Gradient Algorithm for Solution of a Linear Equation," *Numerical Algebra, Control & Optimization*, vol. 6, no. 3, pp. 319-328, 2016.
- [18] G. Shi, B.D.O. Anderson and U. Helmke, "Network Flows That Solve Linear Equations," *IEEE Transactions on Automatic Control*, vol. 62, no. 6, pp. 2659-2674, 2017.
- [19] M. Yang and C.Y. Tang, "A Distributed Algorithm for Solving General Linear Equations Over Networks," *Proc. - 54th IEEE Conference on Decision and Control*, pp. 3580–3585, 2015.
- [20] J. Liu, X. Chen, T. Başar ve A. Nedic , "A Continuous-Time Distributed Algorithm for Solving Linear Equations," *Proc. - 2016 American Control Conference*, pp. 5551–5556, 2016.
- [21] Y. Liu, C. Lageman, B.D.O. Anderson and G. Shi, "An Arrow–Hurwicz–Uzawa Type Flow as Least Squares Solver for Network Linear Equations," *Automatica*, vol. 100, pp. 187-193, 2019.
- [22] P. Wang, S. Mou, J. Lian and W. Ren, "Solving a System of Linear Equations: From Centralized to Distributed Algorithms," *Annual Reviews in Control*, vol. 47, pp. 306-322, 2019.
- [23] GNU Octave, "GNU Octave," 2020. [Online]. Available: <https://www.gnu.org/software/octave/>. [Accessed: 15-May-2020].

Detection of Cervix Cancer from Pap-smear Images

 Fatma Betül AKYOL¹,  Oğuz ALTUN²

¹Corresponding Author; Yıldız Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği;
<https://orcid.org/0000-0002-3836-1317>; akyolfbetul@gmail.com; +90 545 616 16 00

²Yıldız Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği;
<https://orcid.org/0000-0001-8823-3148>; oaltun@yildiz.edu.tr

Received 18 April 2020; Revised 07 May 2020; Accepted 21 May 2020; Published online 28 August 2020

Abstract

Pap-smear test is used to detect cervical cancer, which ranks fourth in the ranking of cancer diseases in women worldwide. In this study, it is aimed to design a computer based decision system that can detect cervical cancer at an early stage. Normal and abnormal cells are found in the cervix images obtained as a result of the pap-smear test and the abnormal cells are marked on the image. The features extracted from the images were examined with pathologists and a dataset was created. For each of the 917 images in the Herlev dataset, these features were extracted and stored in a dataset. Support Vector Machines (SVM), Naive Bayes, Random Forest (RF), Multilayer Perceptron (MLP), Logistic Regression (LR), K- Nearest Neighbor (KNN) methods were applied to the created dataset, and accuracy values between 83% and 92% were obtained.

Keywords: Cervical cancer, Pap-smear, Segmentation, Classification

Pap-Smear Görüntülerinden Rahim Ağzı Kanseri Tespiti

Öz

Dünya çapında kadınlarda görülen kanser hastalıkları sıralamasında dördüncü sırada yer alan rahim ağzı kanserini erken aşamada tespit edebilmek için pap-smear testi kullanılmaktadır. Bu çalışma kapsamında rahim ağzı kanserini erken aşamada tespit edebilecek bilgisayar temelli bir karar sisteminin tasarlanması amaçlanmıştır. Pap-smear testi sonucunda elde edilmiş olan serviks görüntülerinde normal ve anormal özellikli hücreler bulunarak anormal olan hücreler görüntü üzerinde işaretlenmiştir. Görüntülerden çıkarılan özellikler patoloji uzmanları ile incelenmiş ve bir veri seti oluşturulmuştur. Herlev veri setinde bulunan 917 görüntünün her biri için bu özellikler çıkarılmış ve bir veri setine kaydedilmiştir. Oluşturulan veri setine makine öğrenmesi yöntemlerinden Support Vector Machines (SVM), Naive Bayes, Random Forest (RF), Multilayer Perceptron (MLP), Logistic Regression (LR), K- Nearest Neighbor (KNN) yöntemleri uygulanmıştır ve %83 ve %92 arasında doğruluk değerleri elde edilmiştir.

Anahtar Kelimeler Rahim ağzı kanseri, Pap-smear, Segmentasyon, Sınıflandırma

1. Giriş

Kanser, vücuttaki hücrelerin anormal büyümeleri nedeniyle meydana gelen ve günümüzde 100'den fazla çeşidi bulunan bir hastalık türüdür [1]. Bu kanser çeşitlerinden biri olan rahim ağzı kanseri kadınlarda en sık görülen dördüncü kanser türüdür. 2018 yılında dünya çapında yaklaşık olarak 570.000 kadında rahim ağzı kanseri görülmüş ve 311.365 kadın bu hastalık nedeniyle hayatını kaybetmiştir [2]. Servikal intraepitelyal neoplazi (SIN) anormallikleri ile ilişkilendirilen İnsan Papilloma Virüsü (IPV) rahim ağzı kanseri vakalarının %50'sinde hastalık kaynağı olarak gözlenmiştir [3]. Bethesda Sistemi SIN'i, düşük dereceli skuamöz intraepitelyal lezyon (DSIL), yüksek dereceli skuamöz intraepitelyal lezyon (YSIL) ve skuamöz hücre karsinoması olmak üzere üç gruba ayırır. Anormal görünümde olan fakat kanser hücresi olmayan kanser öncüsü olarak görülen hücrelere displazi denir. Bethesda Sistemi'ne göre yukarıda belirtilen hücre anormalliklerinden, DSIL hafif displazi, YSIL ise orta displazi olarak sınıflandırılabilir [4].

Hastalığın teşhisinde farklı yöntemler mevcuttur. Mevcut yöntemler arasında en yaygın olarak kullanılan tekniklerden biri görüntülemedir. Günümüzde rahim ağzı kanseri görüntüleme yöntemi olarak “Pap-Smear Testi” yaygın olarak kullanılmaktadır. [5].

Kanser hastalığında ve kanser kaynaklı ölümlerde meydana gelen artış bu hastalık üzerinde yoğun çalışmalar yapılmasını beraberinde getirmiştir. Yapay zekâ ile birlikte bilgisayar temelli karar sistemleri tasarlanmaya başlanmıştır ve bu karar sistemleri günümüzde tıp dâhil birçok alanda kullanılmaktadır. Makine öğrenmesi yöntemlerini kullanarak tıbbi görüntü analizi alanında yapılan çalışmalar yaygın şekilde devam etmektedir. Literatürde bu alanda yapılan birçok çalışma bulunmaktadır.

Taha vd. [6], pap-smear görüntülerinin analizinde özellik çıkarımı için Evrişimli Sinir Ağları (ESA), sınıflandırma için ise SVM yöntemini kullandılar.

Rahim ağzı kanseri teşhisinde serviks hücrelerinin çekirdek ve sitoplazmaları ayrı olarak ele alınmıştır. Bu nedenle hücreye çekirdek ve sitoplazma segmentasyonu işlemleri uygulanmıştır. Segmentasyon alanında yapılmış olan çalışmalardan, Plissiti vd. [7], pap-smear görüntülerinde hücre çekirdeği ve sitoplazmasını otomatik olarak sınıflandırmak için bulanık c-ortalama (fuzzy c-means) ve SVM tekniklerini kullandılar. Çalışmalarında hücre çekirdeklerini %99,25 doğruluk oranı ile sınıflandırdılar.

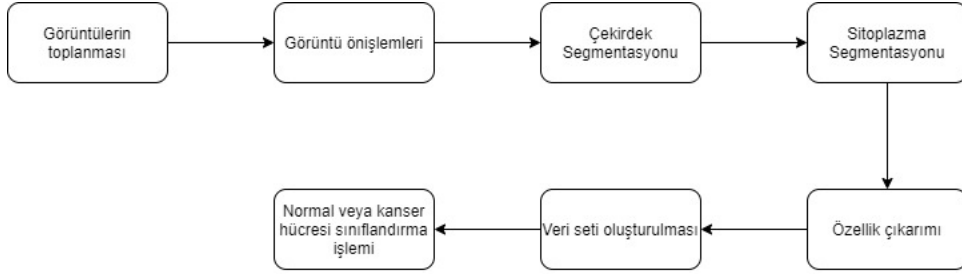
Wang vd. [8], hücre çekirdeğini bulma ve segmentasyonu için matematiksel morfolojiye sahip ortalama kaydırma kümeleme algoritmasını ve sınıflandırıcı olarak SVM yöntemini kullanmışlardır. Diğer bir çalışmada, Marinakis vd. [9], çalışmalarında özellik seçimi için genetik algoritmaları, sınıflandırma için KNN yöntemini kullanmışlardır. Genctav vd. [10], pap-smear görüntüleri üzerinde segmentasyon için watershed algoritmasını ve hücreleri sınıflandırmada ise denetimsiz öğrenme algoritmalarını kullanmışlardır. Phoulady vd. [11], genişletilmiş alan derinliği algoritmasını kullanarak serviks hücrelerinden çekirdek segmentasyonu işlemi yapmışlardır. %96 hassasiyet ile diğer metotlardan daha iyi bir sonuç elde etmişlerdir. Diğer bir çalışmada Win vd. [12] pap-smear görüntüleri üzerinde rahim ağzı kanseri tespiti için sınıflandırıcı olarak Linear Discriminant (LD), SVM, KNN, boosted trees (yükseltilmiş ağaçlar) ve bagged trees (torbalı ağaçlar) yöntemlerinin sonuçlarının kombinasyonu olan bir yöntem kullandılar ve bu yöntem ile iki sınıflı sınıflandırma işleminde %98,27 doğruluk değeri elde ettiler. William vd. [13] çalışmalarında pap-smear görüntülerinden serviks kanseri tespiti için bir araç geliştirme üzerine çalışma yapmışlardır. Çalışmalarında sınıflandırıcı olarak bulanık c-ortalama (fuzzy c-means) yöntemini kullanmışlardır ve %98,8 doğruluk değeri elde etmişlerdir.

Bu çalışma kapsamında pap-smear testi sonucunda elde edilmiş olan görüntülerde ilk olarak serviks hücrelerinin çekirdek ve sitoplazmaları ayrı olarak bulunmuştur. Çekirdek segmentasyonu için iki ayrı yöntem uygulanmıştır. Bu yöntemlerden ilki eşikleme (thresholding) yöntemi iken ikincisi K-ortalama kümeleme yöntemidir. Sitoplazma segmentasyonu için ise eşikleme yöntemi ve morfolojik işlemler kullanılmıştır. Segmentasyon işlemi ile görüntüde bulunan her bir çekirdek ve sitoplazmadan bazı özellikler çıkarılmıştır. Bu özelliklere hücrenin normal veya anormal olarak sınıflandırılmasında ihtiyaç duyulmuştur. Anormal olarak bulunan hücreler sınıflandırılmıştır ve görüntü hücre işaretlenmiştir. Görüntülerden çıkarılan özellikler bir veri setine kaydedilmiş olup patoloji uzmanları ile incelenmiştir. Çıkarılan özellikler bir pap-smear görüntü veri seti ile denenmiştir ve bir veri seti oluşturulmuştur. Oluşturulan veri seti SVM [14], RF [15], MLP [16], KNN [17], NaiveBayes [18] ve LR [19] yöntemleri kullanılarak normal ve anormal hücre sınıflandırılmasında kullanılmıştır.

Kullanılan materyal ve yöntemler 2. Kısımda anlatılmış olup, Sonuçlar 3. Kısımda belirtilmiştir.

2. Materyal ve Metotlar

Çalışmada verilerin elde edilmesi işleminden sınıflandırma işlemine kadar takip edilen adımlar Şekil 1’de gösterilmiştir.



Şekil 1 Çalışmada takip edilen adımlar

Çalışma segmentasyon, özellik çıkarımı ve sınıflandırma olmak üzere üç temel adımdan oluşmaktadır. Görüntüleri segmentasyon işlemine hazır hale getirmek için görüntü üzerinde bazı ön işlemler uygulanmıştır.

2.1 Veri Seti

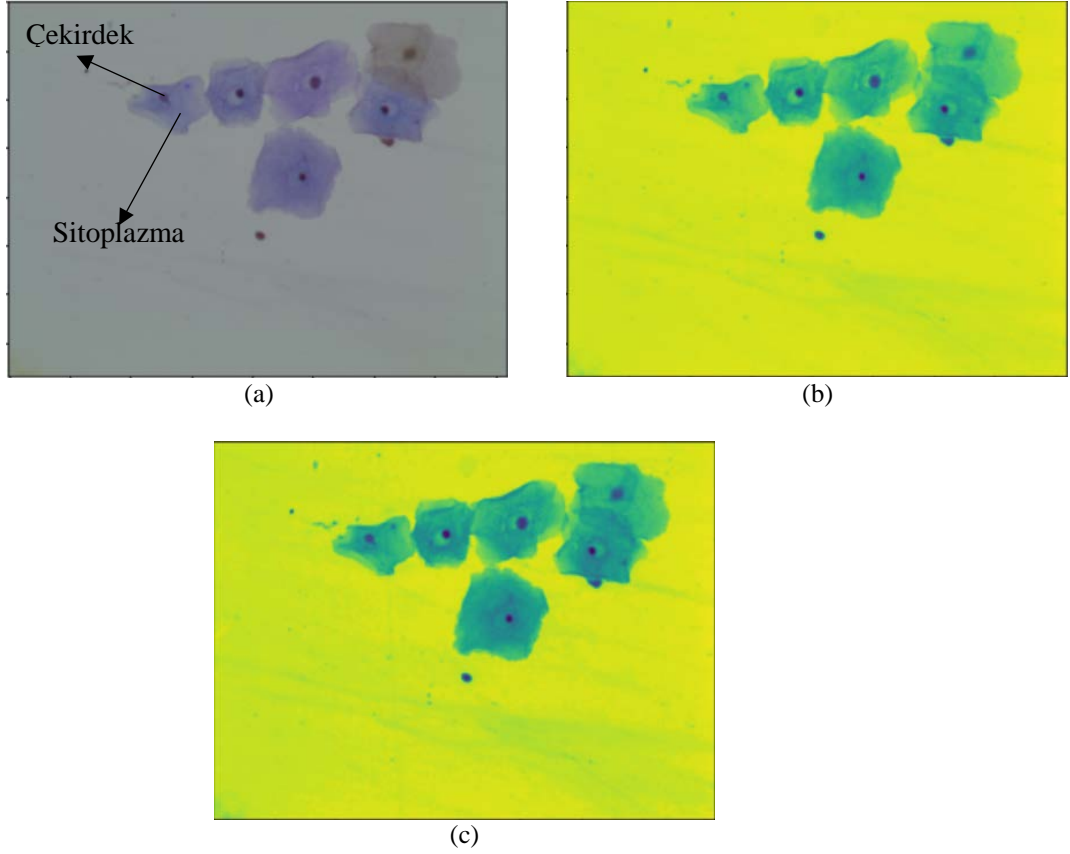
Bu çalışmada pap- smear görüntüleri kullanılmıştır. Çalışmada ele alınan yöntemler Herlev veri seti [20] üzerinde uygulanmıştır. Herlev Üniversitesi Hastanesi tarafından hazırlanmış olan Herlev veri seti 4 tanesi anormal ve 3 tanesi normal olarak belirlenmiş 7 sınıfa ait 917 pap-smear görüntüsünden oluşmaktadır. Anormal olarak belirlenen kategorilerden *carcinoma_in_situ*'ya ait 150, *light_dysplastic*'te 182, *moderate_dysplastic*'te 146 ve *severe_dysplastic*'te 192 pap-smear görüntüsü bulunmaktadır. Normal olarak sınıflandırılan kategorilerden *normal_superficial*'da 74, *normal_intermediate*'ta 70 ve *normal_columnar*'da 98 adet görüntü bulunmaktadır [21]. Görüntüler üzerinde sırasıyla ön işlemler, segmentasyon ve sınıflandırma işlemleri uygulanmıştır.

2.2 Veri Ön İşleme

Rahim ağzı kanseri tespitinde hücre çekirdeğinin önemli bir yeri vardır. Bu nedenle ilk olarak görüntü üzerinde hücre çekirdeği kısmı incelenmiştir. Çekirdeği tek olarak ele alabilmek için sitoplazma kısmı görüntüden çıkarılmıştır. Hücre çekirdeğini sitoplazmadan ayırabilmek için hücreye segmentasyon işlemi uygulanmıştır.

Bu çalışmada veri olarak pap- smear testi sonucu elde edilen serviks görüntüleri kullanılmıştır. Segmentasyon işlemine hazırlık için görüntüler bazı ön işlemlerden geçirilmiştir. İyileştirme işleminde amaç görüntü üzerinde bulunan gürültünün temizlenmesi ve kontrastlığın artırılmasıdır. Kontrastlığın artırılma amacı ise çekirdek ve sitoplazma arasındaki koyuluğun artırılarak çekirdeğin daha koyu sitoplazmanın daha açık hale getirilmesidir. Gürültüyü gidermek için boyutu 5 olan bir medyan filtre uygulanmıştır. Görüntüyü iyileştirmek için ise Ayrık Kosinüs dönüşümü (AKD) ve Ayrık Dalgacık Dönüşümü (ADD) yöntemleri uygulanmıştır.

Ayrık kosinüs dönüşümü görüntü sıkıştırma alanında kullanılan ve bu alanda yüksek performans gösteren bir yöntemdir [22]. AKD yönteminde amaç görüntüyü uzaysal alandan frekans alanına çevirerek iyileştirmektir. AKD yönteminde her pikselin grilik seviyelerinin tutulduğu 8×8 ' lik bir piksel matrisi oluşturulmuştur ve bu matrise AKD yöntemi uygulanmıştır. Girdi olarak verilen görüntüyü yeniden elde etmek için ters ayrık kosinüs dönüşümü uygulanmıştır. SIPAKMED [23] veri setinden alınan Şekil 2 (a)' da gösterilen giriş görüntüsü için AKD işlemi sonucu elde edilen çıkış görüntüsü Şekil 2 (b)'de gösterilmiştir.



Şekil 2 (a) Giriş görüntüsü (b) AKD sonrası görüntü (c) ADD sonrası görüntü

ADD, giriş olarak verilen görüntüyü ilk olarak L ve H olmak üzere 2 banda böler. 2 boyutlu ADD ise görüntüyü LL (low-low), LH (low-high), HL (high-low), HH (high-high) olmak üzere dört banda böler [24]. LH, HL ve HH bantları görüntünün yüksek frekansa sahip alanlarını içerirken LL ise düşük frekanslı alanını içerir [25].

Çekirdek ve sitoplazmayı birbirinden daha iyi ayırabilmek için bu iki görüntü arasındaki zıtlık oranı, çekirdek daha koyu, sitoplazma daha açık hale getirilerek artırılmıştır. ADD işlemi sonucu elde edilen görüntü Şekil 2 (c)'de de gösterilmiştir.

2.3. Çekirdek Segmentasyonu

Bir serviks hücrelerini normal veya anormal olarak sınıflandırabilmek için hücrenin çekirdeği ve sitoplazması ayrı olarak ele alınmıştır. Bu iki yapıyı birbirinden ayırabilmek için segmentasyon işlemi uygulanmıştır.

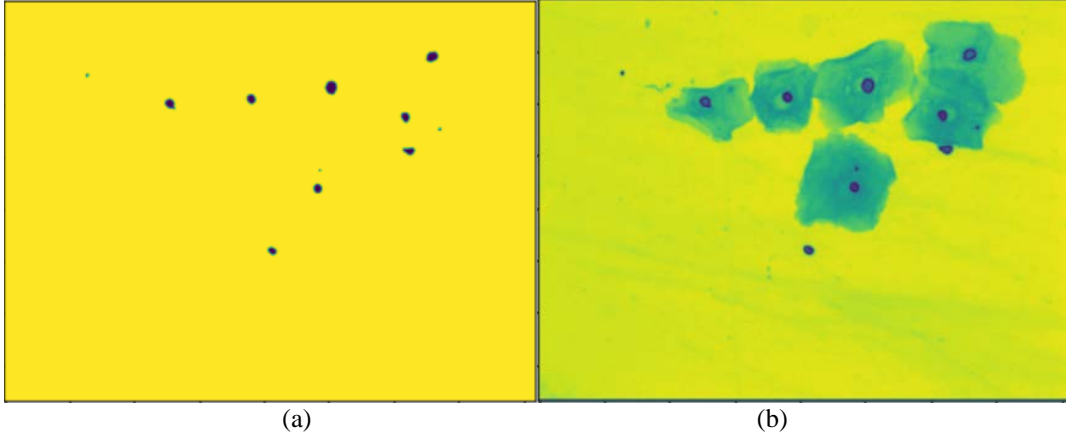
Şekil 2 (c)'de gösterilen ayırık dalgacık dönüşümü ile iyileştirilen görüntüde hücre çekirdeklerinin yaygın olarak mavi renk spektrumunda olduğu görülmektedir. Bu nedenle hücre çekirdeklerini tespit edebilmek için giriş görüntüsünde mavi renkli alanların bulunması gerekmiştir. Bu alanları bulabilmek için ilk olarak RGB renk uzayında bulunan giriş görüntüsü HSV (Hue, Saturation, Value) renk uzayına dönüştürülmüştür. Daha sonra mavi renkli alanları filtrelemek için Denklem 1'de a ile belirtilen mavi renk alt sınırı Denklem 2'de t ile gösterilen mavi renk üst sınırı belirlenmiştir ve görüntüde bu aralıklarda bir maske oluşturulmuştur. Alt ve üst sınır üç değerden oluşmaktadır Her bir değer HSV renk uzayının üç kanalına (h, s ve v) denk gelmektedir. H değeri rengi, s değeri rengin doygunluğunu ve v değeri rengin parlaklığını ifade eder. Renk uzayında bulunan tüm renkler belirli bir alt ve üst sınır renk değerleri arasında bulunmaktadır. a ve t dizilerinin ilk elemanları (0-120) renk uzayında mavi rengin alt(0) ve üst (120) renk değerini, ikinci elemanları rengin canlılık değerinin alt ve üst sınırını ve üçüncü elemanlar ise mavi renginin belirlenen parlaklık alt ve üst sınırlarıdır. Bu çalışmada mavi renkli bölgeler filtreleneceği için renk uzayında mavi renk alt ve üst sınırı [0,120] arası olarak belirlenmiştir. RGB renk

uzayında mavi renk $[0,0,255]$ şeklinde ifade edilmektedir. $[0,0,255]$ değeri HSV renk uzayına çevrilmiştir ve h değeri olarak 0 elde edilmiştir ve bu değer alt sınır olarak kullanılmıştır. Oluşturulan maske ile görüntüye Denklem 3'te belirtilen bit bazında mantıksal işlemler uygulanarak elde edilen m maske görüntüsü Şekil 3 (a)'da gösterildiği gibi maskelenmiştir.

$$a = [0, 0, 55] \quad (1)$$

$$t = [120, 255, 255] \quad (2)$$

$$g = f \& m \quad (3)$$



Şekil 3 (a) Maskeleye işlemi çıktısı (b) Hücre çekirdeklerinin giriş görüntüsü üzerinde işaretlenmesi

Maskeleye işlemi ile bulunan Şekil 2 (c) ile gösterilen f görüntüsü için sırasıyla morfolojik açma ve yayma işlemleri uygulanmıştır.

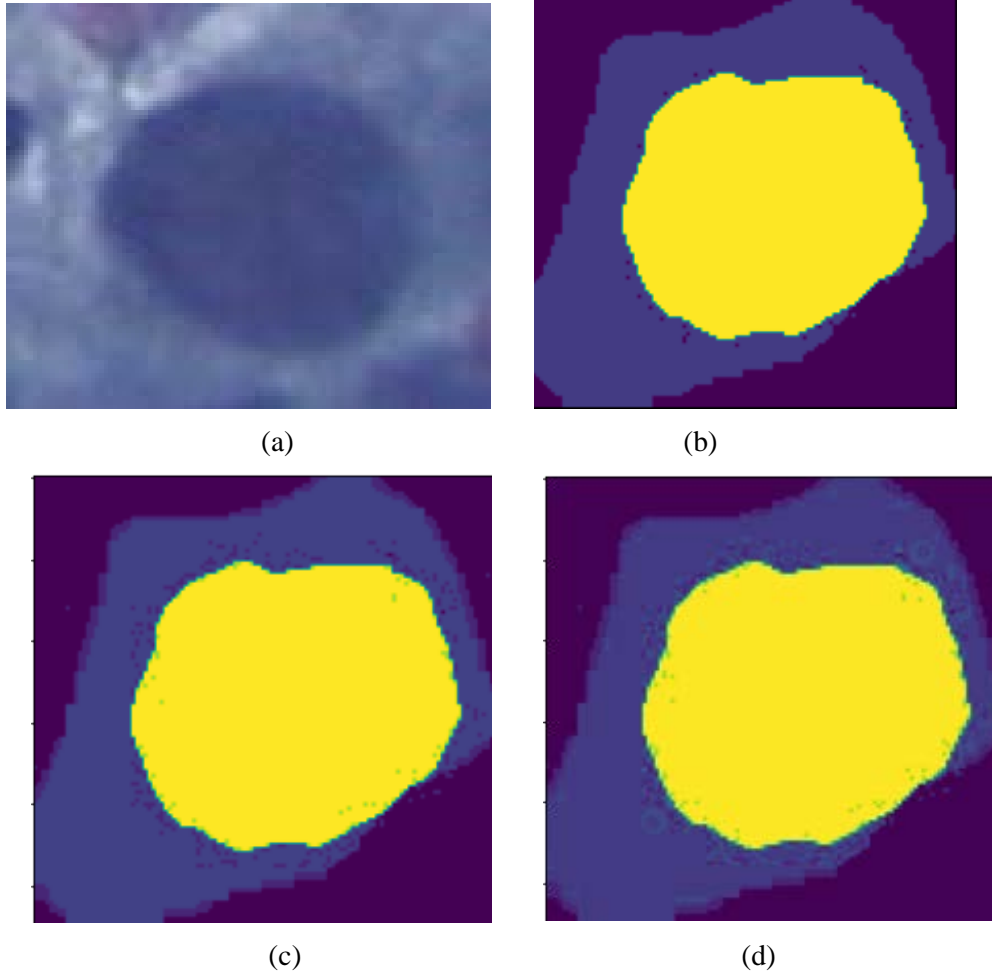
$$m = f \circ s \quad (4)$$

$$g = f \ominus s \quad (5)$$

Denklem 4'te HSV renk uzayında maske uygulanan f imgesi için s yapı elemanı ile uygulanan morfolojik açma işlemi belirtilmiştir. Denklem 5'te ise açma işlemi ve s yapı elemanı ile uygulanan morfolojik yayma işlemi verilmiştir.

Yayma işlemi sonucu elde edilen g görüntüsü Şekil 3 (a)'da gösterilmiştir. Elde edilen çekirdek maskeleri f giriş görüntüsü üzerinde Şekil 3(b)'de gösterildiği gibi işaretlenmiştir.

Çekirdek segmentasyonu işleminde kullanılan diğer bir yöntem ise K- ortalama (K-means) olmuştur. Bu yöntemde $k = 3$ alınmıştır ve Şekil 4 (a)'da verilen giriş görüntüsü için k-ortalama işlemi sonucu elde edilen görüntü Şekil 4 (b)'de gösterilmiştir. Farklı K değeri değiştirildiği zaman görüntü üzerinde yapılan kümeleme işlemi bazı noktaların yanlış sınıflara kümelendiği görülmüştür. Şekil 4 (c)'de görüldüğü gibi K-ortalama yönteminde $k = 5$ değeri kullanıldığında çekirdek içerisinde bulunan bazı noktaların yanlış kümelere atandığı gözlemlenmiştir. K değeri 7 olarak alındığında ise Şekil 4 (d)'de görüldüğü gibi kümeleme işlemi sonucu elde edilen görüntünün bozulduğu ve yanlış kümelere atanan noktaların sayısının arttığı gözlemlenmiştir.



Şekil 4 (a) Giriş görüntüsü (b) $k = 3$ için K-means sonucu (c) $k = 5$ için K-means işlemi sonucu (d) $k = 7$ için K-means işlemi sonucu

2.4. Sitoplazma Segmentasyonu

AKD ve ADD yöntemleri ile iyileştirilen smear görüntülerinde hücre ve sitoplazma arasındaki kontrast artırılmıştır. Çekirdek daha koyu, sitoplazma daha açık hale getirilerek segmentasyon aşamasında bu iki yapının birbirinden daha kolay ayrılabilmesi sağlanmıştır.

Hücrenin sitoplazma bölümlerini bulmak için uygulanan adımlar Denklem 6' da gösterilmiştir.

$$S_{stp} = G - f \quad (6)$$

Kontrastlığı artırılmış G giriş görüntüsünden eşikleme yöntemi kullanılarak bulunan çekirdek kısımları çıkarılmıştır. Kalan bölümlere OTSU eşikleme yöntemi ve Denklem 4 ve 5'te belirtilen morfolojik işlemler uygulanmıştır. Yapılan işlemler sonucu elde edilen görüntü Şekil 5'te gösterilmiştir. Görüntüde sitoplazma alanları beyaz olarak görünmektedir ve her biri kontur olarak kaydedilip özellik çıkarımı alanında belirtilen sınıflandırma kriterleri hesaplanmıştır.



Şekil 5 Segmentasyon işlemi sonucu sitoplazma

2.5. Özellik Çıkarımı

Bir serviks hücrelerini anormal veya normal olarak sınıflandırabilmek için görüntüden çıkarılan bazı özellikler sınıflandırma kriteri olarak kullanılmıştır. Anormal hücrelerde genellikle; konturda düzensizlik ve çekirdekte genişleme gözlemlenir. Koyu renk çekirdeğe sahip olup yüksek çekirdek-sitoplazma oranına sahiptirler [26].

Yukarıda belirtilen bilgilere dayanarak görüntü üzerinde aşağıdaki özellikler hesaplanmış ve sınıflandırma kriteri olarak belirlenmiştir:

Çekirdek alanı: Çekirdek olarak belirlenen her bir konturun alanı

Sitoplazma alanı: Sitoplazma olarak bulunan her konturun alanı

Çekirdek koyuluk değeri: Çekirdek konturlarının grilik seviyeleri

Çekirdek çevresi: Çekirdek konturlarının çevre hesabı

Sitoplazma çevresi: Sitoplazmaların hesaplanan çevre değeri

Çekirdeğin merkez koordinatları: Her bir çekirdeğin merkezinin bulunduğu (x, y) koordinatları

Sitoplazmanın merkez koordinatları: Her bir sitoplazmanın merkezinin bulunduğu (x, y) koordinatları

Çekirdek/ Sitoplazma (Ç/S) oranı: Hücrede bulunan çekirdek alanının sitoplazma alanına oranı

Çekirdek sitoplazma arası mesafe: Her bir hücre çekirdeğinin o hücrenin sitoplazması ile arasındaki uzaklık

Çekirdek yuvarlaklığı: Çekirdek konturunun yuvarlaklığının hesaplanması Denklem 7 ile gösterilmiştir.

$$Rnd = \frac{4 \cdot \pi \cdot alan}{cevre^2} \quad (7)$$

Sitoplazma yuvarlaklığı: Sitoplazmanın yuvarlaklık değerinin Denklem 7'ye göre hesaplanan değeri

Her hücrede bulunan çekirdek sayısı: Bazı anormal hücreler birden fazla çekirdekte oluşabildiği için sınıflandırma işleminde her bir hücrede kaç tane çekirdek bulunduğu hesaplanmıştır.

2.6. Sınıflandırma ve Değerlendirme

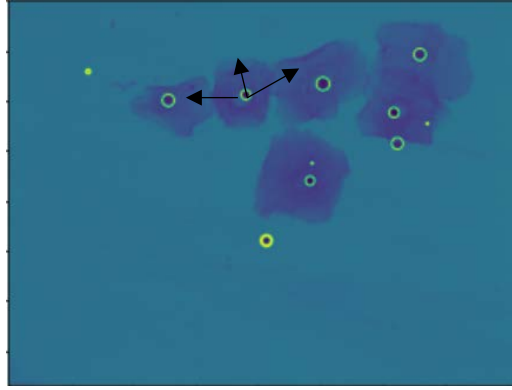
Çekirdek ve sitoplazma alanının hesaplanması çekirdek genişlemesi durumunun ölçülebilmesine olanak sağladığı için hücreyi anormal olarak sınıflandırmada büyük önem taşır. Bu nedenle her bir sitoplazma ve çekirdek konturu için görüntüden alan özelliği çıkarılmıştır. Elde edilen alan bilgilerinden önemli bir teşhis kriteri olan Ç/S oranı hesaplanmıştır. Segmentasyon işlemi sonucunda görüntüde birden fazla çekirdek ve sitoplazma konturları bulunduğu için her çekirdeğin içerisinde bulunduğu sitoplazması ile

eşleştirilmesi gerekmiştir. Bu eşleştirme işleminin doğru yapılmaması sonucunda yanlış Ç/S oranı hesaplanacağı için eşleştirme işlemi doğru oran bulmada önemli yer tutmaktadır. Eşleştirme işlemi aşamasında her bir çekirdek ve sitoplazma için çıkarılan merkez koordinatları kullanılmıştır.

Her bir çekirdek için (x_n, y_n) ikilisi ve her sitoplazma için (x_s, y_s) ikilisi konturun merkezinin bulunduğu x ve y koordinatlarını belirtmektedir. Her sitoplazma içinde çekirdek bulduđu için sitoplazmanın merkez koordinatları ile içinde bulduđu çekirdeğin merkez koordinatları arasında bulunan mesafenin az olması gerektiđi düşünölmüştür. Bu iki nokta arasındaki mesafeyi ölçmek için Denklem 8’te gösterilen Euclidean mesafesi yöntemi kullanılmıştır.

$$d = \sqrt{(x_n - x_s)^2 + (y_n - y_s)^2} \quad (8)$$

Şekil 6’da verilen referans görüntüsü için Denklem 8’ te belirtilen d uzaklık değeri 1.00 değeri eşit veya daha küçük olması doğru çekirdek- sitoplazma eşleştirme sonucu verdiđi gözlemlenmiştir. d uzaklık değeri olarak hesaplama işleminde doğru eşleştirme sonucunu veren minimum değer kabul edilmiştir. Minimum değer alınmasının nedeni ise d mesafesinin yüksek bir değer alınması durumunda Şekil 6’da üzerinde örnek olarak alınan bir çekirdek oklar ile gösterilen üç sitoplazma ile de eşleştirilebilir ve bu durum yanlış çekirdek-sitoplazma eşleştirme sonucunu ortaya çıkarmaktadır. Sitoplazması ile eşleştirilen her bir çekirdek yuvarlak içine alınmıştır.

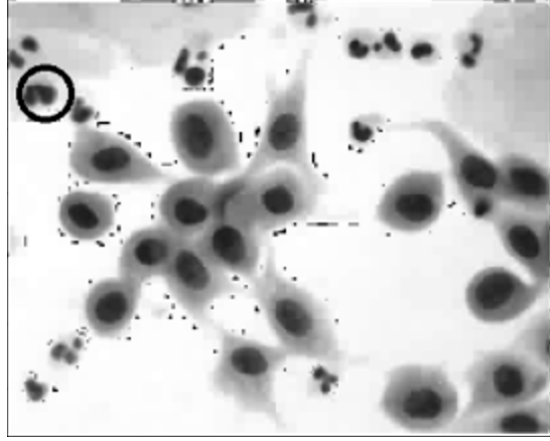


Şekil 6 Çekirdek-Sitoplazma eşleştirme

Eşleştirilen her bir çekirdek ve sitoplazma için Ç/S oranı hesaplanmıştır ve veri setine kaydedilmiştir. Veri setinde mesafesi 1.2 ve altında olan hücrelerde Ç/S oranı olarak hesaplanan değer, diğer durumlarda 0 olarak alınmıştır.

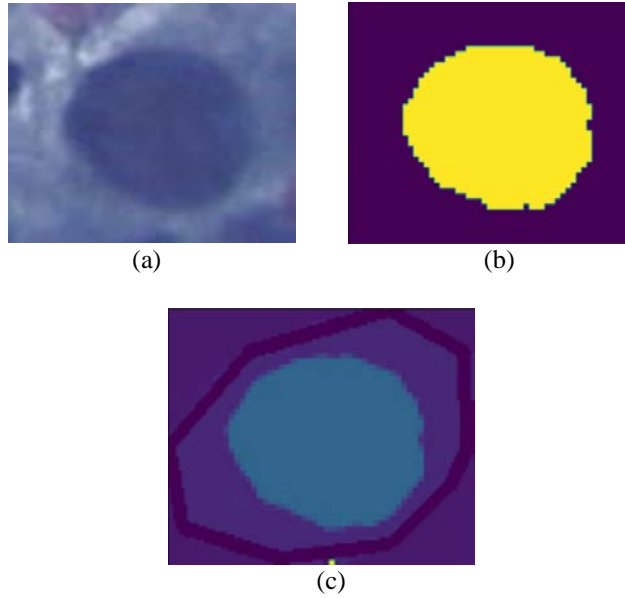
Koyu renk çekirdeđe sahip olan hücreleri bulmak için her bir çekirdek konturunun ortalama koyuluk değeri hesaplanmış ve ilgili çekirdek ile ilgili verilere kaydedilmiştir. Daha sonra veri setinde yoğunluk özelliğinin altında bulunan tüm değerlerin ortalaması bulunmuştur. Bir çekirdeğin konturu hesaplanan ortalama değeri büyükse hücre koyu renk çekirdeđe sahip olduđu düşünölmüştür.

Anormal hücre sınıflandırma kriterlerinden biri olan sitoplazmada bulunan hücre sayısı hesaplanmıştır. Eğer bir sitoplazmada birden fazla çekirdek bulunuyorsa bu hücreler anormal olarak değerlendirilmiştir. Bu hesaplama işleminde çekirdek konturlarının merkezlerinin arasındaki mesafe bulunmuştur. Mesafe hesaplama işleminde her (x_n, y_n) çekirdek merkez koordinatları için Denklem 8 ile verilen mesafe bulunmuştur ve bu d mesafesinin 5.00’in altında olması durumunda sitoplazma içerisinde bulunan çekirdek sayısı artırılmıştır. Şekil 7’de verilen görüntü [27] için birden fazla çekirdeđe sahip olan hücre bulunmuştur ve görüntü üzerinde yuvarlak içerisine alınmıştır.



Şekil 7 İki çekirdeğe sahip hücrenin görüntü üzerinde görünümü

Makine öğrenmesi yöntemleri ile başarı oranı yüksek bir sınıflandırma yapabilmek için veri sayısının fazla olması gerekmektedir. Bu nedenle Herlev veri seti kullanılarak bu veri setinde bulunan 917 görüntü için ilk olarak çekirdek ve sitoplazma segmentasyon işlemleri yapılmıştır. Şekil 8 (a)'da belirtilen görüntü için çekirdek ve sitoplazma segmentasyonu işlemlerinin sonucu Şekil 8 (b) ve (c)'de gösterilmiştir.



Şekil 8 (a) Giriş görüntüsü (b) Çekirdek segmentasyonu (c) Sitoplazma segmentasyonu

Çekirdek ve sitoplazma segmentasyonu işlemlerinden sonra sınıflandırma kriteri olarak belirlenmiş özellikler veri setinde bulunan 917 görüntü için bulunmuştur. Görüntülerden çıkarılan özellikler ile oluşturulan veri setinden normal ve anormal hücrelere ait bazı veriler Tablo 1'de gösterilmiştir. Herlev veri setinde bulunan 7 kategori bu çalışma kapsamında normal ve anormal olmak üzere iki kategoriye indirgenmiştir. Tablo 1' de görüldüğü gibi normal hücrelerin sınıf özelliği 1 olarak, anormal hücrelerin ise 0 olarak etiketlenmiştir. Herlev veri setinde, Şekil 8 (a)' da görüldüğü gibi her sitoplazma bir çekirdek içerdiği için sitoplazmada bulunan çekirdek sayısı ve çekirdek sitoplazma arası mesafe sınıflandırma kriterleri içinde kullanılmamıştır.

Tablo 1'de gösterilen ÇekirdekÇevre ve StpCevre özellikleri çekirdek ve sitoplazma konturlarının çevre hesabını ifade etmektedir. Nrnd ve SRnd özellikleri ise Denklem 7'de verilen denklem ile hesaplanmış olan çekirdek ve sitoplazmaların yuvarlaklık oranını belirtmektedir. Anormal hücrelerin çekirdeklerinde genişleme görüldüğü için ÇekirdekAlan özelliği ile belirtilen çekirdek alanları ile StpAlan özelliği belirtilen sitoplazma alanının birbirine yakın olduğu bu nedenle NC_Oranı olarak belirtilen ÇekirdekAlan/StpAlan oranlarının yüksek çıktığı görülmüştür.

Tablo 1 Normal ve anormal hücelere ait bazı veriler

NC_Oranı	CekirdekAlan	StpAlan	CekirdekÇevre	StpCevre	NRnd	SRnd	Class
1.189425684	4825.5	4057	267.8650047	393.8893892	0.84512359	0.3285	0
0.80497289	3415.5	4243	228.5512966	341.060962	0.8216687	0.4583	0
0.01559438	569	36487	92.76955187	752.055909	0.8906882	0.8106	1
0.15582719	3639.5	23356	599.168608	691.712764	0.6928079	0.6134	1

Oluşturulan veri seti üzerinde makine öğrenmesi yöntemlerinden olan SVM, LR, MLP, RF, NaiveBayes ve KNN yöntemleri uygulanmıştır ve elde edilen sonuçlar Tablo 2’de gösterilmiştir. Değerlendirme işleminde ölçüt olarak doğruluk (accuracy), hassasiyet (precision), duyarlılık (sensitivity) ve özgünlük (specificity) değerleri kullanılmıştır.

SVM yönteminde çekirdek fonksiyonu olarak linear kullanılmıştır ve düzenleme parametresi (C) 1.00 olarak alınmıştır.

MLP yönteminde aktivasyon fonksiyonu olarak ReLU (Rectified Linear Unit) kullanılmıştır ve gizli katman sayısı 3 olarak belirlenmiştir. Her bir gizli katmanda bulunan düğüm sayısı 10 ve 100 değerleri arasında değiştirilerek başarı değerleri ölçülmüştür. Ölçümler sonucunda düğüm sayısının 40 ve üstüne çıktığı sınıflandırma işlemlerinde eğitim setinde başarı artarken test setinde başarı azalmıştır. Bu nedenle her katmanda bulunan düğüm sayısı 30 olarak alınmıştır.

RF yönteminde ormanda bulunan ağaç sayısını belirten *n_estimators* değişkeni 10 olarak alınmıştır.

Tablo 2 Sınıflandırma yöntemleri ve doğruluk değerleri

	SVM	RF	MLP	LR	KNN (k=9)	NaiveBayes
Doğruluk (%)	89.2	92.04	90	87.5	90	83.8
Hassasiyet (%)	82	80	93.1	100	85.2	72.9
Duyarlılık (%)	91.2	95.5	85.7	85.8	90.8	86.3
Özgünlük (%)	82	80	93.1	100	85.2	72.9

Sınıflandırma başarısı olarak RF yöntemi en iyi sonucu vermiştir. Naive Bayes yönteminin başarı oranının düşük olması bu yöntemin istatistiksel veriler ile daha iyi sonuç vermesi olarak düşünülmüştür. MLP yönteminde başarı oranının %90 olarak kalması veri sayısının yeterli miktarda olmaması olarak düşünülmüştür.

Literatürde bu alanda yapılan çalışmalar, kullanılan yöntemler ve elde edilen doğruluk değerleri Tablo 3’te verilmiştir.

Tablo 3 Literatürde Herlev veri seti ile son yıllarda yapılan çalışmalar

Çalışma adı	Sınıflandırıcı	Doğruluk değeri (%)
Chankong vd. [28]	Linear discriminant analysis, KNN, Bayes, SVM, ANN	99.27
William vd. [29]	Fuzzy c-means	98.8
Zhang vd. [30]	ESA	98.3
Win vd.	LD+SVM+KNN+ boosted trees + bagged trees	98.27
Önerilen yöntem	SVM, KNN, RF, LR, MLP, NaiveBayes	92.04

Çalışmanın sınıflandırma başarısı Tablo 3 ile verilen Herlev veri seti üzerinde yapılan diğer çalışmalara göre düşük kalmıştır. Bunun nedeni olarak ise görüntülerden çıkarılan özellik sayısının bu çalışmalarda bulunan özellik sayısına göre az olması düşünülmüştür.

3. Sonuçlar

Bu çalışma kapsamında pap-smear görüntüleri kullanılarak serviks hücrelerinin normal veya anormal olma durumu bulunmuştur. Çalışma segmentasyon, özellik çıkarımı ve sınıflandırma olmak üzere üç ana bölümden oluşmaktadır. Segmentasyon işleminde için renk- tabanlı eşikleme ve K-means olmak üzere iki farklı yöntem kullanılmıştır. Özellik çıkarımında ise görüntülere matematik temelli işlemler uygulanarak sınıflandırma kriterleri belirlenmiştir. Önerilen renk-tabanlı segmentasyon ve matematiksel işlemler ile yapılan özellik çıkarımı yöntemleri doğru çekirdek- sitoplazma eşleşmesi yaptığı için ve hücrelerde bulunan fazla çekirdek sayılarını tespit edip anormal hücreleri görüntü üzerinde bulduğu için birden fazla hücre içeren pap-smear test sonuçlarında patoloji uzmanlarına hastalığın teşhisinde katkı sağlayacaktır.

Oluşturulan veri seti genişletilerek DSIL, YSIL için daha belirleyici ve sayıca fazla kriterler ve daha fazla pap-smear görüntüsü eklenerek rahim ağzı kanserini daha erken ve daha başarılı sonuçlar ile tespit eden bilgisayar-temelli bir arayüz geliştirilerek çalışma tekrar yapılabilir.

Teşekkür

Bu çalışmada rahim ağzı kanseri ve hastalığın teşhisi hakkında bilgi veren İstanbul Zeynep Kâmil Kadın ve Çocuk Hastalıkları Eğitim ve Araştırma Hastanesi Patoloji Bölümü'ne teşekkür ederim.

Referanslar

- [1] "National Institutes of Health homepage", 2020. [Online]. Available: <https://www.cancer.gov/about-cancer/understanding/what-is-cancer>. [Accessed: 01-Feb-2020].
- [2] Cancer Today, "Global Cancer Observatory homepage", 2018. [Online]. Available: <http://gco.iarc.fr/today/online-analysis-table>. [Accessed: 02-Feb-2020].
- [3] DB. Cooper, C.E. McCathran, "Cervical Dysplasia", *StatPearls [Internet]*. StatPearls Publishing, 2019.
- [4] RJ. Kurman, D. Solomon, *The Bethesda System for Reporting Cervical/Vaginal Cytologic Diagnoses*. New York: Springer-Verlag, 1994.
- [5] S.E. Waggoner, "Cervical cancer", *The Lancet*, vol. 361, no. 9376, pp. 2217-2225, 2003.
- [6] T. Bilal, J. Dias, and N. Werghi, "Classification of cervical-cancer using pap-smear images: a convolutional neural network approach", *Annual Conference on Medical Image Understanding and Analysis*, Springer, Cham, 2017.
- [7] M.E. Plissiti, N. Christophoros, and A. Charchanti, "Automated detection of cell nuclei in pap smear images using morphological reconstruction and clustering", *IEEE Transactions on information technology in biomedicine*, vol. 15, no.2, pp. 233-241, 2010.

- [8] P. Wang, L. Wang, Y. Li, Q. Song, S. Lv, X. Hu, "Automatic cell nuclei segmentation and classification of cervical Pap smear images", *Biomedical Signal Processing and Control*, vol. 48, pp. 93-103, 2019.
- [9] Y. Marinakis, G. Dounias, and J. Jantzen, "Pap smear diagnosis using a hybrid intelligent scheme focusing on genetic algorithm based feature selection and nearest neighbor classification", *Computers in Biology and Medicine*, vol. 39, no.1, pp. 69-78, 2009.
- [10] A. GençTav, S. Aksoy, and S. Önder, "Unsupervised segmentation and classification of cervical cell images", *Pattern recognition*, vol. 45, no.12, pp. 4151-4168, 2012.
- [11] H.A. Phoulady, "A framework for nucleus and overlapping cytoplasm segmentation in cervical cytology extended depth of field and volume images", *Computerized Medical Imaging and Graphics*, vol. 59, pp. 38-49, 2017.
- [12] K.P. Win, Y. Kitjaidure, K. Hamamoto, T. Myo Aung, "Computer-Assisted Screening for Cervical Cancer Using Digital Image Processing of Pap Smear Images", *Appl. Sci.*, vol. 10, no.5, pp.1800, 2020.
- [13] W. William, A. Ware, A.H. Basaza-Ejiri, & J. Obungoloch, "A pap-smear analysis tool (PAT) for detection of cervical cancer from pap-smear images", *Biomedical engineering online*, vol.18, no.1, pp.16, 2019.
- [14] V. Vapnik, *The nature of statistical learning theory*. New York: Springer-Verlag, 1995.
- [15] M. Pal, "Random forest classifier for remote sensing classification", *International Journal of Remote Sensing*, vol.26, no.1, pp. 217-222, 2005.
- [16] M.W. Gardner, S.R. Dorling, "Artificial neural networks (the multilayer perceptron) —a review of applications in the atmospheric sciences", *Atmospheric Environment*, vol. 32, no. 14–15, pp. 2627-2636,1998.
- [17] Y. Liao, V. R. Vemuri, "Use of k-nearest neighbor classifier for intrusion detection", *Computers & security*, vol. 21, no.5, pp. 439-448, 2002.
- [18] M. M. Saritas, A. Yasar, "Performance Analysis of ANN and Naive Bayes Classification Algorithm for Data Classification", *International Journal of Intelligent Systems and Applications in Engineering*, vol. 7, no. 2, pp. 88-91, 2019.
- [19] S.K. Shevade, S. S. Keerthi, "A simple and efficient algorithm for gene selection using sparse logistic regression", *Bioinformatics*, vol. 19, no. 17, pp. 2246-2253, 2003.
- [20] MDE-lab, "MDE-lab downloadspage,"2011.[Online]. Available: <http://mde-lab.aegean.gr/downloads>. [Accessed: 30-Sep-2019].
- [21] J. Jantzen, J. Norup, Jonas, G. Dounias, B. Bjerregaard, "Pap-smear Benchmark Data For Pattern Classification", *Nature Inspired Smart Information Systems (NiSIS)*, pp. 1-9, 2005.
- [22] N. Nill, "A visual model weighted cosine transform for image compression and quality assessment", *IEEE Transactions on communications*, vol.33, no. 6, pp. 551-557, 1985.
- [23] M. E. Plissiti, P. Dimitrakopoulos, G. Sfikas, C. Nikou, O. Krikoni, A. Charchanti, SIPAKMED: A new dataset for feature and image based classification of normal and pathological

cervical cells in Pap smear images, *IEEE International Conference on Image Processing (ICIP) 2018*, Athens, Greece, 7-10 October 2018.

[24] H. Demirel, G. Anbarjafari, "Discrete Wavelet Transform-Based Satellite Image Resolution Enhancement", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 6, pp. 1997-2004, 2011.

[25] H. Demirel, G. Anbarjafari, "Image resolution enhancement by using discrete and stationary wavelet decomposition", *IEEE transactions on image processing*, vol. 20, no. 5, pp. 1458-1460, 2010.

[26] E.S. Cibas, B.S. Ducatman, *Cytology E-Book: Diagnostic principles and clinical correlates*, Elsevier Health Sciences, 2013.



[27] "Eurocytology Cervical Cytology homepage", 2020. [Online]. Available: <https://www.eurocytology.eu/en/course/1292>. [Accessed: 15-Mar-2020].

[28] T. Chankong, N. Theera-Umpon, & S. Auephanwiriyaikul, "Automatic cervical cell segmentation and classification in Pap smears", *Computer methods and programs in biomedicine*, vol.113, no.2, pp.539-556, 2014.

[29] W. William, A. Ware, A.H. Basaza-Ejiri, & J. Obungoloch, "Cervical cancer classification from Pap-smears using an enhanced fuzzy C-means algorithm", *Informatics in Medicine Unlocked*, vol.14, pp. 23-33, 2019.

[30] L. Zhang, L. Lu, I. Nogues, R.M. Summers, S. Liu, & J. Yao, "DeepPap: deep convolutional networks for cervical cell classification", *IEEE journal of biomedical and health informatics*, vol. 21, no.6, pp. 1633-1643, 2017.

Classification of Robust and Rotten Apples by Deep Learning Algorithm

 Kıyas KAYAALP¹,  Sedat METLEK²

¹Corresponding Author; Isparta Uygulamalı Bilimler Üniversitesi, Uluborlu Selahattin Karasoy Meslek Yüksekokulu, Bilgisayar Teknolojileri Bölümü, Isparta, Türkiye; kiyaskayaalp@isparta.edu.tr; <https://orcid.org/0000-0002-6483-1124>; +902465312622

²Burdur Mehmet Akif Ersoy Üniversitesi, Teknik Bilimler Meslek Yüksekokulu, Elektronik ve Otomasyon Bölümü, Burdur, Türkiye; sedatmetlek@mehmetakif.edu.tr; <https://orcid.org/0000-0002-0393-9908>

Received 09 April 2020; Revised 18 May 2020 Accepted 29 May 2020; Published online 28 August 2020

Abstract

In the study, it is aimed to classify the apples as rotten and robust by using the deep learning algorithm of the apple images taken from the CAPA database. In the proposed model, the processing steps are image reading, preprocessing and classification of apples, respectively. In the image reading stage, images taken from the image database were used. The applied deep learning architecture consists of introduction, convolutional, activation, pooling, memorization, full connection and conclusion layers. The data used in this architecture are divided into two as 80% training and 20% test data. Four different wavelength, 16 kinds of image combinations were used for the training and testing of the system. At the classification stage, a success rate of 91.25% was achieved in detecting rotten and robust apples. As a result, it is predicted that the proposed model can be used in the fruit processing industry to automatically classify rotten and robust apples.

Keywords: apple classification, deep learning, image processing, rotten, apple

Sağlam ve Çürük Elmaların Derin Öğrenme Algoritması ile Sınıflandırılması

Öz

Yapılan çalışmada, CAPA veri tabanından alınan elma görüntülerinin derin öğrenme algoritması kullanılarak, elmaların çürük ve sağlam olarak sınıflandırılması amaçlanmıştır. Önerilen modelde işlem adımları sırasıyla görüntü okuma, ön işleme ve elmaların sınıflandırılmasıdır. Görüntü okuma aşamasında, görüntü veri tabanından alınan görüntüler kullanılmıştır. Uygulanan derin öğrenme mimarisi giriş, evrimsel, aktivasyon, havuzlama, ezberleme, tam bağlantı ve sonuçlandırma katmanlarından oluşmaktadır. Bu mimaride kullanılan veriler, %80 eğitim ve %20 test verisi olmak üzere ikiye ayrılmıştır. Sistemin eğitim ve test işlemleri için dört farklı dalga boyunda, 16 çeşit görüntü kombinasyonu kullanılmıştır. Sınıflandırma aşamasında, çürük ve sağlam elmaların tespit edilmesinde %91.25 başarı oranına ulaşılmıştır. Sonuç olarak, önerilen modelin meyve işleme sanayisinde çürük ve sağlam elmaların otomatik olarak sınıflandırılmasında kullanılabileceği ön görülmektedir.

Anahtar Kelimeler: elma sınıflandırma, derin öğrenme, görüntü işleme, çürük, elma

1. Giriş

Elmanın, içerdiği mineral ve vitaminler açısından beslenmedeki rolü büyüktür. Taze bir elmanın %84'ünü su oluşturmaktadır. Geri kalan %16'lık kuru kısmında ise karbonhidratlar, proteinler, vitaminler, pektinler ve çeşitli mineraller bulunmaktadır. Bu minerallerin kandaki asit-baz dengesi üzerinde olumlu etkileri bulunmaktadır. Özellikle günde bir elma yemenin kanser riskini dahi önemli ölçüde azalttığı tespit edilmiştir [1].

Elma, dünya üzerindeki toplam meyve üretiminin yaklaşık %12'sini oluşturarak, muzdan sonra ikinci sırada yer alan önemli bir meyve türüdür. Bununla beraber elma gıda sektöründe, reçel, elma suyu, elma

şırası ve sirke gibi ürünlerin ham maddesi olarak kullanılmaktadır. Ayrıca, üretimin son aşamasında atık olarak elde edilen elma posasında bile insan sağlığı için önemli olan pektin maddesi bulunmaktadır [2].

Elma posasının bile bu kadar önemli olmasının yanı sıra meyve olarak tüketilmesi de oldukça önemlidir. Bu nedenle elmaların meyve olarak tüketilmesi için ilk olarak sağlam ve çürük elmaların ayırt edilmesi gerekmektedir. Bu ayırım işlemi, teknolojinin gelişmesine bağlı olarak sınıflandırma makineleri tarafından yapılmaktadır. Bu makineler görüntü işleme tabanlı çalışmakla beraber, kontrol mekanizmalarının içerisinde birçok sınıflandırma algoritması bulunmaktadır.

Teknolojinin gelişmesine paralel olarak sınıflandırma algoritmalarında da sürekli yeni yöntemler geliştirilmektedir. Bu sınıflandırma algoritmaları da görüntüden elde edilen öznelikleri temel alarak sınıflandırma yapmaktadır. Bu nedenle elma görüntülerinde bulunan sap ve taç kısımları ile çürükler karıştırılmakta ve bu kısımlar bazen çürük olarak nitelendirilebilmektedir. Bu durum ile ilgili olarak literatürde yapılan sınıflandırmalarda iki yaklaşım mevcuttur. Bunlardan ilkinde elma görüntülerinde sap ve taç bölgeleri bulunmamaktadır. Böylece geri kalan elma yüzey görüntüleri üzerinde sağlam ve çürük sınıflandırılması yapılmaktadır. Diğer yaklaşımda ise elma görüntülerinde sap ve taç bölgeleri bulunmaktadır. Bu yaklaşımda görüntülerden ilk olarak sap ve taç bölgeleri ayırt edilmekte ve akabinde sınıflandırma işlemi gerçekleştirilmektedir [3].

Leemans ve arkadaşları 2002 yılında yaptıkları çalışmada, 1170 adet elma görüntüsünden (528 Golden Delicious, 642 Jonagold) oluşan görüntü veri tabanını kullanarak, sağlam ve çürük meyve sınıflandırmasını gerçekleştirmişlerdir. Elma sınıflandırma işlemini; görüntü elde etme, zemin renk sınıflandırması, kusur bölümlendirme, sap-taç tanıma, kalite sınıflandırma ve meyve sınıflandırma olarak altı adımda gerçekleştirmişlerdir. Golden Delicious'da %78 ve Jonagold'da %72 oranında sınıflandırma başarısı elde etmişlerdir [4].

Sofu ve arkadaşlarının 2013 yılında, tasarladıkları sistem ile elma görüntülerini elde etmişlerdir. Geliştirdikleri yazılım ile elmaların renk tahmininde %99 ve cins tahmininde %95.52 doğrulukla sınıflandırma başarısını yakalamışlardır. Uygulanan algoritma sayesinde, elma üzerindeki lekelerin tespitini herhangi bir filtreye gereksinim duymadan %70 başarı ile gerçekleştirmişlerdir [5].

2016 yılında, Cömert ve arkadaşları 600 adet sağlam, 600 adet çürük elma görüntüsünden oluşan bir veri tabanı kullanarak, elmaları sınıflandırmışlardır. Görüntü alma-önişleme, çürük bölgelerin tespit edilmesi ve elmaların sınıflandırması olmak üzere 3 adımda oluşturdukları modelde, evrimsel sinir ağı yöntemlerinden olan Faster R-CNN yöntemini kullanmışlardır. Derin öğrenme ve görüntü işlemenin birlikte kullanıldığı bu yaklaşım sayesinde %84.95 doğru sınıflandırma oranına ulaşmışlardır [3].

Lu ve arkadaşı 2018 yılındaki çalışmalarında, elmaların yüzey ve yüzey altı kusurlarının daha iyi algılanması için aydınlatmanın önemi üzerinde durmuşlardır. Bu nedenle aydınlatmanın direkt bileşeni, genlik bileşeni ve bunların kombinasyonlarını kullanarak elde ettikleri görüntüler ve esnek hesaplama yöntemleri sayesinde sınıflandırma işlemlerini gerçekleştirmişlerdir [6].

Siddiği 2019 yılında yaptığı çalışmasında, 244 çürük elma görüntüsünden oluşan bir görüntü veri tabanını kullanarak elmaları sınıflandırmıştır. Bu sınıflandırmada, çürük elmalardaki çürüğün yerini bulmak için YOLO (you only look once) ve SSD (single shot multibox detector) yöntemlerini kullanmıştır. SSD tabanlı sistemin YOLO tabanlı sisteme göre daha başarılı olduğunu tespit etmiştir [7].

Yu ve arkadaşları 2019 yılında, elma görüntüleri üzerindeki sap ve taç bölgeleri ile çürük bölgelerin ayırımını yapmak için bir çalışma gerçekleştirmişlerdir. Bu çalışmada elmaların alt, üst ve diğer yüzey görüntülerinin bulunduğu gri formatta bir görüntü veri tabanı üzerinde çalışmışlardır. Bu veri tabanındaki 60 adet elma görüntüsünü eğitim, 30 adet elma görüntüsünü de test için kullanmışlardır. Hazırladıkları algoritmada K-means kümeleme yöntemi sayesinde %96 sınıflandırma başarı oranını elde etmişlerdir [8].

Görüntü işleme ile elma sınıflandırılması işlemlerinde kullanılan teknolojilerin gün geçtikçe geliştiği görülmektedir. Bunun ile birlikte, donanım teknolojisinin değişmesine paralel olarak gelişen birçok güncel sınıflandırma algoritması bulunmaktadır. Bu çalışmada güncel sınıflandırma algoritmalarından olan derin öğrenme algoritması kullanılmıştır. Gerçekleştirilen çalışmada, CAPA veri tabanından alınan

görüntülerin derin öğrenme algoritması kullanılarak, elmaların çürük ve sağlam olarak sınıflandırılması sağlanmıştır.

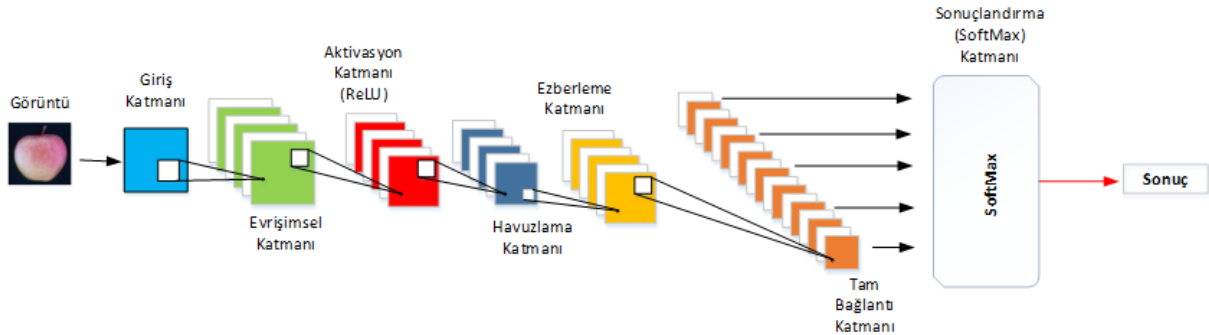
2. Materyal ve Metot

Çalışmada, Gembloux Tarım Üniversitesinde bir araştırma sonucunda, Jonagold elma görüntülerinin oluşturduğu CAPA görüntü veri tabanı kullanılmıştır. Bu veri tabanındaki görüntüler 450, 500, 750 ve 800 nm dalga boyunda alınmıştır. Alınan görüntüler sağlam, çürük, sap ve taç kısımlarının görüntüleri olmak üzere dörde ayrılmaktadır. Bu görüntülerden 280 adeti sağlam, 246 adeti çürük, 148 adeti sap ve 145 adeti taç görüntüleridir. Belirtilen her bir görüntünün de dört farklı dalga boyutunda kombinasyonu bulunmaktadır. Bu nedenle görüntü veri tabanı toplamda 3276 adet görüntüden oluşmaktadır. Çalışmada kullanılan çürük görüntüleri çeşitli dış yüzey kusurlarına (pas rengi, morarma, çürüme, güneş yanığı, dolu hasarı gibi) sahiptir. Bu görüntüler 430×560 piksel boyutlarında ve her bir piksel başına 8 bit çözünürlüğe sahiptir [9]. Gerçekleştirilen çalışmada sadece 280 adet sağlam, 246 adet çürük olmak üzere toplam 526 adet elma görüntüsü kullanılmıştır. Kullanılan görüntüler içerisinde sap ve taç kısımları bulunmamaktadır.

Literatürde son yıllarda yapılan, elma görüntülerinden sağlam ve çürük olarak sınıflandırma ile ilgili yapılan çalışmalar incelendiğinde, hiperspektral ve multispektral görüntüleme sistemlerinin kullanıldığı görülmektedir [10-16]. Fakat tercih edilen sistemlerin genellikle laboratuvar ortamında çalışabilecek nitelikte olduğu ve maliyetlerinin ise çok fazla olduğu görülmektedir. Bu sebeple, ilerleyen dönemlerde maliyeti daha düşük sistemlerle bu görüntülerin alınacağı aşikardır. Ancak, geliştirilen yazılımların bu sistemler ile entegrasyonunun sağlanması daha da önem kazanmaktadır. Günümüzde birçok sisteme entegre edilebilen derin öğrenme algoritmalarının tarım sektöründe de birçok alt alana uygulanmaktadır. Bu nedenle çalışmamızda güncel sınıflandırma algoritmalarından olan derin öğrenme ile çürük ve sağlam elmaların sınıflandırma işlemleri test edilmiş ve uygulanmıştır.

2.1 Derin Öğrenme

Günümüzün popüler konularından olan derin öğrenme, esnek hesaplama yöntemlerinin en son geldiği noktalardan birisidir. İlk esnek hesaplama modellerinden olan yapay sinir ağları (YSA), zamanla geliştirilerek Evrimsel Sinir Ağı (ESA) modelleri oluşturulmuştur. Derin öğrenme mimarisinin temelini de geliştirilen ESA'lar oluşturmaktadır. İlk YSA modellerinde geliştiriciler sırasıyla, öznelik elde etme ve en iyi öznelikleri seçme işlemlerini gerçekleştirdiler [17]. Sonrasında geliştirilen ESA modellerinde ise, geliştiriciler ellerindeki verilerden öznelik çıkarma işlemini, ESA'nın içerisinde bulunan filtrelerle yaptırmışlardır. Ayrıca, bu filtreler sayesinde elde edilen veriler ESA'da doğrudan kullanılmaktadır. ESA sonucunda elde edilen veriler istenildiğinde farklı sınıflandırma algoritmalarında da kullanılabilmektedir. Derin öğrenme algoritmalarında ise, tüm öznelik çıkarma işlemleri geliştirilen ESA'larda olduğu gibi filtreler yardımı ile otomatik olarak yapılmaktadır [18].



Şekil 1 Derin öğrenme mimarisini


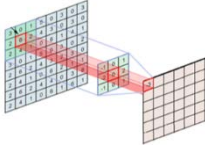
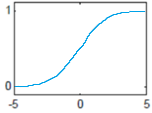
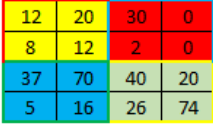
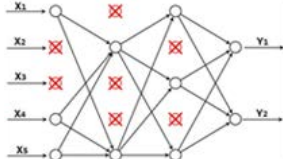
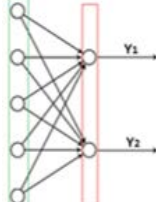
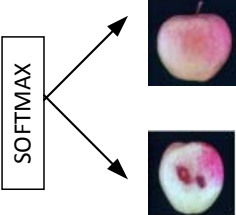
Bunun ile birlikte sınıflandırma algoritması da bu ağ yapısının içerisinde bulunan softmax katmanında gerçekleştirilmektedir [19]. Derin Öğrenme algoritmalarında, geliştiricilerin sadece dışarıdan uygun veri tiplerini ağa sunması yeterlidir. Kullanımının bu denli kolay olmasından dolayı pek çok farklı

araştırma alanında olduğu gibi uygulama ağırlıklı olan tarım teknolojilerinde de derin öğrenme algoritmaları yaygın olarak kullanılmaya başlanmıştır [20].

Derin Öğrenme mimarilerinin özellikle görüntü işleme ve sınıflandırma uygulamalarında yaygın olarak kullanıldığı görülmektedir [21-23]. Bu çalışmada da görüntüler üzerinden sınıflandırılma yapılacak olması nedeniyle derin öğrenme algoritması tercih edilmiştir. Kullanılan derin öğrenme mimarisinin genel yapısı Şekil 1’de verilmiştir.

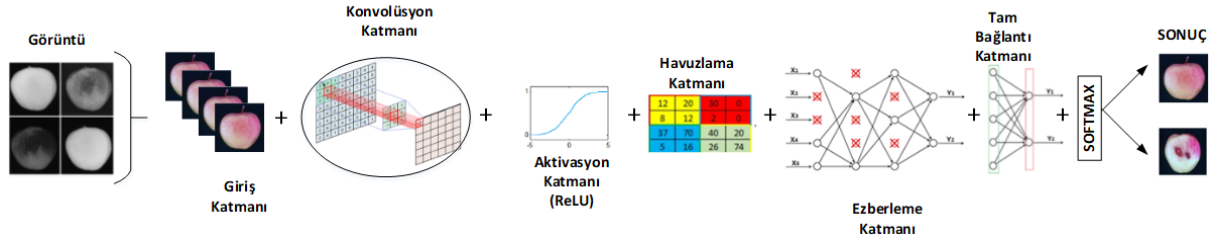
Şekil 1’de gösterildiği üzere derin öğrenme mimarisi kendi içerisinde katmanlı bir yapıya sahiptir. Bu katmanlar sırasıyla; giriş, evrimsel, aktivasyon, havuzlama, ezberleme, tam bağlantı ve sonuçlandırma katmanlarıdır. Bu katmanların detayları Tablo 1’de gösterilmiştir.

Tablo 1 Derin öğrenme modeli katmanları ve özellikleri

	<p>Giriş Katmanı; Derin Öğrenme algoritmalarında giriş katmanına genellikle görüntüler aktarılmaktadır. Aktarılan bu görüntüler farklı tür, boyut ve içerikteki verilerden oluşabilmektedir. Bu nedenle bu katmanda veriler belli bir standarda getirilerek kendisinden sonraki katmana aktarılmaktadır.</p>
	<p>Evrimsel Katmanı; Giriş katmanından gelen belli standarttaki veriler üzerinde, boyutlarını geliştiricilerin belirlediği $n \times n$ boyutlarındaki filtrelerin dolaylı olarak yeni özneliklerin çıkartılarak daha küçük bir veri matrisinin elde edildiği katmandır.</p>
	<p>Aktivasyon Katmanı; Bu katmanın görevi evrimsel katmanından elde edilen matris değerlerinin kullanılan algoritmaya bağlı olarak negatif değerlerin pozitifeye dönüştürülmesidir. Bunun için, sigmoid, tanjant hiperbolik, adım ve ReLU gibi farklı aktivasyon fonksiyonları kullanılmaktadır.</p>
	<p>Havuzlama Katmanı; Aktivasyon katmanından gelen verileri daha küçük boyutlardaki matrislere indirmek için kullanılan katmandır. Bu veri indirgeme işlemi ağı daha hızlı çalışmasını sağlamasına rağmen, bazı durumlar için verilerde az miktarda da olsa kayba neden olabilmektedir. Veri indirgeme için ortalama, en küçük değer ve en büyük değer matrisleri gibi farklı birçok matris de kullanılabilir.</p>
	<p>Ezberleme Katmanı; Ağın eğitim setini ezberleyebildiği az sayıda veriye sahip eğitim setlerinde, ağın ezberlenmesinin önüne geçmek için kullanılan katmandır.</p>
	<p>Tam Bağlantı Katmanı; Kendisinden önce gelen tüm verileri alarak tek boyutlu bir matrise dönüştüren katmandır. Örneğin, tam bağlantı katmanı 2048×1 olarak seçilirse, $32 \times 32 \times 100$ boyutundaki veriler bu katmanda, 102400×2048 boyutlarında bir ağırlık matrisi oluşturur.</p>
	<p>Sonuçlandırma Katmanı; Derin Öğrenme mimarisinin son katmanı olan bu katman, tam bağlantı katmanından gelen verilerin değerlendirilerek ağın çıkışının oluşturulduğu katmandır. Burada, sınıflandırma bilgisinin yanında gerçekleştirilen uygulamaya bağlı olarak herhangi bir eğrinin katsayıları da hesaplanabilmektedir. Genellikle derin öğrenme algoritmalarında sonuçlandırma katmanında softmax sınıflandırıcısı kullanılmaktadır.</p>

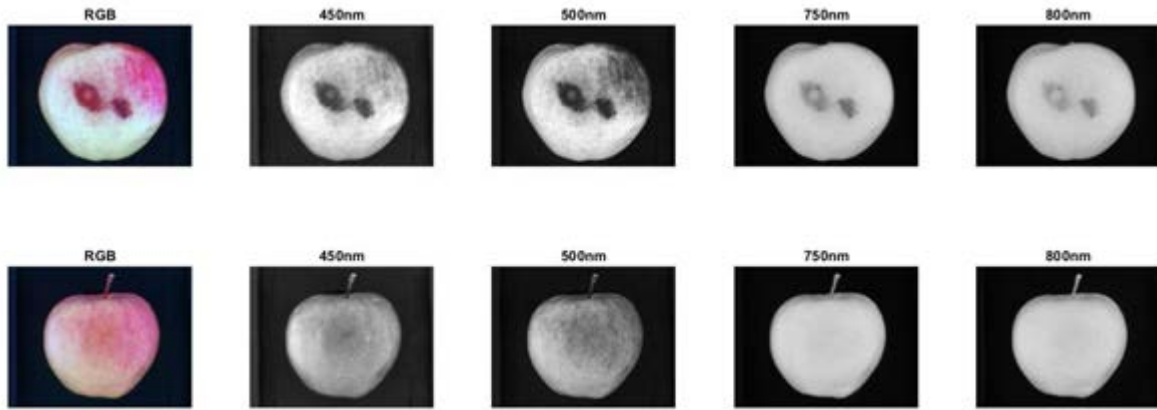
3. Bulgular ve Tartışma

Gerçekleştirilen uygulama Matlab ortamında, CAPA görüntü veri tabanı üzerinde Şekil 2’de gösterilen derin öğrenme mimarisi kullanılarak sağlam ve çürük elmaların sınıflandırılması sağlanmıştır. Bu aşamada ilk olarak sistemdeki 4 farklı dalga boyunda kaydedilen görüntüler alınarak sistemin girişini oluşturacak bir veri tabanı oluşturulmuştur.



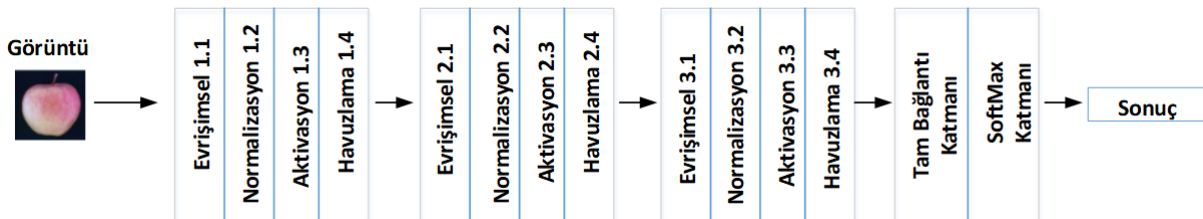
Şekil 2 Elma sınıflandırmasında kullanılan genel derin öğrenme mimarisi

Elmaların farklı dalga boylarındaki görüntülerine ek olarak, RGB formatındaki sonuçların daha iyi çıkabilme ihtimali için görüntülerin RGB formatı elde edilerek, hazırlanan yeni veri tabanına eklenmiştir. Şekil 3’de veri tabanında bulunan, iki farklı sınıftaki elmaya ait görüntüler, sırasıyla gösterilmiştir.



Şekil 3 Çürük ve sağlam elmaya ait RGB ve tüm dalga boylarındaki görüntüleri

Veri tabanı oluşturma işleminden sonra, alınan görüntüler katmanlı mimariye sahip derin öğrenme algoritmasına gönderilmiştir. Çalışmada kullanılan mimari arka arkaya üç adet evrişimsel yapının eklenmesi ile oluşturulmuştur. Oluşturulan her bir evrişimsel yapıda (Şekil 4) sırasıyla; evrişimsel (3×8), normalizasyon, aktivasyon (sigmoid), havuzlama (2×2 enbüyük), evrişimsel (3×16), normalizasyon, aktivasyon (sigmoid), havuzlama (2×2 enbüyük), evrişimsel (3×32), normalizasyon, aktivasyon (sigmoid), havuzlama (2×2 enbüyük) katmanları bulunmaktadır. Mimarinin son katmanında ise, tam bağlantı ve softmax katmanı bulunmaktadır.



Şekil 4 Çalışmada kullanılan derin öğrenme sınıflandırma mimarisi

Tasarlanan yapının eğitilmesinde k-fold 4 değeri için verilerin %80’i eğitim, %20’si test verisi olarak kullanılmıştır. Bunun yanında, öğrenme katsayısı olarak 0.005 değeri alınmıştır. Sistemin eğitim ve test işlemleri için RGB renk uzayındaki değeri ve Tablo 3’de kullanılan farklı dalga boylarındaki değerleri sırasıyla test edilmiştir.

Hata (Confusion) matrisi bir sınıflandırma modelinin tahminlerinin, ne kadar başarılı olduğunu özetleyen $M \times M$ şekilde bir tablodur. Bu tablonun bir ekseninde tahmin edilen sonuçlar, diğer ekseninde de gerçek değerler gösterilir. Hata matrisi, sınıflandırıcı tarafından yapılan hataları ve hata türleri hakkında, aynı zamanda da model performansı konusunda genel bilgi verir. Bu nedenle gerçekleştirilen çalışmanın hata matrisi Tablo 2’de sunulmuştur.

Tablo 2 Hata matrisi

		Gerçek Değerler	
		Pozitif	Negatif
Tahmin Değerleri	Pozitif	(TP) 260	(FP) 20
	Negatif	(FN) 26	(TN) 220

Tablo 2’de, “TP” tahmin sonucu pozitif gerçek değeride pozitif olanı, “FN” tahmin sonucu negatif gerçek değeride pozitif olanı, “TN” tahmin sonucu negatif gerçek değeride negatif olanı, “FP” tahmin sonucu pozitif gerçek değeride negatif olan değerleri ifade etmektedir.

Doğruluk (Accuracy), doğru tahmin edilen örneklerin (TP ve TN), toplam tahmin edilen örneklere bölünmesiyle elde edilir. Yapılan çalışmanın doğruluğu Denklem 1 ile hesaplanmış ve %91.25 olarak bulunmuştur.

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

Tablo 3’de test görüntüleri ile tüm sistemden elde edilen sınıflandırma başarısı ayrıntılı olarak verilmiştir. Tablo 3’de de görüldüğü üzere 450nm ile 500nm dalga boylarında %91.25 başarı oranı ile sınıflandırma işlemi gerçekleştirilmiştir. Elde edilen başarı oranının literatürde bulunan benzer çalışmalarla ayrıntılı olarak kıyaslaması Tablo 4’de sunulmuştur. Önerilen model aynı veri tabanını kullanan diğer çalışmalar [4,26,27] ile karşılaştırıldığında, sadece üç evrimsel katman ile %91.25 başarı oranını yakalamış bulunmaktadır. Bu nedenle, çalışmaya daha fazla sayıda evrimsel katman eklenerek bu oranın çok daha yukarıya çekilebileceği açıkça ön görülmektedir. Şu anki çalışmada temelde sadece üç evrimsel katman kullanılarak, birçok gömülü sisteme de kolayca adapte edilebilecek bir yapı sunulmuştur.

Tablo 3 Sınıflandırma başarı sonuçları

No	Kullanılan Değerler	İterasyon	Başarı Oranı %
1	RGB	1000	75.91
2	450nm, 500nm, 750nm, 800nm	1000	86.35
3	450nm, 500nm, 750nm	1000	88.87
4	450nm, 500nm, 800nm	1000	87.06
5	450nm, 750nm, 800nm	1000	84.33
6	500nm, 750nm, 800nm	1000	85.56
7	450nm, 500nm	1000	91.25
8	450nm, 750nm	1000	88.70
9	450nm, 800nm	1000	87.09
10	500nm, 750nm	1000	86.53
11	500nm, 800nm	1000	85.16
12	750nm, 800nm	1000	75.97
13	450nm	1000	89.45
14	500nm	1000	90.71
15	750nm	1000	71.38
16	800nm	1000	72.40

Tablo 4 Elmalarda çürük tespiti için literatürde yapılmış çalışmalar

Yazar	Yıl	Elma Türü	Görüntü Sayısı	Yöntem	Başarı Oranı (%)
Leemans ve diğ. [4]	2002	Jonagold	642	Tek gizli katmanlı YSA	72
Leemans ve diğ. [4]	2002	Golden Delicious	528	Tek gizli katmanlı YSA	78
Leemans ve Destain [24]	2004	Jonagold	400	Kmeans kümeleme	73
Xing ve Baerdemaeker [25]	2005	Jonagold	160	Yumuşatma, Eşikleme, Hyperspectral görüntüleme	77.5
Unay ve Gosselin [26]	2006	Jonagold	526	CNN, SOM, k-NN, LDC, QDC, LR, SVM, MLP, CFNN, ENN, LVQ	84.8
Unay ve diğ. [27]	2011	Jonagold	526	LDC, k-NN, Fuzzy k-NN, SVM, C4.5	93.5
Mohana ve Prabhakar [28]	2015	Jonagold	200	Multifractal, Fourier ve Radon dönüşümleri, SVM, ANN, k-NN, LDC ve AdaBoost	94
Dubey ve Jalal [29]	2015	Farklı türler	320	Lab-HSV Renk Uzayları, Kmeans kümeleme, SVM	93
Lu ve Lu [30]	2017	Fuji	120	Hyperspectral görüntüleme, Ridler, Unimode ve Otsu eşikleme	90
Zhang ve diğ. [31]	2017	Fuji	155	SVM, Hyperspectral	92.5
Cömert ve diğ. [3]	2019	Farklı türler	1200	Faster R-CNN	84.95
Önerilen Metot	2020	Jonagold	526	Deep Learning (Softmax)	91.25

4. Sonuç

Bu çalışmada gıda sektöründe oldukça fazla alanda kullanılan bir ürün olan elmanın sağlam ve çürük olarak sınıflandırılması amaçlanmıştır. Bu amaç doğrultusunda, CAPA görüntü veri tabanı ve derin öğrenme algoritması kullanılarak, sağlam ve çürük elmalar sınıflandırılmıştır. Kullanılan görüntü veri seti tekrar düzenlenmiş ve farklı dalga boylarında 16 değişken durum göz önünde bulundurularak eğitim ve test işlemleri gerçekleştirilmiştir. 450nm ve 500nm dalga boylarının birlikte kullanıldığı görüntülerden elde edilen öznelikler ile başarı oranının %91.25 olduğu tespit edilmiştir.

Laboratuvar ortamlarında oluşturulan görüntü işleme tabanlı çalışan otomasyon sistemleri genellikle kamera, aydınlatma ve bilgisayardan oluşmaktadır. Bunların kontrolü ise çoğu zaman manuel olarak yapılmaktadır. Halbuki üretim ortamlarında çalışan otomasyon sistemleri ise kamera, endüstriyel bilgisayar, aydınlatma üniteleri, motor sürücü üniteleri gibi kompleks bir sistemden oluşmaktadır. Bunların kontrolünde otomatik olarak yapılmaktadır. Üretim ortamındaki sistemlerin hızlı bir şekilde çalışabilmesi için yazılım ve donanım uyumunun olması şarttır. Halbuki laboratuvar ortamında gerçekleştirilen sistemlerde böyle bir uyum ve hız aranmamaktadır. Günümüzde ise teknolojinin gelişmesine bağlı olarak işlem kapasitesi yüksek birçok gömülü sistem geliştirilmiştir. Buna paralel olarak da daha öncesinde ayrı yazılım ve donanım sistemleri birleştirilerek tek platformda gerçekleştirilmektedir [32]. Gerçekleştirilen algoritma gömülü sistem üzerinde çalıştırılabilecek nitelikte olduğu için kameradan görüntü alma, alınan görüntüyü işleme ve dış ortamda bulunan kontrol kartları ve sürücülerin kontrol edilmesi yazılımsal olarak aynı platformda geliştirilebilecektir. Bu nedenle çalışmanın başarı oranı, tüm sistemin gerçek üretim ortamlarındaki durumu göz önüne alındığında kullanılabilir niteliktedir.

Referanslar

- [1] H. Kurt, Ş. Ö. Keşkek, T. Çil, and A. Canataroğlu, “Meme kanserli hastalarda tamamlayıcı/alternatif tedavi kullanımı,” *Türk Onkol. Derg.*, vol. 28, no. 1, pp. 10–15, 2013.
- [2] C. A. Perussello, Z. Zhang, A. Marzocchella, and B. K. Tiwari, “Valorization of apple pomace by extraction of valuable compounds,” *Compr. Rev. Food Sci. Food Saf.*, vol. 16, no. 5, pp. 776–796, 2017.
- [3] O. Cömert, M. Hekim, and K. Adem, “Faster R-CNN Kullanarak Elmalarda Çürük Tespiti,” *Uluslararası Mühendislik Araştırma ve Geliştirme Derg.*, vol. 11, no. 1, pp. 335–341.
- [4] V. Leemans, H. Magein, and M.-F. Destain, “On-line fruit grading according to their external quality using machine vision,” *Biosyst. Eng.*, vol. 83, no. 4, pp. 397–404, 2002.
- [5] M. M. Sofu, O. Er, M. C. Kayacan, and B. Cetişli, “Elmaların görüntü işleme yöntemi ile sınıflandırılması ve leke tespiti,” *Gıda Teknol. Elektronik Derg.*, vol. 8, no. 1, pp. 12–25, 2013.
- [6] Y. Lu and R. Lu, “Detection of surface and subsurface defects of apples using structured-illumination reflectance imaging with machine learning algorithms,” *Trans. ASABE*, vol. 61, no. 6, pp. 1831–1842, 2018.
- [7] R. Siddiqi, “Automated apple defect detection using state-of-the-art object detection techniques,” *SN Appl. Sci.*, vol. 1, no. 11, p. 1345, 2019.
- [8] Y. Yu, S. A. Velastin, and F. Yin, “Automatic grading of apples based on multi-features and weighted K-means clustering algorithm,” *Inf. Process. Agric.*, 2019.
- [9] O. Kleynen, V. Leemans, and M.-F. Destain, “Development of a multi-spectral vision system for the detection of defects on apples,” *Journal of Food Engineering*, vol. 69, no. 1, pp. 41–49, 2005.
- [10] B. Zhang *et al.*, “Principles, developments and applications of computer vision for external quality inspection of fruits and vegetables: A review,” *Food Res. Int.*, vol. 62, pp. 326–343, 2014.
- [11] S. Cubero, W. S. Lee, N. Aleixos, F. Albert, and J. Blasco, “Automated systems based on machine vision for inspecting citrus fruits from the field to postharvest—a review,” *Food Bioprocess Technol.*, vol. 9, no. 10, pp. 1623–1639, 2016.
- [12] A. Folch-Fortuny, J. M. Prats-Montalbán, S. Cubero, J. Blasco, and A. Ferrer, “VIS/NIR hyperspectral imaging and N-way PLS-DA models for detection of decay lesions in citrus fruits,” *Chemom. Intell. Lab. Syst.*, vol. 156, pp. 241–248, 2016.
- [13] X. Zhu and G. Li, “Rapid detection and visualization of slight bruise on apples using hyperspectral imaging,” *Int. J. Food Prop.*, vol. 22, no. 1, pp. 1709–1719, 2019.
- [14] Z. Du, X. Zeng, X. Li, X. Ding, J. Cao, and W. Jiang, “Recent advances in imaging techniques for bruise detection in fruits and vegetables,” *Trends Food Sci. Technol.*, 2020.
- [15] X. Zeng, Y. Miao, S. Ubaid, X. Gao, and S. Zhuang, “Detection and classification of bruises of pears based on thermal images,” *Postharvest Biol. Technol.*, vol. 161, p. 111090, 2020.
- [16] M. Zhang, Y. Jiang, C. Li, and F. Yang, “Fully convolutional networks for blueberry bruising and calyx segmentation using hyperspectral transmittance imaging,” *Biosyst. Eng.*, vol. 192, pp. 159–175, 2020.

- [17] Y. LeCun *et al.*, “Backpropagation applied to handwritten zip code recognition,” *Neural Comput.*, vol. 1, no. 4, pp. 541–551, 1989.
- [18] R. P. Lippmann, “Review of neural networks for speech recognition,” *Neural Comput.*, vol. 1, no. 1, pp. 1–38, 1989.
- [19] B. Yuan, “Efficient hardware architecture of softmax layer in deep neural network,” in *2016 29th IEEE International System-on-Chip Conference (SOCC)*, 2016, pp. 323–326.
- [20] A. Kamilaris and F. X. Prenafeta-Boldú, “Deep learning in agriculture: A survey,” *Comput. Electron. Agric.*, vol. 147, pp. 70–90, 2018.
- [21] J. Wan *et al.*, “Deep learning for content-based image retrieval: A comprehensive study,” in *Proceedings of the 22nd ACM international conference on Multimedia*, 2014, pp. 157–166.
- [22] M. I. Razzak, S. Naz, and A. Zaib, “Deep learning for medical image processing: Overview, challenges and the future,” in *Classification in BioApps*, Springer, 2018, pp. 323–350.
- [23] C. Tian, Y. Xu, L. Fei, and K. Yan, “Deep Learning for Image Denoising: A Survey,” *Int. Conf. Genet. Evol. Comput.*, pp. 563–572, 2018.
- [24] V. Leemans and M.-F. Destain, “A real-time grading method of apples based on features extracted from defects,” *J. Food Eng.*, vol. 61, no. 1, pp. 83–89, 2004.
- [25] J. Xing and J. De Baerdemaeker, “Bruise detection on ‘Jonagold’ apples using hyperspectral imaging,” *Postharvest Biol. Technol.*, vol. 37, no. 2, pp. 152–162, 2005.
- [26] D. Unay and B. Gosselin, “Automatic defect segmentation of ‘Jonagold’ apples on multi-spectral images: A comparative study,” *Postharvest Biol. Technol.*, vol. 42, no. 3, pp. 271–279, 2006.
- [27] D. Unay, B. Gosselin, O. Kleynen, V. Leemans, M.-F. Destain, and O. Debeir, “Automatic grading of Bi-colored apples by multispectral machine vision,” *Comput. Electron. Agric.*, vol. 75, no. 1, pp. 204–212, 2011.
- [28] S. H. Mohana and C. J. Prabhakar, “Stem-calyx recognition of an apple using shape descriptors,” *arXiv Prepr. arXiv1501.01083*, 2015.
- [29] S. R. Dubey and A. S. Jalal, “Apple disease classification using color, texture and shape features from images,” *Signal, Image Video Process.*, vol. 10, no. 5, pp. 819–826, 2016.
- [30] Y. Lu and R. Lu, “Histogram-based automatic thresholding for bruise detection of apples by structured-illumination reflectance imaging,” *Biosyst. Eng.*, vol. 160, pp. 30–41, 2017.
- [31] S. Zhang, X. Wu, S. Zhang, Q. Cheng, and Z. Tan, “An effective method to inspect and classify the bruising degree of apples based on the optical properties,” *Postharvest Biol. Technol.*, vol. 127, pp. 44–52, 2017.
- [32] “Gömülü kamera sistemleri”, [Online]. <https://www.matrix-vision.com/smart-cam-compact-application-camera.html> [Erişim tarihi: 21.05.2020.]

An Investigation into the Relationship between Curse of Dimensionality and Dunning-Kruger Effect

 Mehmet Cem Catalbas

Electronic Technology, Vocational High School Department, OSTIM Technical University Ankara/Turkey;
mccmatalbas@gmail.com; <https://orcid.org/0000-0002-9291-1180>

Received 26 April 2020; Revised 6 July 2020; Accepted 24 July 2020; Published online 28 August 2020

Abstract

This study addresses a novel perspective for analyzing the source of confidence in human behavior. The concept of confidence was examined via the relationship between two phenomena in the area of machine learning and psychology, namely the Dunning-Kruger effect and the curse of dimensionality. A relationship was established between these two phenomena which were investigated in the light of neuroscience. This study claims that confidence is highly related with the total time it takes to reach specific information and this relationship is inversely proportional. Image gender classification algorithm was used to analyze this relationship for this study and the curves which were obtained as a result of this analysis was compared with the curve of Dunning-Kruger effect and curse of dimensionality. This relationship has been explained by the knowledge of human's problem-solving ability and mathematical models of memory.

Keywords: Neuroscience, Artificial Intelligence, Dunning-Kruger, Memory Model, Cognitive Science

Boyutun Laneti ve Dunning-Kruger Etkisi Arasındaki İlişkinin İncelenmesi

Öz

Bu çalışma insan davranışındaki güvenin kaynağının incelenmesi ile ilgili yenilikçi bir bakış açısı içermektedir. Güven kavramı, makine öğrenmesinin ve psikoloji biliminin iki fenomeni olan boyutun laneti ve Dunning-Kruger etkisine göre incelenmiştir. Bu iki fenomen arasındaki ilişki sinir bilimi ışığında incelenmiştir. Bu çalışmaya göre bireyin güveni o konudaki özel bilgiye erişme süresi ile alakalıdır ve bu ilişki ters orantılıdır. Cinsiyet imgelerinin sınıflandırılması sürecinde bu ilişki analiz edilmiştir ve bu süreç sonucunda elde edilen eğriler Dunning-Kruger etkisi ve boyutun laneti kavramı ile karşılaştırılmıştır. Bu ilişki insanoğlunun problem çözme yeteneği ve hafızanın matematiksel modellerine ait bilgiler kullanılarak açıklanmıştır.

Anahtar Kelimeler: Sinir Bilimi, Yapay Zeka, Dunning-Kruger, Hafıza Modeli, Bilişsel Bilim

1. Introduction

Throughout history, humankind has explored the causes of their behavior and brain activity which has been a black-box since the ancient Greeks. Neuroscience, which is known as the science of the brain, has opened a way to systematically analyze human behavior [1]. Neuroscience is a branch of science that examines the relationship of the nervous system with behavior and learning. Research into human behavior will never end. As neuroscience can develop, measure and observe living beings, its popularity increases day by day. Additionally, neuroscience is an interdisciplinary science which is related to areas such as, mathematics, linguistics, engineering and computer science [2,3]. Neuroscience provides an understanding of various diseases in the brain including Parkinson's disease and brain tumors [4]. Nowadays, research on the brain has reached a new era thanks to brain imaging techniques such as Magnetic resonance imaging (MRI), especially functional MRI [5,6]. Neuroscience also investigates how the brain processes new information and stores it which is highly important for the field of

education and there are several studies conducted in this area to increase the learning ability of a student in the classroom [7-9]. Another area of interest in neuroscience investigates the reasons behind the decision one makes regarding social life. This trending topic opens a new perspective in the area of advertising as it helps to analyse the reasons customers prefer certain products. Customers' product preferences are analyzed in many different ways by neuroscientists to make products more attractive [10]. Another area which is investigated in the scope of neuroscience is the vision ability of human beings.

Neuroscientists try to understand the stages of how a real object is transformed into a representation in the brain which is a process that also leads to the understanding and development of the areas of computer vision and artificial intelligence [11,12]. According to research in this area, the human brain performs edge detection and determines the shape at first contact after which the colors and other information available are processed. In this study, the relationship between psychology and machine learning was investigated and the source of confidence was analyzed according to this relationship. Firstly, the Dunning-Kruger effect, a phenomenon in psychology, was explained and then another term in the area of machine learning which is known as the curse of dimensionality was explained briefly.

The common part of these two terms was discussed in the light of computational neuroscience. Although these two phenomena occur at different times and in different areas, their graphical representations are quite similar. The area of computational neuroscience creates a bridge between these two phenomena and explains briefly the reason for similarity in terms of graphical representation. Additionally, this study gives information in a different perspective to understand the main reason for confidence in human behavior. According to this study, the main reason of self-confidence is not only related with the general knowledge of humans but also related with the amount of time it takes to reach a piece of information in a specific subject.

2. Dunning-Kruger Effect

The Dunning-Kruger effect, a theory received a Nobel Prize in 2000, was first proposed by Justin Kruger and David Dunning in 1999. In general, this theory examines the relationship between self-confidence and achievements with human behaviors [13, 14]. It is said that people who suffer from the Dunning-Kruger effect have high self-confidence in a subject despite having inadequate knowledge in that specific subject whereas people who do have knowledge don't declare this fact. The basic illustration of the Dunning-Kruger effect is shown in Figure 1.

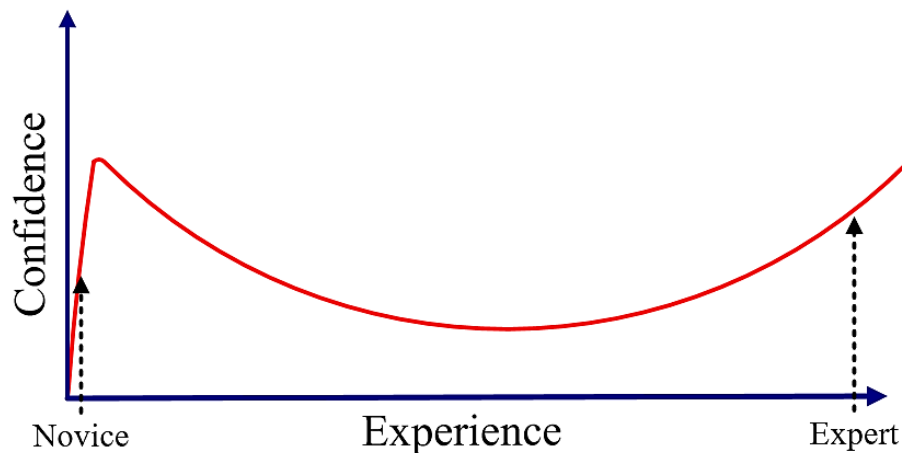


Figure 1 Dunning-Kruger confidence curve

In this study, the relationship between these two phenomena in the fields of machine learning and human behavior was investigated. The relationship between the amount of information and self-confidence is similar in various ways to the curse of dimensionality, an important phenomenon in machine learning. According to Bellman's curse of dimensionality, if there is a larger data set than there should be ("required") to solve a problem, this will decrease the performance of the solution to the problem.

3. Bellman's Curse of Dimensionality

The term of the curse of dimensionality was first introduced by Belman in 1961 [15,16]. According to this term, number of feature must be an optimal size to represent data in the storage area and increase the number of the feature do not always give the satisfying result on classification, clustering or regression. In general, the cost of the algorithm increases exponentially with the dimension of the information. This cost makes the calculation impossible or too difficult for large space dimensions.

The problems with increasing dimensions (d) in the dataset are as follows:

- Sparsity on the input becomes more obvious and the redundant part in the dataset increases rapidly.
- The decreasing performance of the algorithm dramatically depends on the increase of dimensions.
- The similarity measurement or classification process of the algorithms becomes redundant due to input and test data becoming the same as each other. The purpose of using a method such as a classifier or a similarity measurement becomes useless.
- Even though the computation and data storage ability of computers have evolved significantly since the term curse of dimensionality was first discussed, computation time and area of usage in the storage are still critical issues in the area of machine learning.

The curve of the curse of dimensionality is shown in Figure 2 and the classifier performance is represented via Equation 1.

$$\text{ClassifierPerformance} = \frac{\text{Correct ClassificationPercentage}(\%)}{\text{ComputationTime}} \quad (1)$$

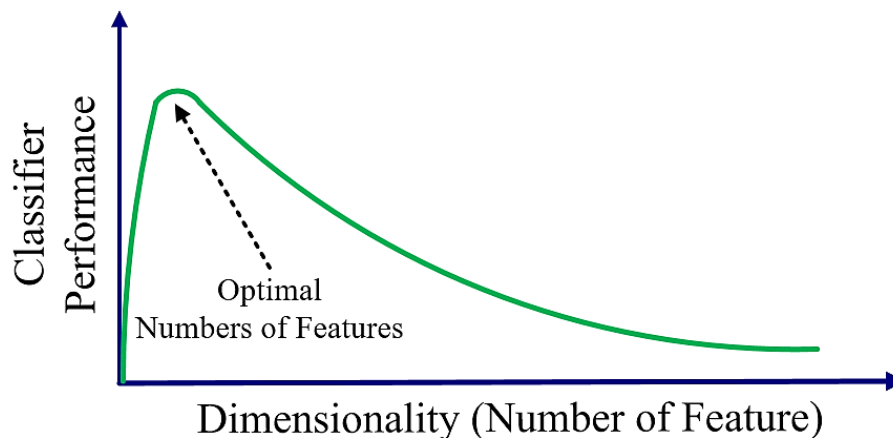


Figure 2 Bellman's curse of dimensionality

As shown in Figure 2, there is a number about an optimal number of feature for classifier performance and this peak point gives information about the optimal representation of features in a dataset with this classifier algorithm.

4. Human Problem Solving and Cognitive Science

According to recent studies in the area of neuroscience, the main reason for the decrease in cognitive ability is the size of information received over time. In a study by Ramscar et al., the processing speed of the brain was found to decrease over time with the size of the stored information and the cognitive model they proposed was tested with linguistic models for the simulation of the human brain [17]. The present study consolidated the relationship between the human brain and machine learning via classification algorithms. As noted at the beginning of this study, processing speed not only affects learning ability but also manipulates human behaviour such as confidence. Processing new information must have a computation cost. This cost is highly related to the size and location of information.

Old studies have suggested that the brains of adult individuals have become stable and do not develop over time. However, it has been observed with current studies in neuroscience that this situation is contrary to what is believed and the neural clusters of human memory are dynamically reconstructed especially at lower level clusters. In the basic studies in the field of human memory, it is foreseen that the human brain consists of three basic parts and these are as follow: after image, primary and secondary memory [18, 19]. Thanks to the developing cognitive and neuroscience today, human memory is predicted to consist of four categories: 1. sensory buffer memory (SBM), 2. short-term memory (STM), 3. long-term memory (LTM), and 4. T action buffer memory (ABM). These categories are classified according to data saving period and the type and size of the stored data [20]. For example, LTM is defined as a category of memory in which knowledge and skills are kept permanently and the storage ability of LTM is about limitless due to large number of neurons (at 100×10^9 level) and much more potential synaptic connections (at 108,432 level). The basics of human knowledge and long-term memory can be represented by an Object-Attribute-Relationship (OAR) model based on the synaptic structure of human memory [21]. The model of OAR is representing dynamic and hierarchical information clusters which are stored in the memory and this model provides to develop a logical model of cognitive knowledge bases. The mathematical model of OAR model which is stored at LTM is given in Equation 2 [18].

$$OAR \triangleq (O, A, R) \quad (2)$$

where O is a finite set of objects identified by unique symbolic names, A is a finite set of attributes for characterizing each object, and R is a set of relations between objects and attributes. The illustration of OAR model is shown in Figure 3.

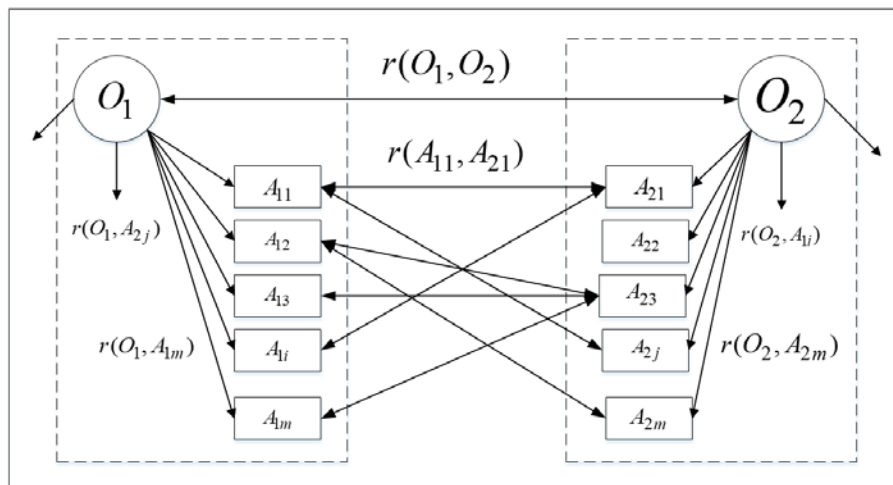


Figure 3 The OAR model of human memory architecture

As shown in the OAR model, the memory architecture of human is consist of the enormous number of networks. These memory structures are similar to the neuron structure, which is the basic building block of the process a, dynamically solving problems. Because neuron structures are also a network structure that contains many interactive and complex relationships with each other. These neuron structures are at the basis of artificial neural networks that form the basis of today's machine learning.

In this study, radial basis function (RBFN) neural networks were used to model the classification process of the human brain [22]. RBF neural networks are generally used to create a nonlinear function between inputs and outputs and the topology of RBF is a form of three-layered feed-forward neural networks. The first layer is defined as the input layer, the hidden layer consists of structures containing a large number of nonlinear RBF activation functions and the third layer is defined as the output layer. The

preliminary structure that separates the RBF neural network from traditional multilayer neural networks is the structure of the activation function and Gaussian transfer function is used as an activation function in RBFN structure [23]. The illustration of Gaussian transfer function is given in Figure 4. The value of radius (σ) is determines the behavior of the model to be created on RBFN. The larger value of selection on sigma causes underfitting, while the low sigma value leads to overfitting on the stage of function approximation or classification.

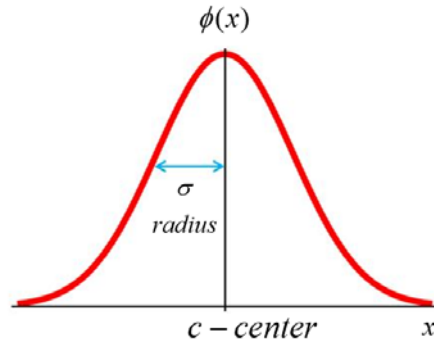


Figure 4 Gaussian Activation Function

The structure of RBFN is given in Figure 5.

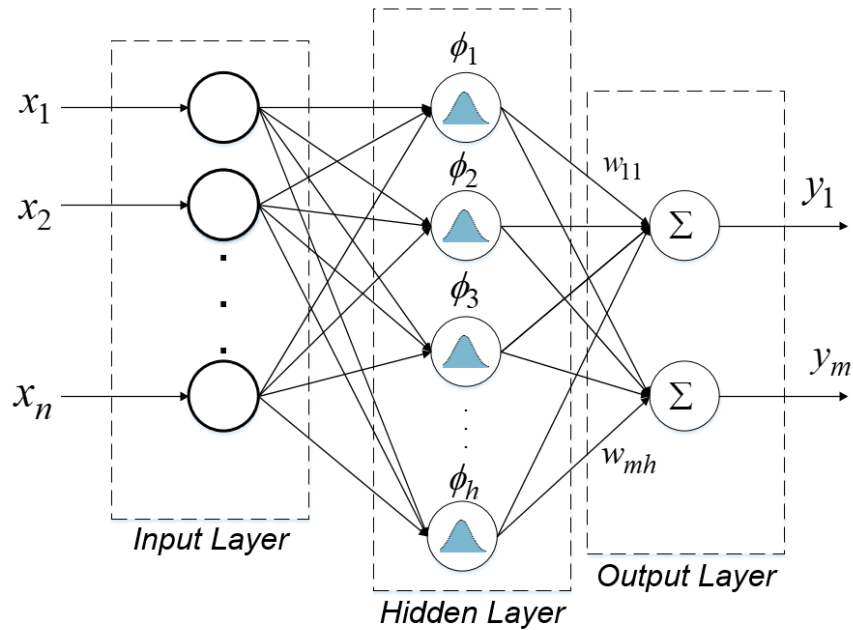


Figure 5 Generic illustration of RBFN structure

The output of activation function ϕ_i is given in Equation 3 and x is equal to input value, c_i is i th center, σ_j is spread of j th transfer function, and $\| \cdot \|$ is Euclidean norm.

$$\phi_i(\|x - c_i\|) = \exp\left(-\frac{\|x - c_i\|^2}{2\sigma_j^2}\right) \quad (3)$$

The output (y_k) of the node m at the output layer can be determined as Equation 4.

$$y_k = \sum_{j=1}^h w_{jk} \phi_j(x) \quad (4)$$

RBFN is a popular alternative to MLP structure due to rapid training process and a simple structure and it has the capability of universal approximation.

5. Image Classification and Curse of Dimensionality

In this study, the optimal numbers of the feature selection process were realized with an image classification problem. In this problem, gender classification was performed with principal component analysis (PCA) and Radial Basis Function (RBFN) [24,25]. The proposed technique was tested using Aberdeen facial images from Psychological Image Collection at Stirling (PICS image database) [26]. In this dataset, there were 130 input images for each gender and 70 of them were selected for training image while the rest of the images were used for the testing process. The detection of classifier performance was obtained as a result of the cross-validation of the dataset. The flowchart of the classification process is shown in Figure 6.

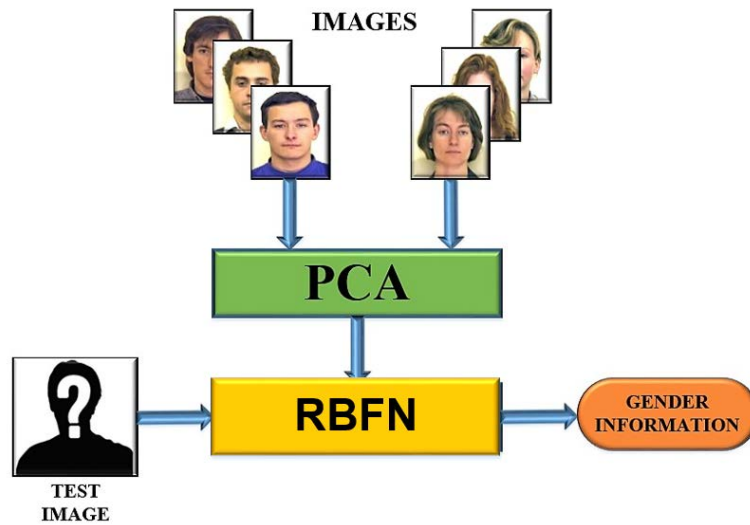


Figure 6 Gender detection from the images

The error percentage of gender classification according to the number of features dimension is shown in Figure 7. As can be seen from the figure, the increase of the selected feature number is meaningless after the first four features after which this kind of selection only causes a decreasing performance in the algorithm.

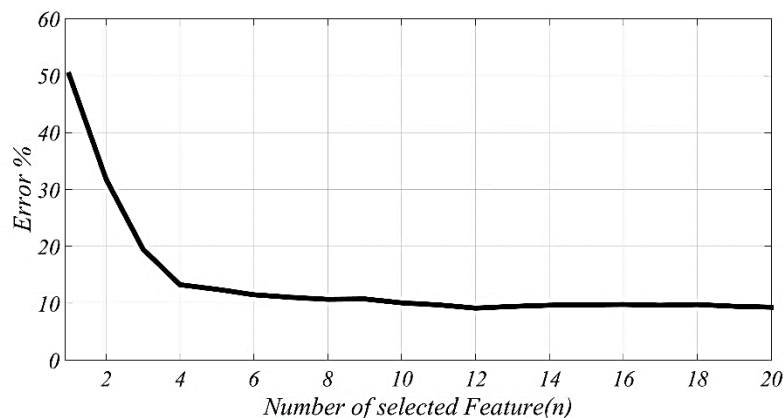


Figure 7 Percentage of error

The computation time of the algorithm versus the selected feature size is shown in Figure 8. As expected, the total computation time of the classifier increased with feature size.

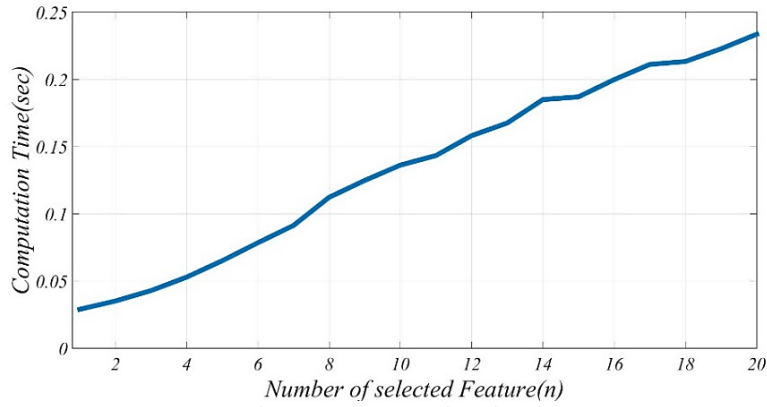


Figure 8 Computation time

The classifier performance of the gender classification process was performed according to Equation 1. As a result of this computation, a specific curve very similar to the curve of dimensionality curve was obtained as shown in Figure 9.

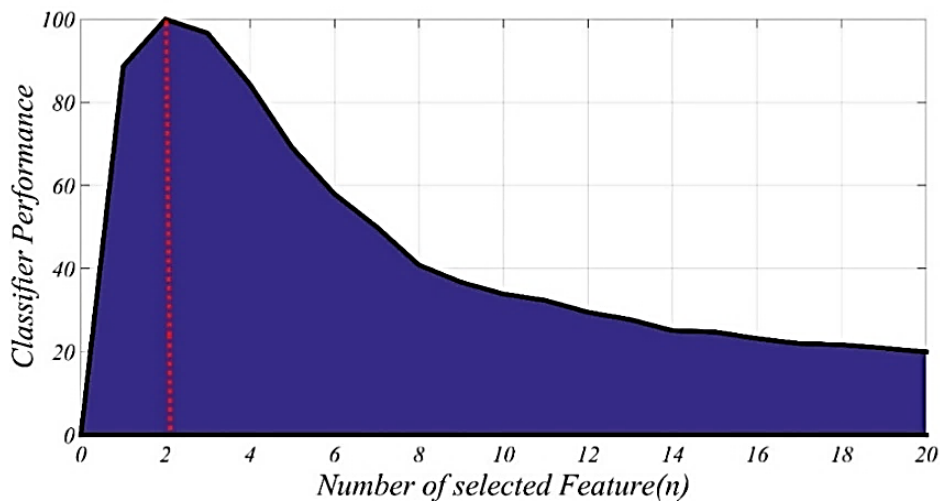


Figure 9 Classifier performance

The selection of the first two components is necessary to achieve optimal classification performance for this gender classification process. Additionally, this curve is also very similar to the curve of the Dunning-Kruger effect. In the image classification process, using the RBFN structure, it was provided to simulate the classification skills of people on computer systems [27]. The results of studies in the field of cognitive science are supportive of this similarity. Hong et al conducted analyzes on students to examine the relationships between cognitive parameters and self-confidence in the area of education [28]. In the study in which 6 different hypotheses were analyzed and the partial least squares (PLS) method was used to analyze the research hypotheses. The hypothesis which is related to our research paper is 6th hypothesis and the hypothesis question is analyses the relationship between self-confidence enhancements (SCE) relevant to learning progress (LP) is examined in the 6th hypothesis. Path graphic showing the relations obtained as a result of the analysis is given in Figure 10.

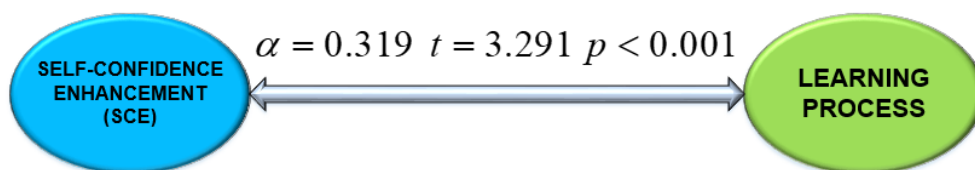


Figure 10 Relationship between SCE and LP

As a result of the analysis, there is a positive correlation between SCE to the learning process and this result supports the relationship between the self-confidence and performance of data access as we suggest in this article.

6. Results

In this study, the source of confidence in human behaviour was explained according to the curse of dimensionality. Image gender classification and a number of component selections were performed to obtain classifier performance. The obtained curve was found to be very similar to the curse of dimensionality. This study explained the reason for confidence being not only inadequate amounts of information about a specific subject but also the processing time information on the brain. This study provides a new perspective to the Dunning-Kruger effect in psychology and also uses neuroscience to explain the source of confidence. In this study, the storage area was defined as the human brain. This enhancement is the increase in total area and time it takes to reach specific information on the brain like a storage area on computer system. This means there must be a way to the optimal representation of data on the brain to increase confidence. The phenomena about machine learning and psychology are not as far apart as it is thought. In fact, they are highly correlated with each other and can be used to describe the origin of human behavior. Human memory structure and types are explained. The results of the studies supporting the proposed relationship were analyzed within the scope of the article. As a result of this study, processing speed and the total time it takes to reach correct information are highly correlated with confidence in human behavior and this result was explained in this multidisciplinary study utilizing information theory, machine learning and neuroscience.

Conflicts of Interest

No conflict of interest was declared by the authors.




References

- [1] S. Finger, *Origins of neuroscience: a history of explorations into brain function*. Oxford University Press, USA, 2001.
- [2] J. Ward, *The student's guide to cognitive neuroscience*. Psychology Press, 2015.
- [3] P. Dayan and F. A. Laurence, "Theoretical neuroscience: computational and mathematical modeling of neural systems", *Journal of Cognitive Neuroscience*, vol.15, no.1, pp.154-155.
- [4] T. C. Napier, J. C. Corvol, A. A. Grace, J.D. Roitman, J. Rowe, J., V. Voon and A. P. Strafella, "Linking neuroscience with modern concepts of impulse control disorders in Parkinson's disease", *Movement Disorders*, vol.30, pp.141-149, 2015.
- [5] M. Filippi, *fMRI techniques and protocols*. Humana press, 2016.
- [6] A. Poldrack Russell, "The role of fMRI in cognitive neuroscience: where do we stand?", *Current opinion in neurobiology*, 18223-227, 2008.
- [7] U. Goswami, "Neuroscience, education and special education", *British Journal of Special Education*, vol.31, pp.175-183, 2004.
- [8] Y. Immordino, M. Helen, A. Damasio, "We feel, therefore we learn: The relevance of affective and social neuroscience to education", *Mind, brain, and education*, vol.1, pp. 3-10, 2007.

- [9] B. Aberšek, *Cognitive science in education and alternative teaching strategies*. Cambridge Scholars Publishing, 2017.
- [10] B. Sung, N. J. Wilson, J. H. Yun, and E. J. LEE, “What can neuroscience offer marketing research?”, *Asia Pacific Journal of Marketing and Logistics*, pp. 4-23, 2019.
- [11] S. Ullman, “Using neuroscience to develop artificial intelligence”, *Science*, vol. 363, no. 6428, pp. 692-693, 2019.
- [12] G. Deco, E. T. Rolls, *Computational neuroscience of vision*. Oxford university press, 2007.
- [13] D. Dunning,” The Dunning–Kruger effect: On being ignorant of one's own ignorance”, *In Advances in experimental social psychology*, vol.44, pp. 247-296, 2011.
- [14] G. Pennycook, R. M. Ross, D. J. Koehler and J. A., Fugelsang,” Dunning–Kruger effects in reasoning: Theoretical implications of the failure to recognize incompetence”, *Psychonomic Bulletin & Review*, vol.24, no.6, pp. 1774-1784, 2017.
- [15] N. Altman, M. Krzywinski, ”The curse (s) of dimensionality”, *Nat Methods*, vol.15, pp. 399-400, 2018.
- [16] V. Zelenyuk, “Aggregation of inputs and outputs prior to Data Envelopment Analysis under big data”, *European Journal of Operational Research*, vol.282, no.1, pp. 172-187, 2020.
- [17] M. Ramscar, P. Hendrix, C. Shaoul, P. Milin, H. Baayen, “The myth of cognitive decline: Non-linear dynamics of lifelong learning”, *Journal of Experimental Psychology*, pp.5-42, 2014.
- [18] Y. Wang, and V. Chiew, “On the cognitive process of human problem solving”, *Cognitive systems research*, vol.11, no.1, pp.81-92, 2010.
- [19] H. Kuai, X. Zhang, Y. Yang, J.Chen, B. Shi, N. Zhong, “THINKING-LOOP: The Semantic Vector Driven Closed-Loop Model for Brain Computing”, *IEEE Access*, vol.8, pp.4273-4288, 2020.
- [20] Y. Wang et al., “Brain-Inspired Systems: A Transdisciplinary Exploration on Cognitive Cybernetics, Humanity, and Systems Science Toward Autonomous Artificial Intelligence”, *IEEE Systems, Man, and Cybernetics Magazine*, vol. 6, no. 1, pp. 6-13, Jan. 2020
- [21] R. Shivhare, A. K.Chelukuri, and J. Li, ,”Establishment of cognitive relations based on cognitive informatics”, *Cognitive Computation*, vol.9, no.5, pp.721-729, 2017.
- [22] H. Faris, I. Aljarah and S. Mirjalili, *Evolving radial basis function networks using moth–flame optimizer*. In Handbook of neural computation, Academic Press, pp. 537-550, 2017.
- [23] C. Zhang, H. Wei, L. Xie, Y. Shen, and K. Zhang, “Direct interval forecasting of wind speed using radial basis function neural networks in a multi-objective optimization framework”, *Neurocomputing*, vol. 205, pp.53-63, 2016.
- [24] I.T. Jolliffe and J. Cadima, “Principal component analysis: a review and recent developments”, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol.374, no.2065, 20150202, 2016.

- [25] G. R. Naik, *Advances in Principal Component Analysis: Research and Development*. Springer, 2017.
- [26] I. Craw, The aberdeen set of Psychological Image Collection at Stirling (PICS) database (PICS), URL: <http://pics.psych.stir.ac.uk/zips/Aberdeen.zip> [Accessed: 19-July-2020].
- [27] N. Kriegeskorte, "Deep neural networks: a new framework for modeling biological vision and brain information processing", *Annual review of vision science*, vol.1, pp.417-446, 2015.
- [28] J. C. Hong, , M. Y. Hwang, , K. H. Tai and C. R. Tsai, "An exploration of students' science learning interest related to their cognitive anxiety, cognitive load, self-confidence and learning progress using inquiry-based learning with an iPad", *Research in Science Education*, vol.47, no.6, pp.1193-1212, 2017.

The Efficacy of Frontal and Temporal Lobes in Detecting 2D&3D Video Transition Using EEG

 Negin MANSOURI¹,  Mesut MELEK (Masoud MALEKI)²,  Temel KAYIKCIOGLU³

¹Department of Electrical and Electronics Engineering, Karadeniz Technical University, Trabzon, Turkey; negin.manshouri@ktu.edu.tr; <https://orcid.org/0000-0001-5297-5545>; 05535791640

² Department of Electronics and Automation, Gumushane University, Gumushane, Turkey; mesutmelek@gumushane.edu.tr; <https://orcid.org/0000-0002-7152-7788>

³Department of Electrical and Electronics Engineering, Karadeniz Technical University, Trabzon, Turkey; tkayikci@ktu.edu.tr; <https://orcid.org/0000-0002-6787-2415>

Received 18 June 2020; Revised 20 July 2020; Accepted 28 July 2020; Published online 28 August 2020

Abstract

Due to the three-dimensional (3D) structure of the human eye, 3D technology was used in this research. Transition to 2D and 3D analysis is important, claiming that binocular vision will lose dimension during fatigue. Thus, a single-stream video consisting of random 2D&3D parts was watched by nine volunteers. The dynamic behavior and power spectral density (PSD) of the volunteers' brain signals were considered for a comprehensive quantitative analysis. The dominant EEG bands and time intervals were selected in 2D to 3D (2D_3D) and 3D to 2D (3D_2D) transitions by analyzing power differences based on short-time Fourier transformation (STFT). Taking into account this information, applying effective feature extraction and classification techniques, the behavioral patterns of channels representing the brain lobes of the volunteers were analyzed. Hjorth parameters and statistical methods were used as feature extraction methods. The k-nearest neighbors (k-NN) and linear discriminant analysis (LDA) algorithms were applied to classify 2D_3D and 3D_2D transitions. The results revealed that, thanks to the comprehensive classification analysis of the 2D_3D and 3D_2D transitions, the change in the activity power of the brain cortex can be represented. The dominance of the temporal and frontal lobes between the channels representing these transitions cannot be excluded.

Keywords: EEG, 2D&3D, transition, classification, anaglyph

EEG Kullanarak 2D&3D Video Geçişinin Tespitinde Frontal ve Temporal Lobların Etkinliği

Öz

İnsan gözünün üç boyutlu (3B) yapısı nedeniyle, bu araştırmada 3B teknolojisi kullanılmıştır. Binoküler görmenin yorgunluk esnasında boyut kaybedeceğini iddia ederek 2B ve 3B'ye geçiş analizi önemlidir. Böylece rastgele 2B ve 3B parçalardan oluşan tek akışlı bir video dokuz gönüllü tarafından izlenilmiştir. Gönüllülerin beyin sinyallerinin dinamik davranışı ve güç spektral yoğunluğu (GSY), kapsamlı bir nicel analiz için dikkate alınmıştır. Baskın EEG bantları ve zaman aralıkları, kısa zamanlı Fourier dönüşümüne (KZFD) dayalı güç farklılıklarını analiz ederek 2B'den 3B'ye (2B_3B) ve 3B'den 2B'ye (3B_2B) geçişlerde seçilmiştir. Bu bilgiler dikkate alınarak, etkili özellik çıkarımı ve sınıflandırma teknikleri uygulanarak, gönüllülerin beyin loblarını temsil eden kanalların davranış kalıpları analiz edilmiştir. Öznelik çıkarma yöntemleri olarak Hjorth parametreleri ve istatistiksel yöntemler kullanılmıştır. 2B_3B ve 3B_2B geçişlerini sınıflandırmak için en yakın komşular (e-YK) ve doğrusal diskriminant analiz (DDA) algoritmaları uygulanmıştır. Sonuçlar, 2B_3B ve 3B_2B geçişlerinin kapsamlı sınıflandırma analizi sayesinde beyin korteksinin aktivite gücünün değişiminin temsil edilebileceğini ortaya koymuştur. Bu geçişleri temsil eden kanallar arasındaki temporal ve frontal lobların baskınlığı göz ardı edilemez.

Anahtar Kelimeler: EEG, 2B&3B, geçiş, sınıflandırma, anaglif

1. Introduction

Studies in 2D&3D areas have been rapidly developing because these areas form real feelings and a sense of sensibility in people. This development is manifested in different domains of the scientific world, from the medical to the entertainment and education sectors. These areas of application have been taken into account by researchers in the category of entertainment [1], health [2]-[3], industry, and architecture [4]. Thanks to the excitement created by this technology, it attracts a great deal of attention from young people [5]. In addition to being the focus of attention of 3D technology, it has some disadvantages. In case of frequent use, it can lead to habits as well as health problems such as headaches, eye fatigue, and physical discomfort [6]. 3D technology has a long history, spanning from the invention of stereoscope in the 18th century to the 3D TV in the 20th century [7]-[8].

Some studies have evaluated the visual differences resulting from watching 2D&3D images using the electroencephalogram (EEG) [9]-[10]. Measurement of ionic voltage fluctuations of brain neurons is defined as EEG. The EEG method is preferred because it is the most important and reliable physiological measurement in evaluating mental activities. This method is the physiological method used to record the electrical activity produced by the brain through electrodes placed on the scalp surface. By using the frequency region of signals, especially frequency bandwidths, EEG enables this information to help reveal the functional states of the brain [11]-[12].

In order to grasp the logic of 3D technology, the human eye structure must be taken into account. It is also useful to learn how the eyes and the brain work. People see the world with eyes looking in two directions at a distance of about 6 cm [13]. Unlike some animals [14], people have the ability to see binoculars due to the eye structures. Each eye has a certain field of view when the other eye is closed. Although the visual information in the environment can only be perceived with one eye, the fields of vision of the eyes overlap. This information is transmitted to a region of the brain behind the head, and synthesizes overlapping visual information to create a single coordinated image. In summary, what we see is the result of signals sent from the eyes to the brain. In general, the brain receives signals from both eyes (ocular) at the same time. The information contained in the signal for each eye is slightly different, and with a well-functioning binocular vision, the brain can understand these differences; Also it can be used to assess distances and coordinate eye movements [15]-[16]. This difference in the retina is known as binocular disparity. In 1838, Wheatstone showed that these subtle differences between retinal images are detected by the brain and provide a powerful clue to the depth structure of the world around us [16]. This ability is known as (binocular) stereoscopic or stereo vision. It is also often described as a 3D vision. One of the reasons why binocular vision is so important is that it allows to perceive depth and relationships between objects.

EEG analysis of 2D and 3D technology can generally be classified as eye fatigue detection [17], game analysis [18], [19], detection of power spectrum differences in brain waves [20], [8], and investigation of the effects of stereoscopic disparity on event-related potentials [21]. Few of these studies have focused on a brief look analysis of brain signals in 2D and 3D video watching instead of comprehensive analysis, and none of them has examined the 2D and 3D transitions of video watching modes.

Taking into account the basis of 2D and 3D technology, until 2011 very few studies in this area actually focus on the quantitative benefits of 3D visualization [22]. Many studies have proved that despite the complexity of 3D technology, it performs better than 2D in terms of model visualization [23]. Thanks to the excitement and realistic feelings created by 3D technology, its expansion continues rapidly in today's research. Therefore, it is important to understand the comprehensive analysis of brain activation patterns underlying 2D and 3D studies [24].

The viewing of the television screen from different distances and how this distance affects the human visual system was also evaluated in detail [17]. Time interval and screen distance are among the important factors in 3D watching. The visual fatigue survey for 2D versus 3D has been carried out in [25]. By using EEG-based 2D and 3D concentrated games, a biomedical feedback system has been developed in [3]. In this study, by utilizing 2D and 3D technology and nonlinear methods concentration check on human individuals was performed. In five physiological status analysis of [18], EEG signal complexity was evaluated using two parameters. In the case of openness of the eye and in 3D mode,

complexity increasing was observed. The depth degree of 3D images and the relationship between visual disturbance were also explored in detail in [21]. As a result of the depth of 3D technology, a considerable increase in ERP components (P1) was observed compared to 2D. The amplitude of this component can be adjusted by changing the degree or perception of depth. Brain areas sensitive to binocular vision and perception of depth were analyzed in [9], [26] based on Brodmann areas (BA). These regions were determined as temporal, parietal and occipital.

This study is a basic step towards evaluating our hypothesis which is based on the human eye structure. In this hypothesis, the analysis of the moment of transition to 2D_3D and 3D_2D is important, claiming that 3D vision will lose dimension during fatigue. Behavior analysis and classification of brain signals in these transitions are the main goals of this research. In the process of watching a single anaglyph video consisting of random 2D and 3D parts, electrophysiological responses were discussed in detail. In a brief description of this detailed analysis, the power spectrograms of the electrophysiological responses achieved on the presentation of the 2D_3D and 3D_2D transitions were obtained using the STFT. Thus, it is possible to get an idea about the behavior of each channel by considering the power spectrum difference of transition states. In addition, based on this visual information, important frequency bands, time windows, and channels that maximize the power difference of the transitions were brought to the fore. In order to obtain an acceptable performance classification technique, effective feature extraction methods were applied by selecting dominant frequency and time intervals.

By evaluating this hypothesis, the effects of 2D_3D and 3D_2D transitions on brain signals in terms of time, frequency, and channel can be determined in a single video of 2D and 3D parts. In this way, it is hoped that the researchers will be able to provide information about the parameters to be considered in the studies on 2D and 3D technology (visual processing, classification, early detection of drivers falling asleep, etc.). The results of the study will contribute to the literature on the comprehensive brain dynamics underlying 2D and 3D technology and the parameters affecting these dynamics.

2. Material and Method

In this section, information was given about volunteers, video production, EEG recording, data processing, and analysis methods.

2.1. Volunteers

Nine volunteers (four women and five men) without neurophysiological disorders participated in the experiments. The average age and standard deviation (SD) were 34.12 and 2.072 years, respectively. The experiments were carried out according to the experimental procedures No. 24237859-806 approved by the Ethics Committee of the Faculty of Medicine, Karadeniz Technical University. Volunteers were informed about this protocol in advance, and their written consent was obtained.

2.2. Video Production

A 135-second video was watched by the volunteers five times. This video consisted of 8-second 2D and 3D random parts. In the adjustment of this anaglyph video, the 3D version of the SAW video [27] was converted to 2D using the Xilisoft 3D Video Converter program [28]. Later, the 3D version was converted to a 3D anaglyph format using the IQmango 3D converter program [29]. The resulting parts were combined by applying the Idoos video program [30]. The selected glasses must be independent of any imaging system to view 2D and 3D video segments in a single video, considering TV features [20]. Thus, anaglyph glasses [31] were found to be suitable for this scenario according to the working principle.

2.3. EEG Recording

A 32-channel Micromed Italy Brian Quick device was used for EEG data recording. The experimental environment and the system are presented in Figure 1. The EEG cap consisted of 21 electrodes, and the

animation graphic of this cap is presented in Figure 2. In the experiments, the Cz electrode was determined as a ground electrode. EEG recordings were made by trying to keep the electrode resistance values below 10 K Ohms on average. Considering the EEG bandwidth of 0-100 Hz, sampling frequency (fs) was chosen as 512 Hz to ensure sufficient time resolution and prevent data exchange in long-term recordings.

The experiments were conducted in a dimly lit room. The subjects were asked to sit in a comfortable chair about 85 cm from the TV stand, relax, and minimize body movements during the experiment. Each recording lasted 135 seconds and was repeated five times. There was insulation isolation in the EEG system used to protect each experiment from the electrical impact. In the device used, a filter that passes between 0.1-120 Hz was applied to clear the recorded EEG signals from internal and external artifacts. In addition, a 50Hz notch filter was employed to eliminate mains noise.

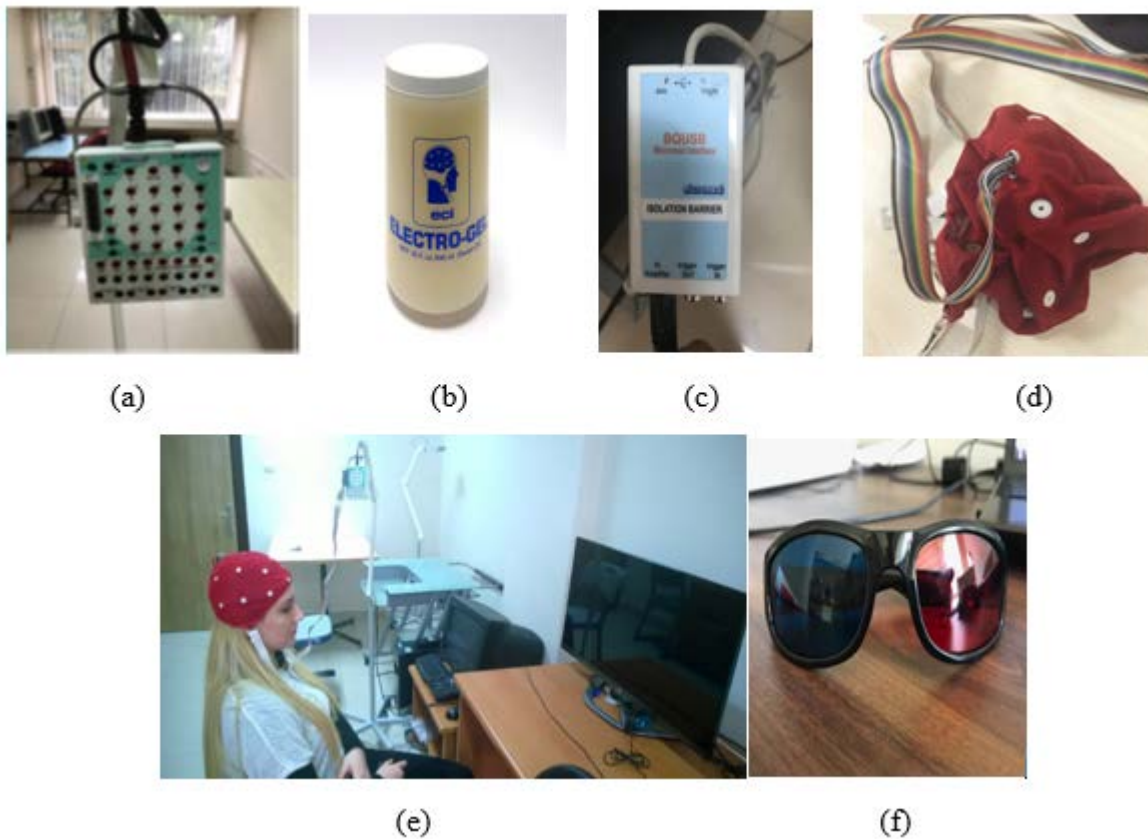


Figure 1 EEG system a) EEG amplifier, (b) gel, (c) EEG isolation, (d) EEG cap, (e) EEG recording area, (f) Anaglyph glasses

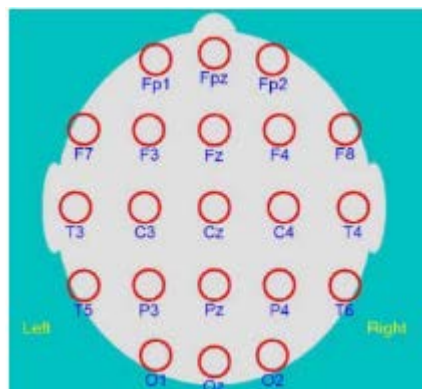


Figure 2 The animation graphic of the EEG cap

2.4. Scenario

In this scenario, the movie consisting of random 2D and 3D videos was watched by the volunteers. The aim was to analyze and classify 2D and 3D transitions. An EEG recording example of the sequence of the video, consisting of 8-second 2D and 3D random parts, is illustrated in Figure 3. In Figure 3, the orange-colored arrows represent the 2D_3D transitions, and the green arrows the 3D_2D transitions. In the transition analysis, there is a total of nine transitions (five 2D_3D and four 3D_2D transitions). The window length of five seconds was taken into account in terms of epoch category. This category is presented in Table 1. In this table, according to Figure 3, window lengths of 5 seconds were selected for analysis, considering the critical transitions from 2D to 3D (2D_3D) and from 3D to 2D (3D_2D). As mentioned before, each recording was repeated five times. Thus, due to the window length of 5 seconds and the existence of nine epochs, a total of 45 epochs was included in the analysis.

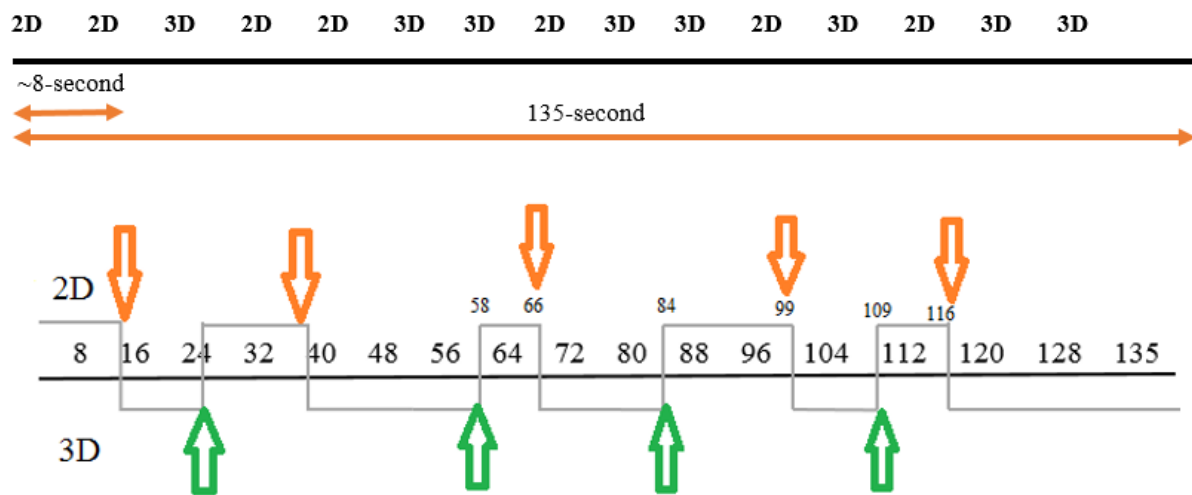


Figure 3. Details of the video

Table 1. Epoch category

5-second window length from 2D_3D	5-second window length from 3D_2D
13-18, 37-42, 63-68, 96-101, 113-118	21-26, 55-60, 81-86, 106-111

2.5. Preprocessing

Following the EEG signal recording, the recordings were split into two sub-data sets, namely 5-second window lengths covering 2D_3D and 3D_2D transitions. Since there were nine transitions in each of the five trials, 45 epochs were taken into account. In addition, each sub-data set represents 20 channels of nine volunteers. To take the first step of data analysis, the average of five trials was calculated in order to get an idea of the choice of dominant bands and to minimize noise. In addition to applying a notch filter at 50 Hz to clear the network noise [32], a third-order bandpass Butterworth filter to clean the noise signal in the frequency range of 1-55 Hz was employed. The order of the filter was chosen by trial and error. All the data had to be normalized in order to eliminate this problem and provide the opportunity to compare different volunteers. The z-score normalization method [33] was adopted to minimize amplitude changes and enhance the classification performance.

2.6. Time-Frequency Analysis Process and Band Selection

Time-frequency analysis was applied to get a more effective idea about the processed data. Thus, in the 2D_3D and 3D_2D transition analysis, the time intervals representing the maximum difference of PSD and the dense frequency range were taken into account. In other words, the behavior change between 2D_3D and 3D_2D transitions is well demonstrated. The spectrogram method based on STFT [34] was used to accomplish this. The STFT is an expanded version of the traditional Fourier analysis used for the analysis of non-stationary data. This technique analyzes the local frequency content of the signal

with floating windowing [35]. By adjusting the width of this window, valuable information about unpredictable EEG signals can be obtained. Due to the unknown nature of the EEG signals, the soft-behavior Hanning window was chosen. One of the suitable features of this window is that it has a wide peak unlike its narrow side lobes. Thus, in order to capture the related time and frequency resolution, taking into account the frequency range desired, the window size was considered as 512 sample lengths. The overlapping of the window was noted as 'window size -1'[8].

The transition effect of 2D_3D and 3D_2D anaglyph video watching was evaluated on the power of brain signals for different frequency bands. To gain the goal, the average EEG power spectrogram of all the epochs belonging to the nine volunteers in each sub-data set was achieved separately for 20 channels. As a result, when the average EEG power spectrograms were investigated for the effective time windows, differences were observed between the average power spectrograms of the 2D_3D and 3D_2D transitions on the dominate frequency and channels. Based on the spectrogram graph, in 2D_3D and 3D_2D transition analysis, the delta band was determined as the dominant band in approximately all EEG channels representing brain lobes during the comparison phase. Considering the epoch averages, this band selection was calculated according to the average results of the volunteers in each channel. In the next step, taking into account the dominant band and time intervals obtained from this section, the feature extraction process was performed for 5-second epochs by selecting 1-1.5 and 1.5-3-second time intervals at 2, 3, and 4Hz frequencies.

2.7. Feature Extraction

The main purpose of the feature extraction method in EEG signals is the extraction of features that effectively reflect the frequency, time, and spatial properties of the signal. Here, selecting effective values of time-frequency windows that best represent the differences between the power spectra showing the 2D_3D and 3D_2D transitions can be defined as the feature extraction method. Considering the STFT-based spectrogram graph, two feature extraction methods were utilized in the dominant band and time intervals. One is the SD and max of statistical functions, and the other is Hjorth parameters [36]. Three parameters of the Hjorth method are activity, mobility, and complexity. The mathematical representation of these parameters is shown respectively in Equations 1, 2, and 3.

$$\mathbf{Activity} = \frac{1}{N} \sum_{n=0}^{N-1} [x(n) - \bar{x}]^2 = \mathbf{variance}(x(n)) \quad (1)$$

$$\mathbf{mobility} = \sqrt{\frac{\mathbf{Activity}(\dot{x}(n))}{\mathbf{Activity}}} = \sqrt{\frac{\mathbf{variance}(\dot{x}(n))}{\mathbf{variance}(x(n))}} \quad (2)$$

$$\mathbf{Complexity} = \sqrt{\frac{\mathbf{variance}(\dot{x}(n))\mathbf{variance}(x(n))}{\mathbf{variance}(\dot{x}(n))^2}} \quad (3)$$

where $x(n)$ and $\dot{x}(n)$ are the first and second derivatives of the signal, respectively. While these three parameters contain information on the frequency spectrum of a signal, they also help analyze the signals in the time domain. In addition, lower computational complexity can be achieved by their use.

Ultimately for each channel, in the SD and max feature extraction method, considering the dominant frequency and time windows, the feature vector dimension was determined as 3*2*2 and in the Hjorth method as 3*3. The band selection and feature extraction method is summarized in Figure 4 as block diagram representation.

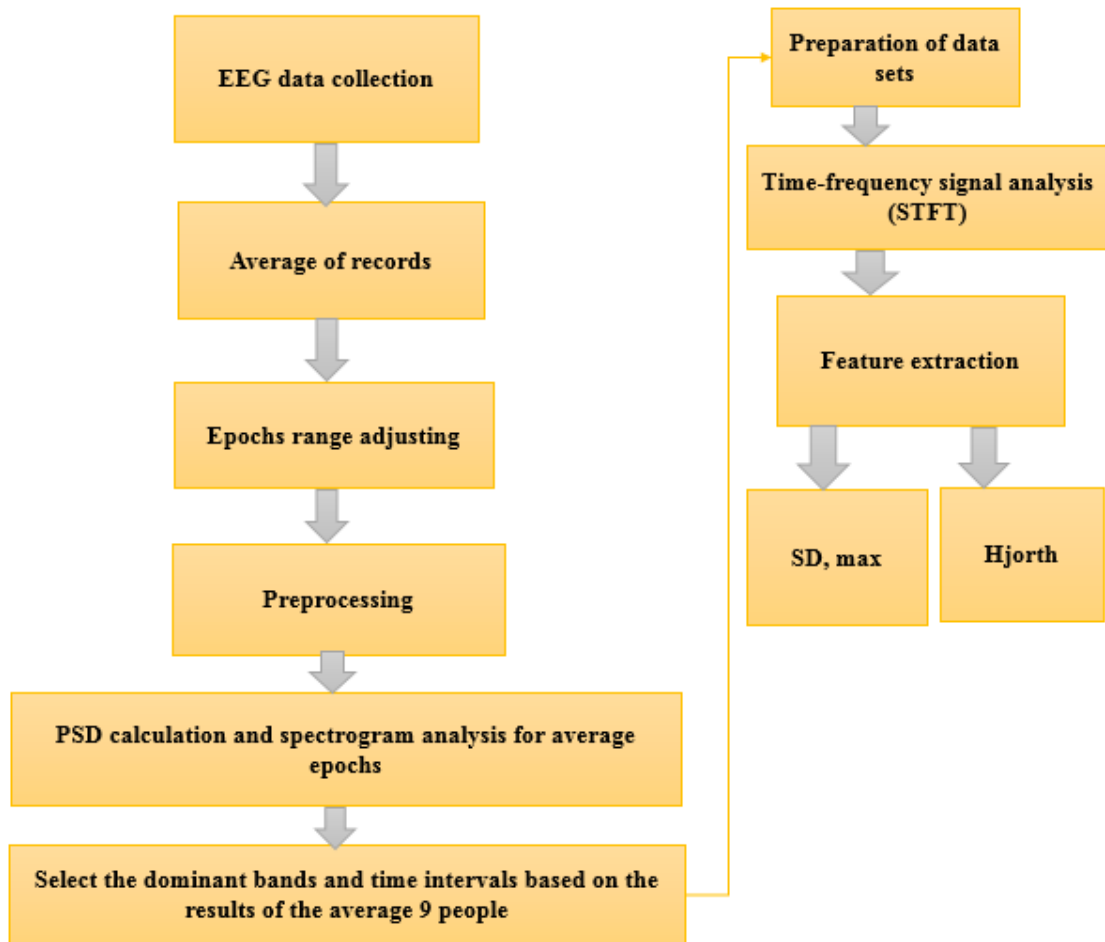


Figure 4 Flow diagram summarizing band selection and feature extraction method

2.8. Classification Techniques and Performance Evaluation of Classification

To achieve the goals of this study, namely 2D_3D and 3D_2D transition classification, k-NN and LDA classification techniques were employed. Since this study has two classes, the chance level is 50%. Additionally, sensitivity and specificity [16] were used to evaluate classifier performance and classification accuracy. The basic definition of these techniques is briefly presented below.

The supervised and simple k-NN technique that is sensitive to the local structure of the data is also used in pattern recognition for different classification and regression processes [27]. This algorithm has a lazy learning mechanism, and it tries to make estimations by memorizing the training dataset instead of learning it. During the estimation process, it looks for the nearest neighbors, taking into account the entire data set [28]. The class of the new sample is determined by looking at the closest k neighbor's class according to the distance in the learning set. In this study, these distances were calculated using the Euclidean distance function. Setting the k or neighborhood parameter directly affects the performance of this algorithm. If k is too small, it affects noise sensitivity, and if it is too large, it affects the neighbors of other classes. There is no clear method for choosing k. As used in this study, in general, the optimum k is selected by using the cross-validation method.

The LDA algorithm is used in machine learning and pattern recognition to separate two or more classes of problems. Its basic principle is based on the linear combination of features. This algorithm actually reveals the modeling differences of classes by moving the high-dimensional space features to a smaller size [29].

To evaluate the classification result validation, K-fold cross-validation was applied. K-fold cross-validation was employed to evaluate the classification result validation. In this study, the K value was chosen as 10. After cross-validation, the k value of k-NN was obtained. Here, class 2D_3D and 3D_2D

were defined as the positive and negative sample, respectively. To move on to the classification stage, the epochs of each class were randomly divided into two test and training sets. The classification process was repeated 20 times to avoid unlucky splitting.

3. The Research Findings and Discussion

If we want to summarize the outcome of the time-frequency analysis of EEG data, the difference of the average power spectra based on the average epochs in the evaluation of 2D_3D and 3D_2D transitions was shown for 20 different channels. Thus, dominant time intervals and effective frequency band were determined in this transition classification. As a result, the delta and 1-1.5 and 1.5-3 seconds time intervals were the best criteria representing the power difference. The k-NN classification results of 9 volunteers (V) in each channel were shown in Tables 2 and 3 for SD, max and Hjorth feature extraction methods, respectively. Likewise, LDA classification results were presented in Tables 4 and 5, respectively.

When looking at the classification results of 9 volunteers on 20 channels, it is not overlooked that the LDA and Hjorth methods are generally more successful. With a general view in terms of brain lobes, the frontal, temporal and partially parietal lobes appear to be highly successful. Considering the general average classification result in the LDA classification algorithm, T3 and Pz were chosen as the channels with the best results in SD, max and Hjorth methods with success of 66.67 % and 66.25 %, respectively. Since this study consists of two classes, it is observed that the classification results are low. In this case, it is necessary to increase these results by applying another method.

Table 2 k-NN classification results for nine volunteers in each channel, and the average classification accuracy (SD, max)

	V1	V2	V3	V4	V5	V6	V7	V8	V9	Average
Fp1	60	55	50	65	55	60	60	60	60	58.33
Fpz	50	50	55	60	60	70	65	65	45	57.78
Fp2	45	60	55	65	65	60	70	65	50	59.45
F3	50	65	55	50	65	50	55	65	70	58.33
F4	65	50	50	65	55	65	55	60	55	57.78
F7	55	50	55	60	45	65	50	60	55	55
F8	45	30	65	55	65	65	60	45	35	51.67
C4	65	45	60	55	65	50	45	45	60	54.45
C3	75	60	50	50	60	55	85	35	55	58.33
Fz	50	50	65	45	85	50	65	60	60	58.88
P3	65	55	60	60	70	70	55	55	70	62.23
P4	50	60	65	60	40	70	55	55	65	57.78
Pz	60	50	55	70	70	55	70	60	75	62.78
O1	50	60	55	50	60	45	55	55	60	54.45
O2	60	50	60	50	50	75	55	55	65	57.78
T3	60	60	45	45	45	55	50	60	60	53.33
T4	55	55	50	75	50	55	50	50	75	57.22
T5	60	55	40	55	50	60	65	50	50	53.12
T6	65	55	55	65	55	60	50	60	60	58.33
Oz	60	65	45	65	60	45	55	65	65	58.33

Table 3. The k-NN classification results of nine volunteers in each channel, and the average classification accuracy (Hjorth)

	V1	V2	V3	V4	V5	V6	V7	V8	V9	Average
Fp1	60	60	55	65	55	60	65	60	60	60
Fpz	55	70	50	60	60	70	65	65	45	60
Fp2	40	65	50	65	60	60	70	65	50	58.34
F3	65	55	55	60	65	50	55	65	75	60.55
F4	55	65	60	65	60	65	55	60	55	60
F7	35	70	65	60	40	50	40	60	50	52.22
F8	60	60	50	55	75	65	60	40	35	55.56
C4	35	50	55	55	65	50	45	45	60	51.12
C3	60	55	70	55	60	55	80	25	50	56.67
Fz	50	50	50	45	75	60	65	60	50	56.11
P3	50	55	45	65	70	70	55	55	70	59.44
P4	60	45	55	60	40	70	50	50	65	55
Pz	70	65	60	70	70	55	70	60	60	64.45
O1	60	30	60	60	60	55	70	60	60	57.22
O2	60	40	75	55	40	75	55	55	55	56.67
T3	50	65	60	45	45	55	50	60	60	54.45
T4	50	55	65	80	50	60	50	50	75	59.44
T5	65	60	50	55	50	60	65	50	65	57.78
T6	60	60	45	65	50	60	65	55	55	57.22
Oz	70	50	55	65	60	55	60	65	65	60.55

Table 4. LDA classification results of nine volunteers in each channel, and the average classification accuracy (SD, max)

	V1	V2	V3	V4	V5	V6	V7	V8	V9	Average
Fp1	55	45	50	50	60	70	55	45	55	53.89
Fpz	40	50	50	55	60	65	55	60	65	55.56
Fp2	40	50	55	70	65	60	45	65	55	56.11
F3	55	50	50	55	60	60	55	45	70	55.56
F4	45	40	55	65	70	70	60	55	60	57.78
F7	60	45	45	50	60	60	80	65	70	59.45
F8	45	65	50	55	40	55	65	50	60	53.89
C4	55	50	60	55	70	70	60	65	70	61.67
C3	75	55	50	65	55	55	80	75	70	64.45
Fz	55	50	80	50	65	55	65	70	45	59.44
P3	65	75	70	65	70	50	60	60	60	63.89
P4	55	55	60	60	65	50	40	60	60	56.12
Pz	55	80	60	65	55	50	80	90	60	66.11
O1	45	55	55	50	60	50	45	40	75	52.77
O2	50	70	55	50	55	50	45	70	60	56.11
T3	60	60	75	55	85	70	65	60	70	66.67
T4	60	70	45	65	50	60	60	65	75	61.12
T5	50	75	65	65	65	50	70	65	60	62.78
T6	65	60	50	60	55	55	35	55	60	55
Oz	50	60	65	55	65	60	40	60	65	57.78

Table 5. LDA classification results of nine volunteers in each channel, and the average classification accuracy (Hjorth)

	V1	V2	V3	V4	V5	V6	V7	V8	V9	Average
Fp1	55	60	40	50	60	70	55	45	55	54.37
Fpz	70	55	45	70	60	65	55	60	65	60
Fp2	70	55	45	55	65	60	45	65	55	57.5
F3	50	80	60	65	60	60	55	45	70	59.37
F4	45	45	55	50	70	70	60	55	60	56.25
F7	70	70	50	55	60	60	80	65	70	63.75
F8	85	60	60	55	40	55	65	50	60	58.75
C4	65	55	45	65	70	70	60	65	70	61.87
C3	60	60	55	50	55	55	80	75	70	61.25
Fz	75	60	45	65	65	55	65	70	45	62.5
P3	65	65	55	60	70	50	60	60	60	60.62
P4	50	65	65	65	65	50	40	60	60	57.5
Pz	70	65	70	50	55	50	80	90	60	66.25
O1	55	65	60	50	60	50	45	40	75	53.12
O2	55	75	45	55	55	50	45	70	60	56.25
T3	70	45	50	65	85	70	65	60	70	63.75
T4	65	60	65	65	50	60	60	65	75	61.25
T5	50	50	65	60	65	50	70	65	60	59.37
T6	65	65	65	55	55	55	35	55	60	56.25
Oz	75	65	60	55	65	60	40	60	65	60

3.1. Temporal and Frontal Lobe Activity Using Voting

The three channels with the best accuracy rates in all volunteers for all classifiers are listed in Table 6 below. Considering the table of the results presented in the previous section, it can be seen that the channels in the temporal and frontal regions are effective and dominant. Therefore, we can say that these regions are important in 2D and 3D video watching. In order to improve the classification results, three channels from the frontal and temporal lobes with the best accuracy in each volunteer are identified in Table 7. That is, in Table 6, the channels from the parietal, occipital, and central lobes are replaced by the best channels from the frontal and temporal lobes. For a good understanding, for example, in the first volunteer (V1) of Table 2, in the three best channels (C3, F4, T6), the C3 channel was replaced by the channel (Fp1), which gave the best result from the frontal or temporal lobe. The combination of these channels is presented in Table 7.

Table 6 Three channels with the best success in nine volunteers for SD, max, and Hjorth feature extraction methods

	k-NN		LDA	
	SD and max	Hjorth	SD and max	Hjorth
V1	C3, F4, T6	Oz, Pz, F3	C3, T6, P3	F8, Fz, Oz
V2	F3, Oz, T3	Fpz, F7, Fp2	Pz, T5, P3	F3, O2, F7
V3	F8, Fz, P4	O2, C3, T4	Fz, T3, P3	Pz, T5, T4
V4	T4, Pz, Fp1	T4, Pz, Fp1	Fp2, T4, T5	Fp2, F4, T4
V5	Fz, P3, Pz	F8, Fz, P3	T3, F4, P3	T3, F4, C4
V6	O2, Fpz, P4	O2, Fpz, P3	Fp1, F4, T3	Fp1, F4, T3
V7	C3, Fp2, Pz	C3, Fp2, Pz	F7, Pz, C3	F7, C3, Pz
V8	Fpz, Fp2, F3	Fpz, Fp2, F3	Pz, C3, O2	Pz, C3, Fz
V9	T4, Pz, F3	T4, F3, P3	T4, F8, P3	T4, O1, F3

Table 7 The three best channels selected from the frontal and temporal lobes

	k-NN		LDA	
	SD and max	Hjorth	SD and max	Hjorth
V1	F4, T6, Fp1	F3, T5, Fp1	T6, F7, T4	F8, Fz, Fpz
V2	F3, T3, Fp2	Fpz, F7, Fp2	T5, T4, F8	F3, F7, T6
V3	F8, Fz, T6	T4, F7, F4	Fz, T3, T5	T5, T4, T6
V4	T4, Fp1, T6	T4, Fp1, Fp2	Fp2, T4, T5	Fp2, F4, T4
V5	Fz, Fp2, F3	F8, Fz, F3	T3, F4, Fp2	T3, F4, T5
V6	Fpz, F4, F7	Fpz, F8, F4	Fp1, F4, T3	Fp1, F4, T3
V7	Fp2, Fpz, Fz	Fp2, T5, T6	F7, T5, T3	F7, T5, T3
V8	Fpz, Fp2, F3	Fpz, Fp2, F3	Fp2, F7, T4	Fz, T4, T5
V9	T4, F3, Fp1	T4, F3, T5	T4, F3, F7	T4, F3, F7

After selecting the best three channels' combination as a result of the classification process, the result improvement analysis was performed with inter-label voting [37]. According to the decision mechanism, the result is explained for each epoch. This decision mechanism is summarized in Table 8. The label of two (or three channels) must be 1 to be considered the estimated epoch label 1 or "2D_3D". On the other hand, if the estimated epoch label for two channels (or three channels) is 2 and considered "3D_2D", it takes the value of 2.

Thus, in the proposed hypothesis, in both feature extraction methods, only the 2, 3, and 4 Hz from the delta band and the temporal and frontal lobes from the EEG brain lobes were used to classify the 2D_3D and 3D_2D transitions. Meanwhile, the voting flow chart is displayed in Figure 5.

The 2D_3D and 3D_2D transitions' classification results for nine volunteers by using the voting process are presented in Table 9. In this table, the EEG signal analysis of each volunteer has been made for the channels of the volunteer that gives the best classification success. These channels were selected from the frontal and temporal lobes in all the volunteers, and this study is expected to be suitable for use in the biomedical field as a result of reducing the number of channels. Among the nine volunteers, the highest average test performance was calculated in V3 as 90% with SD, max feature extraction method, and LDA classifier. As seen in Table 7, this success was achieved by voting between channels Fz, T3, and T5. In general, based on Table 9, the best average test performance was determined as 76% in both classifications for both feature extraction methods. In addition, the classification performance criteria were largely compatible with each other.

Table 8 Decision mechanism

Channel 1	Channel 2	Channel 3	Estimated label
1	1	1	1
1	1	2	1
1	2	1	1
1	2	2	2
2	1	1	1
2	1	2	2
2	2	1	2
2	2	2	2

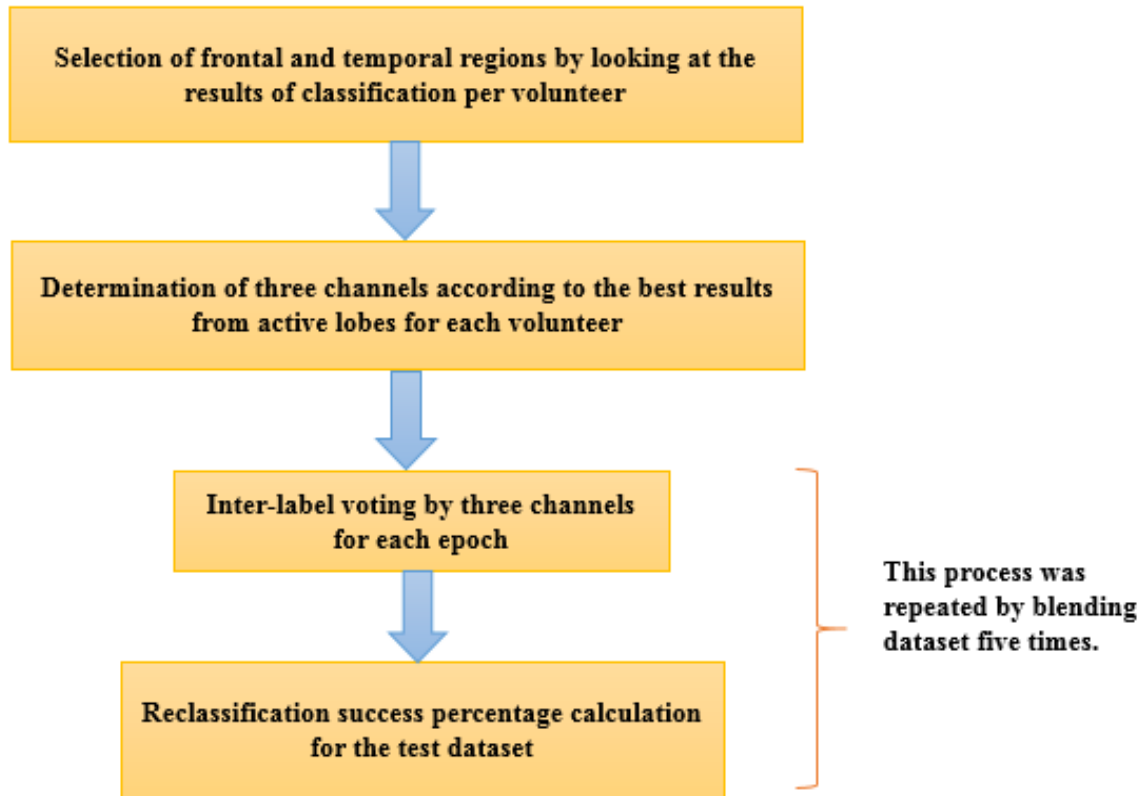


Figure 5 Voting flow chart of temporal and frontal lobe activity

3.2. Classification Techniques and Performance Evaluation of Classification

To achieve the goals of this study, namely 2D_3D and 3D_2D transition classification, k-NN and LDA classification techniques were employed. Since this study has two classes, the chance level is 50%. Additionally, sensitivity and specificity [20] were used to evaluate classifier performance and classification accuracy. The basic definition of these techniques is briefly presented below.

The supervised and simple k-NN technique that is sensitive to the local structure of the data is also used in pattern recognition for different classification and regression processes [38]. This algorithm has a lazy learning mechanism, and it tries to make estimations by memorizing the training dataset instead of learning it. During the estimation process, it looks for the nearest neighbors, taking into account the entire data set [39]. The class of the new sample is determined by looking at the closest k neighbor's class according to the distance in the learning set. In this study, these distances were calculated using the Euclidean distance function. Setting the k or neighborhood parameter directly affects the performance of this algorithm. If k is too small, it affects noise sensitivity, and it is too large, it affects the neighbors of other classes. There is no clear method for choosing k. As used in this study, in general, the optimum k is selected by using the cross-validation method.

The LDA algorithm is used in machine learning and pattern recognition to separate two or more classes of problems. Its basic principle is based on the linear combination of features. This algorithm actually reveals the modeling differences of classes by moving the high-dimensional space features to a smaller size [40].

To evaluate the classification result validation, K-fold cross-validation was applied. K-fold cross-validation was employed to evaluate the classification result validation. In this study, the K value was chosen as 10. After cross-validation, the k value of k-NN was obtained. Here, class 2D_3D and 3D_2D were defined as the positive and negative sample, respectively. To move on to the classification stage, the epochs of each class were randomly divided into two test and training sets. The classification process was repeated 20 times to avoid unlucky splitting.

Table 9 Results of classifying 2D_3D and 3D_2D transitions after the voting process in the temporal and frontal lobe activity

		k-NN (%Success)		LDA (%Success)	
		SD, max	Hjorth	SD, max	Hjorth
V1	Accuracy	63	69	60	85
	Specificity	64.95	61.76	62.5	81.82
	Sensitivity	62	100	58.33	88.89
V2	Accuracy	66	68	80	80
	Specificity	68	100	75	80
	Sensitivity	68.02	61.34	88	80
V3	Accuracy	63	67	<u>90</u>	65
	Specificity	57.6	69.06	83.33	66.67
	Sensitivity	100	66.1	100	63.64
V4	Accuracy	78	79	75	75
	Specificity	76.09	78.55	66.67	66.67
	Sensitivity	80.61	79.56	100	100
V5	Accuracy	74	79	85	85
	Specificity	72.89	74.54	88.89	88.89
	Sensitivity	79.38	88.42	81.82	81.82
V6	Accuracy	80	80	75	75
	Specificity	83.03	82.59	77.78	77.78
	Sensitivity	78.56	78.78	72.73	72.73
V7	Accuracy	76	74	75	75
	Specificity	70.04	67.56	69.23	69.23
	Sensitivity	91.11	92.7	85.71	85.71
V8	Accuracy	76	73	65	65
	Specificity	74.08	73.7	63.64	63.64
	Sensitivity	80.71	74.63	66.67	66.67
V9	Accuracy	82	80	80	80
	Specificity	86.62	80.56	80	80
	Sensitivity	81.78	82.43	80	80
Avearge	Accuracy	73	74	<u>76</u>	<u>76</u>
	Specificity	72.58	76.48	74.11	80.44
	Sensitivity	80.21	74.97	81.42	79.94

It's good to say that in this study, classification performance was evaluated using confusion matrix. The general diagram of this matrix was shown in Table. 10. As an example, the confusion matrices for volunteer 1, 3, 5 and 9 using the Hjorth feature extraction method and LDA algorithm were presented in Figure 6.

Table 10 Confusion matrix

Volunteer		Predicted	
		3D_2D	2D_3D
Observed	3D_2D	TN	FP
	2D_3D	FN	TP

V1		Predicted	
		3D_2D	2D_3D
Observed	3D_2D	9	2
	2D_3D	1	8

V3		Predicted	
		3D_2D	2D_3D
Observed	3D_2D	6	3
	2D_3D	4	7

V5		Predicted	
		3D_2D	2D_3D
Observed	3D_2D	8	1
	2D_3D	2	9

V9		Predicted	
		3D_2D	2D_3D
Observed	3D_2D	8	2
	2D_3D	2	8

Figure 6 The confusion matrices for volunteer 1, 3, 5, 9, using the Hjorth feature extraction method and in LDA algorithm

4. Discussion

In terms of the findings of the scenario, a comprehensive qualitative classification study was obtained after creating EEG records datasets in the analysis of 2D_3D and 3D_2D transitions. In a video consisting of single-stream 2D_3D and 3D_2D random transitions, the behavior of all brain lobes in the EEG frequency bands was studied in detail. In the literature review in the field of 2D and 3D technology, the analysis was started based on some basic information, e.g., by taking into account only the effectiveness of some brain lobes or frequency bands [41], [17], [9]. However, in the proposed study, the dominant EEG band and brain lobes [9] were selected by analyzing all the criteria.

In the literature review phase, no 2D_3D and 3D_2D transition classification hypothesis study was found to capture the moment of transition to sleep. Therefore, this study has great originality and innovation in the field of 2D and 3D technology. Studies in this field have been based on eye fatigue created by 3D televisions, superficial and qualitative brain signal analysis of 2D and 3D game consoles [2], [42], [43], [44]. Although it has been suggested that watching TV in 2D and 3D modes in terms of neurophysiological and cognitive responses does not have major effects on brain dynamics [2], it has been reported in later studies that this result changes in deep analysis [10], [45].

In the inter-volunteer classification results, the channels of the frontal and temporal lobes are the most effective channels in this transition classification analysis of the proposed study. Thus, the delta band and the channels of frontal and temporal lobes appear to be effective in the 2D_3D and 3D_2D transition classification. Based on the overall average results, the parietal lobe is considered to be important in the 2D_3D and 3D_2D transition classification. The presence of channels from the temporal lobe can be explained because of its important role in visual memory. In terms of the frontal lobe channels, this may be because the lobe has an important share of the focus of attention. The function of the parietal lobe is to interpret sensory information from different parts of the brain. Considering the PSD spectrogram graphics based on STFT, frequencies 2, 3, and 4 Hz from the delta band were chosen as the dominant band in these transitions. In this classification study, the importance of transition analysis is based on the eye structure. The human eye sees 3D due to its nature [20]. We proceed in this study, claiming that the ability to see in a sleepy and tired state will lose dimension. Therefore, the analysis and classification of 2D_3D and 3D_2D transition are the main objectives of research.

5. Conclusion

This paper investigated how watching a video consisting of random 2D and 3D parts affects the neural activity of the cortex. One of the innovations of this study was making 2D_3D and 3D_2D transition classification using a single-stream anaglyph video. For these transitions, especially considering the dominant time intervals, the brain activity for different frequency bands and brain lobes was considered. In this study, channels of five brain lobes and all frequency bands were tested in detail and dominant band and effective brain regions were obtained. This stage could be defined as EEG signal preliminary analysis and considered as an important principle of this study compared to other EEG studies. By interpreting the spectrogram charts, the PSD of the brain signals in the transitions was evident at the 2, 3, and 4 Hz of the delta band. It was aimed to reach the target by extracting appropriate features to enhance the performance of classification algorithms from the selected band. The Hjorth parameters and the LDA classification technique yielded promising results in the temporal and frontal lobes in general.

The detailed analysis of the transition moment in a single-stream hybrid video consisting of 2D and 3D parts has the ability to add a new and unique perspective to driver fatigue studies. The results of this study provide a good basis for our future research. It is hoped that this study may shed light on the early diagnosis of the moment of falling asleep when it is conducted comprehensively on tired individuals. Using different feature extraction and classification techniques, the transition moment can be analyzed in professional videos. In addition to these planning, the number of participants can be increased to solve one of the limitations of the study. The classification results of 2D_3D and 3D_2D transition analysis can be improved using deep learning algorithms.

References

- [1] R. N. H. R. Khairuddin, A. S. Malik, and N. Kamel, "EEG Topographical Maps Analysis for 2D and 3D Video Game Play," *2014 5th International Conference on Intelligent and Advanced Systems (ICIAS)*, no. June, pp. 1–4, 2014.
- [2] H.-G. Jeong *et al.*, "The impact of 3D and 2D TV watching on neurophysiological responses and cognitive functioning in adults," *The European Journal of Public Health*, vol. 25, no. 6, pp. 1047–1052, Dec. 2015.
- [3] Q. Wang, O. Sourina, and M. K. Nguyen, "EEG-Based 'Serious' Games Design for Medical Applications," in *2010 International Conference on Cyberworlds*, 2010, pp. 270–276.
- [4] P. Benzie *et al.*, "A Survey of 3DTV Displays: Techniques and Technologies," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 11, pp. 1647–1658, Nov. 2007.
- [5] M. S. Banks, J. C. A. Read, R. S. Allison, and S. J. Watt, "Stereoscopy and the Human Visual System," *SMPTE Motion Imaging Journal*, vol. 121, no. 4, pp. 24–43, May 2012.
- [6] D. Bavelier, C. S. Green, D. H. Han, P. F. Renshaw, M. M. Merzenich, and D. A. Gentile, "Brains on video games," *Nature Reviews Neuroscience*, vol. 12, no. 12, pp. 763–768, Dec. 2011.
- [7] M. Gentzkow, "Television and Voter Turnout*," *Quarterly Journal of Economics*, vol. 121, no. 3, pp. 931–972, Aug. 2006.
- [8] N. Manshouri, M. Maleki, and T. Kayikcioglu, "An EEG-based stereoscopic research of the PSD differences in pre and post 2D&3D movies watching," *Biomedical Signal Processing and*

Control, vol. 55, Jan. 2020.

- [9] N. Zvezdochkina and V. Antipov, "The EEG Activity during Binocular Depth Perception of 2D Images," *Computational intelligence and neuroscience*, vol. 2018, pp. 1–7, 2018.
- [10] N. Manshouri, M. Maleki, and T. Kayıkçıoğlu, "Classification of Human Vision Discrepancy during Watching 2D and 3D Movies Based on EEG Signals," *International Journal of Computer Science and Information Security*, vol. 15, no. 2, pp. 430–436, 2017.
- [11] D. De Waard, "The Measurement of Drivers' Mental Workload," 1996.
- [12] B. J. Fisch and R. Spehlmann, *Fisch and Spehlmann's EEG primer : basic principles of digital and analog EEG*. Elsevier, 1999.
- [13] V. Nityananda and J. C. A. Read, "Stereopsis in animals: evolution, function and mechanisms.," *The Journal of experimental biology*, vol. 220, no. Pt 14, pp. 2502–2512, 2017.
- [14] J. C. A. Read, "What is stereoscopic vision good for?," 2015.
- [15] M.-M. Hamed, E. Marzieh, and Ag. David, "The relationship between binocular vision symptoms and near point of convergence," *Indian Journal of Ophthalmology*, vol. 61, no. 7, p. 325, Jul. 2013.
- [16] C. WHEATSTONE, "On some remarkable and hitherto unobserved phenomena of binocular vision.," *The Optometric weekly*, vol. 53, pp. 2311–5, Nov. 1962.
- [17] M. Z. Ramadan *et al.*, "Effects of Viewing Displays from Different Distances on Human Visual System," *Applied Sciences*, vol. 7, no. 11, p. 1153, Nov. 2017.
- [18] H. R. Khairuddin *et al.*, "Analysis of EEG Signals Regularity in Adults during Video Game Play in 2D and 3D," *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 2064–2067, 2013.
- [19] Z. Minchev, "2D vs 3D Visualization and Social Networks Entertainment Games: A Human Factor Response Case Study," Springer, Berlin, Heidelberg, 2013, pp. 107–113.
- [20] N. Manshouri and T. Kayıkcioglu, "A Comprehensive Analysis of 2D&3D Video Watching of EEG Signals by Increasing PLSR and SVM Classification Results," *The Computer Journal*, May 2019.
- [21] W. Avarvand, Forooz Shahbazi and Bosse, Sebastian and Muller, Klaus-Robert and Schufer, Ralf and Nolte, Guido and Wiegand, Thomas and Curio, Gabriel and Samek, "Objective quality assessment of stereoscopic images with vertical disparity using EEG," *Journal of neural engineering*, vol. 14, no. 4, pp. 1–14, 2017.
- [22] S. Ting, T. Tan, G. West, A. Squelch, and J. Foster, "Quantitative assessment of 2D versus 3D visualisation modalities," in *2011 Visual Communications and Image Processing (VCIP)*, 2011, pp. 1–4.

- [23] A. Carvajal, "Quantitative Comparison between the Use of 3D vs 2D Visualization Tools to Present Building Design Proposals to Non-Spatial Skilled End Users," in *Ninth International Conference on Information Visualisation (IV'05)*, pp. 291–294.
- [24] S. E. Kober, J. Kurzmann, and C. Neuper, "Cortical correlate of spatial presence in 2D and 3D interactive virtual reality: An EEG study," *International Journal of Psychophysiology*, vol. 83, no. 3, pp. 365–374, Mar. 2012.
- [25] Y. Han, H. Y. Lin, and C. Chen, "SP-3 Visual Fatigue for Laser-Projection Light-Field 3D Display in Contrast with 2D Display," *2017 24th International Workshop on Active-Matrix Flatpanel Displays and Devices (AM-FPD)*, pp. 9–12, 2017.
- [26] F. P. S. Fischmeister and H. Bauer, "Neural correlates of monocular and binocular depth cues based on natural images : A LORETA analysis," *Vision research*, vol. 46, no. 20, pp. 3373–3380, 2006.
- [27] (1) *3D Video Chain Saw! - YouTube*. .
- [28] "Xilisoft 3D Video Converter - 3D converter, convert to 3D video." [Online]. Available: <http://www.xilisoft.com/3d-video-converter.html>. [Accessed: 30-Dec-2019].
- [29] "Download Free 3D Video Converter - Convert 2D to 3D | IQmango Free Software." [Online]. Available: http://iqmango.com/3DVideo_Converter.html. [Accessed: 30-Dec-2019].
- [30] "Easiest Video Editing Software Free Download." [Online]. Available: <http://www.idoeditor.com/>. [Accessed: 30-Dec-2019].
- [31] J. Kim *et al.*, "A full-color anaglyph three-dimensional display system using active color filter glasses," *Journal of Information Display*, vol. 12, no. 1, pp. 37–41, Mar. 2011.
- [32] K. Eroğlu, T. Kayıkçıoğlu, and O. Osman, "Effect of brightness of visual stimuli on EEG signals," *Behavioural Brain Research*, vol. 382, p. 112486, Mar. 2020.
- [33] N. Roehri, J. M. Lina, J. C. Mosher, F. Bartolomei, and C. G. Benar, "Time-Frequency Strategies for Increasing High-Frequency Oscillation Detectability in Intracerebral EEG," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 12, pp. 2595–2606, Dec. 2016.
- [34] N. Kehtarnavaz, "Frequency Domain Processing," in *Digital Signal Processing System Design*, Elsevier, 2008, pp. 175–196.
- [35] F. Hlawatsch and G. F. Boudreaux-Bartels, "Linear and quadratic time-frequency signal representations," *IEEE Signal Processing Magazine*, vol. 9, no. 2, pp. 21–67, Apr. 1992.
- [36] S.-H. Oh, Y.-R. Lee, and H.-N. Kim, "A Novel EEG Feature Extraction Method Using Hjorth Parameter."
- [37] T. Kayıkcioglu, M. Maleki, and K. Eroglu, "Fast and accurate PLS-based classification of EEG sleep using single channel data," *Expert Systems with Applications*, vol. 42, no. 21, pp. 7825–

7830, Jun. 2015.

- [38] E. Fix and J. L. Hodges, "Discriminatory Analysis. Nonparametric Discrimination: Consistency Properties," *International Statistical Review / Revue Internationale de Statistique*, vol. 57, no. 3, p. 238, Dec. 1989.
- [39] N. S. Altman, "An introduction to kernel and nearest-neighbor nonparametric regression," *American Statistician*, vol. 46, no. 3, pp. 175–185, 1992.
- [40] A. M. Martinez and A. C. Kak, "PCA versus LDA," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 2, pp. 228–233, Feb. 2001.
- [41] M. Host'ovecký and B. B, "Brain activity: beta wave analysis of 2D and 3D serious games using EEG," *JAMSI*, vol. 13, no. 2, 2017.
- [42] C. Chen, K. Li, Q. Wu, H. Wang, Z. Qian, and G. Sudlow, "EEG-based detection and evaluation of fatigue caused by watching 3DTV," *Displays*, vol. 34, no. 2, pp. 81–88, Apr. 2013.
- [43] C. Chen *et al.*, "Assessment visual fatigue of watching 3DTV using EEG power spectral parameters," *Displays*, vol. 35, no. 5, pp. 266–272, Dec. 2014.
- [44] A. S. Malik *et al.*, "EEG based evaluation of stereoscopic 3D displays for viewer discomfort," *BioMedical Engineering OnLine*, vol. 14, no. 1, p. 21, Dec. 2015.
- [45] S. Kim and D. Kim, "Differences in the Brain Waves of 3D and 2 . 5D Motion Picture Viewers," *arXiv preprint arXiv:1210.2147*.

A Convolutional Neural Network Using Raw EEG Signal Obtained from Single Channel for Automatic Sleep Staging

 Gökse Zekiye ÖZEN¹,  Rayimbek SULTANOV²,  Yunus ÖZEN³,  Zahide YILMAZ GÜNEŞ⁴

¹Corresponding Author; Computer Engineering Department, Kyrgyz–Turkish Manas University, Bishkek, Kyrgyzstan; goksu@goksu.info; <https://orcid.org/0000-0001-7033-0126>; +902623433393

²Computer Engineering Department, Kyrgyz–Turkish Manas University, Bishkek, Kyrgyzstan; rayimbek.sultanov@manas.edu.kg; <https://orcid.org/0000-0001-8802-8021>

³Vocational School Department of Computer Technologies, Yalova University, Yalova; yunus.ozen@yalova.edu.tr; <https://orcid.org/0000-0003-3225-8797>

⁴Derince Education and Research Hospital, Clinic of Neurology, Kocaeli; yilmazzahide@hotmail.com; <https://orcid.org/0000-0001-6948-9324>

Received 4 May 2020; Revised 07 August 2020; Accepted 12 August 2020; Published online 28 August 2020

Abstract

Sleep stages are determined firstly for the evaluation of sleep quality and the diagnosis of sleep diseases. The signals, recorded from sensors connected to various parts of the body, such as electroencephalogram (EEG), electrocardiogram (ECG), electrooculogram (EOG) and electromyogram (EMG) are used for this purpose. After the production of affordable wearable EEG devices for individual use, studies have begun to detect sleep stages from a single channel EEG signal. This paper presents an automated system that can perform sleep staging using a single-channel raw EEG signal. A Convolutional Neural Network (CNN) model was trained with the raw EEG signal for sleep stage detection. The use of CNN does not require any feature extraction. The developed CNN model classifies the sleep data sampled at 250 Hz, divided into 30-second segments according to the 5-class sleep staging system. According to the test results, the performance of the proposed system was found to be 93% macro F1 score and 92% accuracy.

Keywords: Convolutional Neural Network, EEG, Sleep Scoring

Tek Kanallı Ham EEG Sinyali Temelli Otomatik Uyku Evrelemesi Yapan Evrişimsel Sinir Ağı

Öz

Uyku kalitesinin değerlendirilmesi ve uyku hastalıklarının teşhisi için öncelikle uyku evreleri tespit edilmektedir. Bunun için vücudun çeşitli bölgelerine bağlı sensörlerden kaydedilen elektroensefalogram (EEG), elektrokardiyogram (ECG), elektrookülogram (EOG), elektromiyogram (EMG) gibi sinyaller kullanılmaktadır. Bireysel kullanım için uygun fiyatlı giyilebilir EEG cihazlarının üretilmesi ile tek kanallı EEG sinyalinden uyku evreleri tespiti yapılabilmesi için çalışmalar başlamıştır. Bu makalede tek kanallı ham EEG sinyali kullanarak uyku evreleri tespiti yapabilen otomatik bir sistem sunulmaktadır. Bu amaçla ham EEG sinyalleri ile bir Evrişimsel Sinir Ağı (ESA) modeli eğitilmiştir. ESA kullanımı sayesinde herhangi bir özellik çıkarımı yapılmasına ihtiyaç bulunmamaktadır. Geliştirilen ESA modeli 250 Hz’de örneklenmiş, 30 sn’lik segmentlere bölünmüş uyku verisini 5 sınıflı uyku evrelemesi sistemine göre sınıflandırmaktadır. Yapılan testlerin sonuçlarına göre önerilen sistemin başarımları %93 makro F1-skoru ve %92 doğruluk olarak bulunmuştur.

Anahtar Kelimeler: EEG, Evrişimsel Sinir Ağı, Uyku Evrelemesi

1. Giriş

Uyku insan sağlığı için gerekli bir eylemdir. Uyku kalitesinin ölçülmesi ve uyku hastalıklarının tespiti için denekler uyku kliniklerinde belirli bir süre için çeşitli ölçüm aletleri bağlanarak gözetim altında uyutulurlar. Uyku uzmanları vücudun farklı bölgelerine bağlı sensörlerden kaydedilen elektriksel

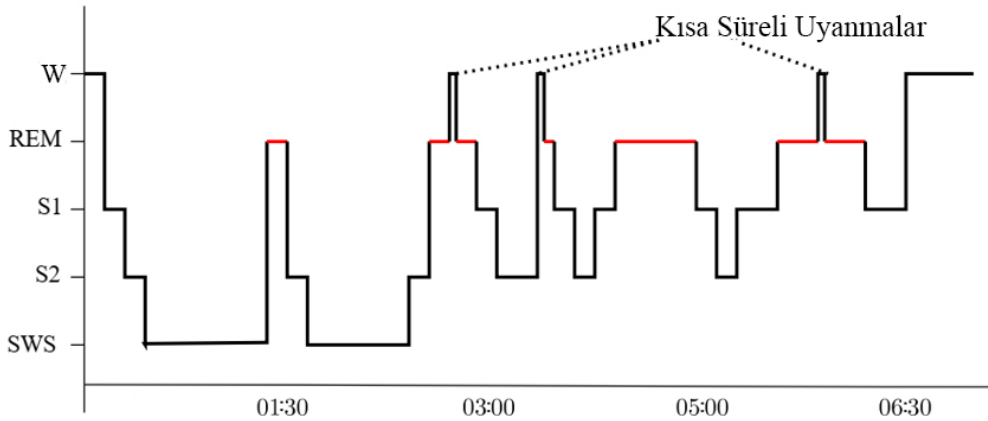
aktiviteyi kullanarak uyku kalitesini değerlendirirler. Bu sensörler yardımıyla ölçülen çoklu fizyolojik sinyallere polisomnogram (PSG) denilmektedir. Bu sinyaller genellikle elektroensefalogram (EEG), elektrokardiyogram (ECG), elektrookülogram (EOG), elektromiyogram (EMG), hava akımı ve oksijen saturasyonundan (SaO₂) oluşmaktadır. PSG, uyku ile ilgili analizlerde en sık kullanılan tanı aracıdır [1]. Bununla birlikte PSG cihazları hastane ve uyku laboratuvarlarında kullanılan pahalı ve özel cihazlardır. Bu cihazlar evde kişisel kullanım için uygun değildir.

PSG ile kaydedilen sinyaller, 30 sn'lik epok adı verilen dilimlere ayrılır ve daha sonra Rechtschaffen ve Kales (R&K) [2] veya Amerikan Uyku Tıbbi Akademisi (AUTA) gibi uyku kılavuzlarına göre uzmanlar tarafından farklı uyku aşamalarına göre sınıflandırılır. AUTA tarafından yayınlanan güncel standart, dört farklı uyku evresi tanımlamaktadır. Bu evreler S1 (uyku hali), S2 (hafif uyku), S3 / SWS (yavaş dalga uyku), REM (hızlı göz hareketi uykusu) ve buna ilave olarak Uyanıklık (W) şeklinde adlandırılmaktadır. AUTA standardı, eski R&K standardının güncellenmiş sürümüdür [3]. AUTA ve R&K arasındaki dönüşüm işlemi Tablo 1'de sunulmaktadır. R&K'da yer alan S3 ve S4 evreleri AUTA'da SWS olarak birleştirilmiştir.

Tablo 1 R&K Standardının AUTA Standardına Dönüşüm Tablosu

Standart	Uyku Evreleri				
R&K	S1	S2	S3 ve S4	REM	W
AUTA	S1	S2	SWS	REM	W

Gece uykusu genellikle birbirini takip eden döngüler halinde gerçekleşmektedir. Bir uyku döngüsü yaklaşık 90 dakika sürmekte ve bu döngü S2, SWS, REM uykusunu ve ardından uyku esnasında ve gece boyunca uyarılmalardan sonra gelen S1'i içermektedir. Bir gece uykusunda bu döngü 3-5 defa tekrarlanmaktadır. Uyku evrelerinin seyri hipnogram adı verilen grafiklerle gösterilir [4]. Şekil 1'de örnek bir hipnogram görülmektedir.



Şekil 1 Örnek Bir Hipnogram Görüntüsü

PSG ölçümü esnasında deneklerin vücudunda farklı bölgelere kablolar yardımıyla çeşitli ölçüm uçları bağlanmaktadır. PSG düzeneği ölçümlerde gürültülere neden olabilmekte ve deneklerin uyku kalitesini etkilemektedir [5]. Bu nedenle, daha az fizyolojik sinyal kullanarak yapılacak bir uyku evreleme sistemine ihtiyaç duyulmaktadır. Böylece ev tabanlı uyku izleme sistemlerinin geliştirilmesi de mümkün olacaktır.

PSG kayıtlarının uzmanlar tarafından elle uyku evrelerine ayrılması işlemi hem zaman alıcı hem de yorucu bir işlemdir [6]. Bu işlemin otomatik olarak bir yazılım tarafından yapılabilmesi, daha az insan gücü kullanmayı ve daha doğru sonuçlar üretmeyi de olanaklı kılacaktır.

Uyku evreleme işleminde evrelerin tespiti için çeşitli özellikler belirlenmektedir. Doğrusal olmayan süreçlerden türetilen bu elle çıkarılan özellikler, zaman domeni [7] özelliklerinden frekans domeni [8] özelliklerine kadar çeşitlilik göstermektedir. Bu özellikleri sınıflandırmak için genellikle geleneksel makine öğrenmesi yöntemleri kullanılmaktadır [9].

Literatürde son yıllarda derin öğrenme yöntemlerinin çeşitli sınıflandırma problemlerine uygulandığı çalışmalar yer almaktadır. Uyku evrelemesi işleminde de derin öğrenme modellerinin kullanılması yaygınlaşmıştır. Derin öğrenmenin gücü verilerden otomatik özellik öğrenme yeteneğine dayanmaktadır. Bir derin öğrenme çeşidi olan Evrişimsel Sinir Ağları (ESA), sınıflandırma problemlerinde başarılı sonuçlar üretmektedir [10].

ESA modelleri bir giriş sinyaline uygulanan aynı ya da farklı boyutta, birden fazla filtre içerir. Evrişim katmanı, havuzlama katmanı, aktivasyon katmanı ve tam bağlı katman olmak üzere dört temel katmandan oluşur. Genellikle katmanlar ardışık olarak bağlanır ve her katman bir önceki katmanın çıkışını alarak bunlardan daha soyut özellik çıkarımları yapabilir. Her bir evrişim katmanındaki yüksek boyutlu özellikler havuzlama katmanı yardımıyla düşük boyutlu özelliklere dönüştürülürler. ESA'da evrişim katmanlarının sonunda sınıflandırma yapmak için tam bağlı katman olarak adlandırılan çok katmanlı bir yapay sinir ağı yer alır. Tam bağlı katmanda düşük boyutlu özellikler ile sınıflar eşleştirilerek sınıflandırma yapılır. Bu adımların gerçekleştirilmesi ile ham veriden soyut özelliklerin otomatik olarak çıkarılması mümkün olmaktadır [11].

Literatürde uyku evrelerini EEG, EOG ve EMG gibi çoklu sinyaller kullanarak otomatik olarak sınıflandıran çeşitli çalışmalar yer almaktadır. Sokolovsky vd. EEG ve EOG sinyallerini kullanarak bir ESA modeli ile otomatik uyku evrelemesi yapan bir sistem geliştirmişlerdir. Modelin makro F1-skoru %77 ve doğruluğu %81 olarak bulunmuştur [12]. Bununla birlikte tek kanallı EEG sinyalleri kullanarak da benzer bir doğruluk başarımına sahip çalışmalar bulunmaktadır. Bu çalışmaların bazılarında geleneksel makine öğrenmesi yöntemleri tercih edilmiştir ve başarımı çok kanallı çalışmalardan düşük değildir. Ghimatgar vd. tarafından yapılan çalışmada tek kanallı EEG sinyali kullanarak bir özellik seçici yöntemiyle uyku epoklarından elle özellik çıkarımı yapılmış, bir makine öğrenmesi yöntemi olan rassal orman sınıflandırıcı ile veri sınıflandırılmıştır. 6 aşamalı R&K ve 5 aşamalı AUTA'ya göre modelin uyku evreleme başarımı sırasıyla en yüksek %87,4 ve %80,4 doğruluk olarak bulunmuştur [13].

Elle özellik çıkarımı yapmak yerine otomatik özellik çıkaran derin öğrenme yöntemleri kullanıldığında da önerilen modellerin başarımı düşmemiştir. Tsinalis vd. tarafından önerilen otomatik uyku evrelemesi yapan tek kanallı ham EEG sinyallerini işleyebilen ESA modeli, genel bir veriseti üzerinde %74 F1-skoru ve %81 doğrulukla evreleme yapmıştır [14]. Zhang vd. iki genel verisetinden alınan tek kanallı EEG sinyallerine Hilbert–Huang dönüşümü uygulamışlar ve 2 boyutlu zaman-frekanslı görüntüler elde etmişlerdir. Çalışmada önerilen Ortogonal ESA ile eğitilen model uyku evrelemesinde sırasıyla %88.4 ve %87.6 toplam sınıflandırma doğruluğu elde etmiştir [15]. Sors vd. tarafından yapılan çalışmada 12 evrişim katmanı ve 2 tam bağlı katmana sahip bir ESA kullanılmıştır. Etiketli ve tek kanallı ham EEG sinyalleri herhangi bir önışlemeye ya da özellik seçimine gerek kalmadan sisteme verilmiştir. 5 sınıflı uyku evrelemesi yapan sistemin doğruluğu %0.81 kapa katsayısıyla, %87 olarak bulunmuştur [16]. Supratak vd. MASS ve Sleep-EDF adlı iki ayrı veriseti üzerinde sınıflandırma yapan iki aşamalı bir model geliştirmiştir. Tek kanallı EEG kullanarak EEG epoklarından özellik çıkarımı için ESA ve uyku evreleri arasındaki geçiş kurallarını öğrenmek için çift yönlü uzun-kısa dönem bellek kullanmışlardır. Modelin doğruluk ve makro F1-skoru sırasıyla MASS için %86.2 ve %81.7 olarak, Sleep-EDF için ise %82 ve %76.9 olarak bulunmuştur [6]. Karmaşık bir yöntem tercih edilmesine rağmen sunulan çalışmanın başarımı literatürdeki tek kanallı EEG kullanan çalışmalardan daha yüksek değildir.

Bu makalede sunulan çalışmada ise yine tek kanallı EEG sinyalleri kullanılarak benzer çalışmalardan daha yüksek başarımlar sağlayacak bir ESA modeli geliştirilmiştir. Aktivasyon fonksiyonu olarak ELU (Üstel Doğrultulmuş Birim) tercih edilmiştir. Evrişimsel katmanlara sistemin başarımını artırmak için normalizasyon katmanı ilave edilmiştir. Sistemin ezberlemesinin engellenmesi için de uygun bir dropout değeri belirlenmiştir.

Makalenin 2. bölümünde, kullanılan veriseti, geliştirilen ESA mimarisi ve değerlendirme ölçekleri sunulmaktadır. 3. bölümde ESA'nın eğitilmesi sonucunda ortaya çıkan sonuçlar literatürde yer alan

diğer çalışmalarla karşılaştırmalı olarak tartışılmaktadır. Son bölümde ise değerlendirmeler ve öneriler yer almaktadır.

2. Materyal ve Metot

Bu bölümde çalışmada kullanılan veriseti ve özellikleri yer almaktadır. Verisetinde yer alan dosyalar ve özellikleri ile ESA modeline verilmeden önce sinyallere uygulanan ön işleme adımları sunulmaktadır. Geliştirilen ESA modeli ve bu modelin başarımını ölçmek için kullanılan değerlendirme ölçekleri de detaylı bir şekilde açıklanmaktadır.

2.1 Veriseti

Bu makalede kullanılan EEG sinyalleri PhysioNet'te yer alan polisomnografik veritabanı MIT-BIH'den alınmıştır. Denekler, Boston'daki Beth Israel Hospital Uyku Laboratuvarı'nda kronik obstrüktif uyku apne sendromunun değerlendirilmesi için izlenmiştir. Veritabanında yer alan çok kanallı PSG kayıtlarındaki her bir kayıt uzmanlar tarafından evrelerine ayrılmış ve çeşitli uyku rahatsızlıkları ile birlikte etiketlenmiştir. Verisetinde toplam 80 saatlik kayıt bulunmaktadır. Her biri farklı derecelerde apneye sahip 32-56 yaşları arasındaki 16 erkek denekten alınan 18 tüm gece PSG kaydı içermektedir. Her kayıta 250 Hz frekansında örneklenmiş EKG sinyali, invazif kan basıncı sinyali, solunum sinyali ve EEG sinyali bulunmaktadır. Her bir kayıt için veritabanında 4 farklı türde dosya bulunmaktadır [17]. Tablo 2'de MIT-BIH uyku verisetinde yer alan dosyaların uzantıları ve açıklamaları sunulmaktadır.

Tablo 2 MIT-BIH Uyku Veriseti Dosya Uzantıları ve Açıklamaları

Dosya Uzantısı	Açıklama
st	Uyku evreleri ile beraber uyku apnesi açıklamaları
ecg	Kalp ritmi açıklamaları
dat	Sinyaller
hea	Başlık dosyaları

MIT-BIH verisetinde yer alan her bir st uzantılı dosyanın içindeki uyku evresi ve uyku apnesi için olan açıklamaların tamamı dosyadaki aux alanında yer almaktadır. Bu dosyada yer alan her bir açıklama notu, açıklama notuna karşılık gelen bir epok veriye aittir. Aux alanındaki uyku evrelerine dair kodlar ve açıklama bilgileri Tablo 3'te yer almaktadır.

Tablo 3 Aux Alanında Yer Alan Uyku Evrelerine Dair Açıklamalara Karşılık Gelen Kodlar

Aux alanında yer alan kodlama kısaltmaları	Açıklaması
W	Uyanıklık
1	Uyku evresi 1 (S1)
2	Uyku evresi 2 (S2)
3	Uyku evresi 3 (S3)
4	Uyku evresi 4 (S4)
R	REM evresi

16 denegin EEG sinyallerinin tamamı kafa derisi üzerinde aynı bölgelerden alınmamıştır. Verisetinde yer alan slp02 kaydı O2-A1 elektrot çiftinden, slp01, slp32, slp37 ve slp41 kayıtları C4-A1 elektrot çiftinden ve diğer kayıtlar C3-O1 elektrot çiftinden kaydedilmiştir [18]. Tüm EEG kayıtları, birbiri ile çakışmayan 30 saniyelik epoklara ayrılmış ve daha sonra bu epokların her biri taşıdıkları özelliklere göre uzmanlar tarafından elle etiketlenmiştir.

Bu makalede sunulan çalışmada MIT-BIH uyku verisetinden alınan toplam 4640 epok'luk EEG sinyali kullanılmıştır. Bu epokların REM, S1, S2, SWS ve W evrelerine göre dağılımı sırasıyla 1198, 1177, 1072, 1193 şeklindedir. Sinyaller 250 Hz ile örneklediği için 30 sn'lik epokların her biri 7500 adet örnek içermektedir.

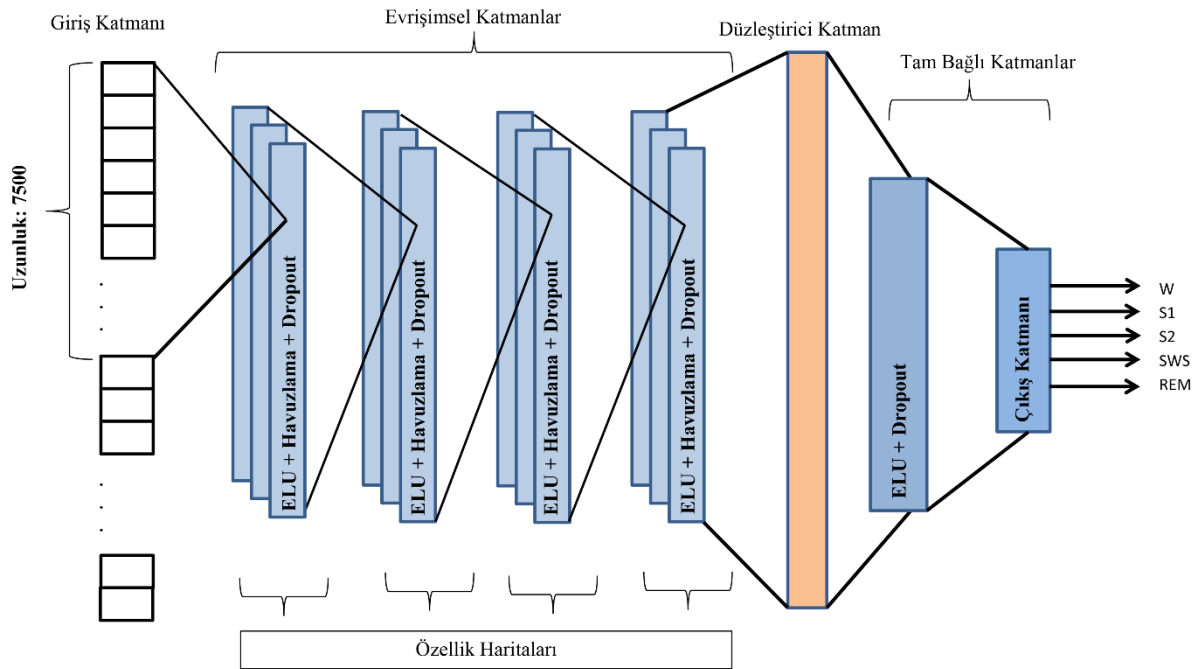
2.2 Önişleme

Deneklerin PSG kayıtlarında etiketlenmemiş olarak yer alan epoklar ESA'ya gönderilmeden önce sinyalden çıkarılmıştır. Verisetini dengeli hale getirebilmek ve gereksiz kaynak kullanımını engellemek için uykuya geçilmeden önceki ve uykudan uyandıktan sonraki uyanıklık evreleri sinyalden çıkarılmıştır. AUTA tarafından tavsiye edilen şekilde uyku evreleri W, S1, S2, SWS (S3 ve S4), REM şeklinde etiketlenmiştir, yani S3 ve S4 evreleri SWS olarak birleştirilmiştir.

2.3 Önerilen ESA Mimarisi

Bu çalışmada giriş olarak tek kanallı ham EEG sinyali alan ve özellikleri kendisi öğrenen bir ESA modeli kullanılmıştır. Önerilen ESA modeli blok diyagramı Şekil 2'de gösterilmektedir. Aktivasyon fonksiyonu olarak tüm katmanlarda ELU (Üstel Doğrultulmuş Birim) [19] kullanılmıştır.

Oluşturulan ESA modelinde 1 epokluk giriş dizisi, farklı sayılarda filtre ve çekirdeklere sahip 4 evrişimsel katman ve her bir evrişim katmanının sonunda aynı boyutta oluşturulan maksimum havuzlama katmanı bulunmaktadır. Evrişim katmanlarının sonunda bir düzleştirici katman ve 25 nörona sahip bir tam bağlı katman ile de sınıflandırma işlemi yapılmaktadır. Tüm evrişimsel katmanlara normalizasyon eklenmiştir. Sistemin ezberlemesinin engellenmesi için tüm katmanlara deneyler sonucu bulunan en uygun dropout değeri eklenmiştir.



Şekil 2 Önerilen ESA Modelinin Blok Diyagramı

Tablo 4'te model mimarisi, çıktı boyutları ve parametre sayıları detaylı olarak sunulmaktadır. Sıralı yapıda bir model kullanıldığı için bir önceki katmanın çıkışı bir sonraki katmanın girişi olacak şekilde çalışmaktadır.

Tablo 4 Kullanılan Modelin Özellikleri

Katman Tipi	Çıktı Biçimi	Parametre Sayısı
Conv1D_1 (Conv1D)	(7500, 40)	3240
Batch_Normalization_1	(7500, 40)	160
Activation_1	(7500, 40)	0
Max_Pooling1d_1	(7500, 40)	0
Dropout_1	(750,40)	0
Conv1D_2(Conv1D)	(750, 60)	48060

Tablo 4 Kullanılan Modelin Özellikleri (Devamı)

Katman Tipi	Çıktı Biçimi	Parametre Sayısı
Batch_Normalization_2	(750, 60)	240
Activation_2	(750, 60)	0
Max_Pooling1d_2	(150,60)	0
Dropout_2	(150,60)	0
Conv1D_3(Conv1D)	(150,80)	48080
Batch_Normalization_3	(150,80)	320
Activation_3	(150,80)	0
Max_Pooling1d_3	(30,80)	0
Dropout_3	(30, 80)	0
Conv1D_4(Conv1D)	(30,100)	32100
Batch_Normalization_4	(30, 100)	400
Activation_4	(30,100)	0
Max_Pooling1d_4	(6,100)	0
Dropout_4	(6, 100)	0
Flatten_1	(600)	0
Dense_1 (Dense)	(50)	30050
Dropout_5	(50)	0
Dense_2(Dense)	(5)	255

ESA modeli 5 katlamalı çapraz doğrulama kullanılarak verinin %80'i eğitim, %20'si test olacak şekilde ayrılmıştır. Çapraz doğrulama kullanıldığı için test setinde yer alan epoklar her bir katlamada rastgele seçilmiş ve böylece her katlamada farklı verilerle doğrulama işlemi yapılmıştır.

Sistemin ezberlemesini engellemek için kullanılan dropout değerleri tüm katmanlarda 0.3 olarak kullanılmıştır. 5 epok boyunca doğrulama setinde hiç iyileşme olmazsa eğitim durdurulmuştur. Eğitim tamamlandıktan sonra çıkışlar 5'li olarak sınıflandırılmıştır. Bu sınıflar REM, S1, S2, SWS ve W şeklindedir.

Deneyler Python dili 3.7 sürümü ile Keras Kütüphanesi ve Tensorflow Kütüphanesi kullanılarak gerçekleştirilmiştir. Eğitimler için Google bulut sunucusu ortamında 15GB disk, 32GB bellek ve 16 çekirdek işlemcili sanal makineler kullanılmıştır. Geliştirme ortamı Ubuntu 18.0.4 LTS işletim sistemi üzerinde kurulmuştur.

2.4 Değerlendirme Ölçekleri

Önerilen yöntemin başarımı Denklem 1-4'te verilen doğruluk, hassaslık, kesinlik, karışıklık matrisi ve F1-skoru değerlendirme ölçekleri kullanılarak ölçülmüştür.

$$\text{Doğruluk} = \% \frac{DT+DN}{DP+YN+YP+DN} \quad (1)$$

$$\text{Hassasiyet} = \frac{DP}{DP+YN} \quad (2)$$

$$\text{Kesinlik} = \% \frac{DP}{DP+YP} \quad (3)$$

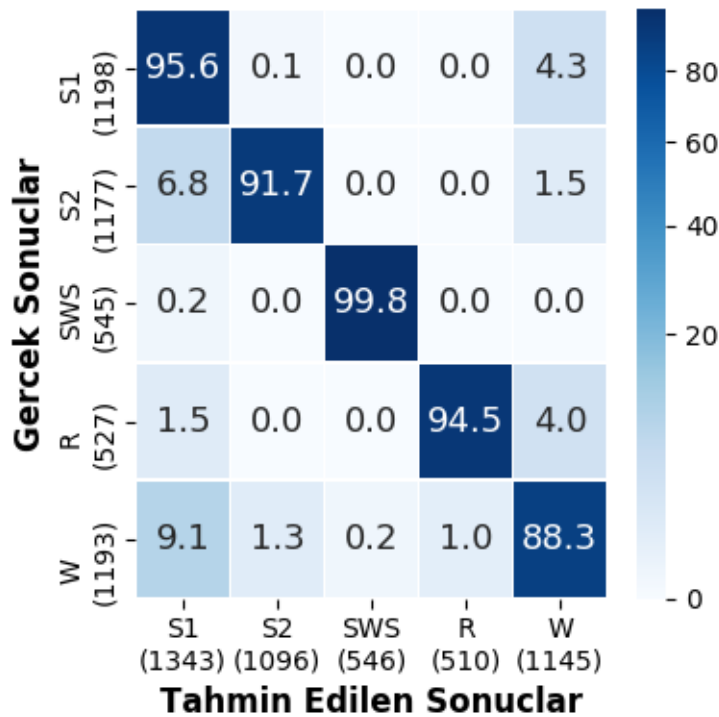
$$\text{F1} = 2 \times \frac{(\text{Hassasiyet} \times \text{Kesinlik})}{\text{Hassasiyet} + \text{Kesinlik}} \quad (4)$$

Denklemlerde yer alan Doğru Pozitif (DP), belirtilen uyku evresinin doğru olarak tahmin edildiği epokların sayısını belirtmektedir. Doğru Negatif (DN), belirtilen uyku evresine ait olmayan ve bu şekilde işaretlenmiş epokların sayıdır. Yanlış Pozitif (YP), belirtilen uyku evresine ait olmasına rağmen

o şekilde işaretlenmeyen epokların sayısıdır. Yanlış Negatif (YN) ise belirtilen uyku evresine ait olmadığı halde, o evreye aitmiş gibi işaretlenen epokların sayısıdır [15].

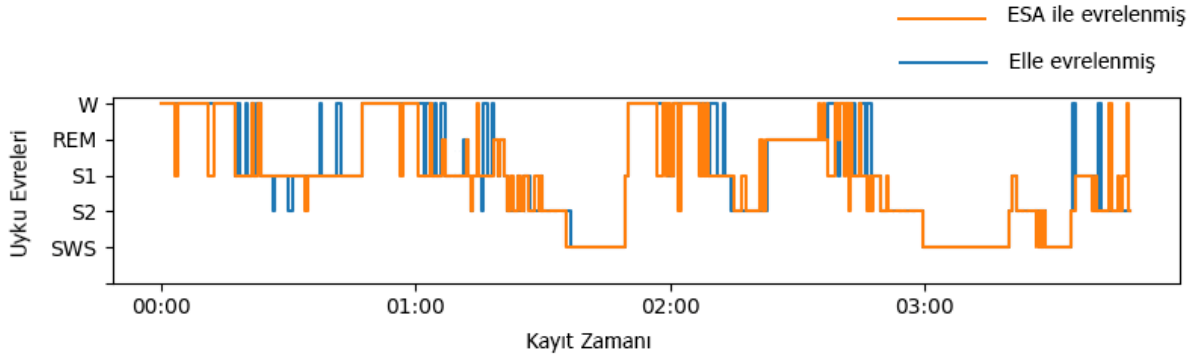
3. Bulgular ve Tartışma

Çalışmada sunulan sisteme ait test sonuçları Şekil 3’de karışıklık matrisi olarak gösterilmektedir. En yüksek doğruluk değeri %99.8 ile SWS evresinde, en düşük doğruluk değeri ise %88.3 ile W evresinde hesaplanmıştır. S1 evresi uyanıklıktan uyku durumuna geçiş evresi olduğu için W ve S2 evresi ile karışabilmektedir. Bunun nedeni S1 evresindeki dalga yapısının kendisinden önce gelen W ve kendisinden sonra gelen S2 evresindeki dalga yapısına olan benzerlikleridir [15]. Önerilen ESA modeli bu karışıklığı yüksek oranda önlemekle birlikte W evresini diğer evrelere göre daha düşük doğrulukla sınıflandırmış ve bu evre %9.1 oranında S1 evresi ile karışmıştır. S2 evresi de aynı şekilde %6.8 oranında S1 evresi ile karışmıştır.



Şekil 3 Uyku Evreleme İşlemine Ait Karışıklık Matrisi

Verisetinden örnek olarak seçilen slp59 kaydı için makalede sunulan ESA modeli tarafından uyku evreleme yapılarak elde edilen evreler ile elle skorlanmış evrelerin birlikte gösterildiği hipnogram görüntüsü Şekil 4’te gösterilmektedir. Turuncu renk ile gösterilen uyku evreleri ESA modeli tarafından, mavi renk ile gösterilen uyku evreleri ise uzmanlar tarafından elle etiketlenmiştir. Evreler yüksek bir oranda birbiri ile örtüşmüştür. Önerilen sistem deneğin tüm gece uyku verisinden otomatik olarak uyku evrelerini işaretleyebilmektedir.



Şekil 4 Slp59 için Elle Etiketlenmiş ve ESA ile Otomatik Etiketlenmiş Evrelerin Hipnogramı

Modeli test etmede kullanılan değerlendirme ölçeklerine göre başarımlarını Tablo 5'te sunulmaktadır.

Tablo 5 Başarımların Sonuçları

Uyku Evreleri	Keskinlik	Hassasiyet	F1-Skorlama
REM	0.85	0.95	0.90
S1	0.96	0.91	0.93
S2	0.99	1.00	0.99
SWS	0.93	0.97	0.95
W	0.93	0.85	0.89
Makro F1 Skorlama	0.93	0.94	0.93
Doğruluk			0.92

Sunulan modelin literatürde yapılan diğer tek kanallı EEG sinyali kullanılarak yapılan çalışmalarla karşılaştırması Tablo 6'da gösterilmiştir. Her bir çalışma için kullanılan veriseti, bu verisetinde kullanılan kanal sayısı ve denek sayıları, evreleme çeşidi, çalışmalar için kullanılan yöntemler ve çalışmaların başarımlarını tablodaki yer almaktadır.

Tablo 6 Yapılan Çalışmanın Sonuçlarının Literatürde Yapılan Diğer Çalışmaların Sonuçlarıyla Karşılaştırılması

Çalışma	Veriseti	Kanal Seçimi	Denek Sayısı	Evreleme Tipi	Kullanılan Yöntem	Doğruluk	F1-skoru
Sors	SHHS-1	Tek Kanal EEG	5728	R&K	ESA	%87	%87
Tsinalis	Sleep-EDF	Tek Kanal EEG	20	R&K	ESA	%74	%81
Zhang	MIT-BIH	Tek Kanal EEG	16	R&K	Hilbert Huang Dönüşümü, OESA	%88.4	%85.8
Supratak	MASS	Tek Kanal EEG	32	AUTA	ESA - UKDB	%86	%86
Bu çalışma	MIT-BIH	Tek Kanal EEG	16	AUTA	ESA	%92	%93

Makalede sunulan tek kanallı EEG kullanan ESA modelinin hem doğruluk hem de F1-skoru değeri literatürdeki rakiplerine göre yüksek çıkmıştır. Zhang vd. tarafından sunulan çalışmada da MIT-BIH veriseti kullanılmış, sinyaller Hilbert Huang Dönüşümü ile resim işleme aşamasından geçirilmiş olmasına rağmen modelin doğruluk başarımlarını %88.4'te kalmıştır. Ham EEG sinyalinde dönüştürme işlemleri yapmadan da uygun bir ESA modeli ile daha yüksek başarımlarını alınabilmektedir.

4. Sonuçlar

Bu makalede sunulan çalışmada tek kanallı ham EEG sinyal örnekleri kullanılarak otomatik uyku evrelemesi yapan bir sistem sunulmaktadır. Bu sistem için tek kanallı ham EEG sinyalleri alan bir Evrişimsel Sinir Ağı modeli geliştirilmiştir. Geliştirilen model PhysioNet'te yer alan MIT-BIH

polisomnografik veritabanında yer alan 16 erkek denekten alınan 18 tüm gece PSG kaydı ile eğitilmiştir. Gerçekleştirilen deneylerin sonuçlarına göre elle özellik çıkarımı yapmadan bir derin öğrenme ağının yüksek başarımla eğitilebildiği görülmektedir. Önerilen sistemin test başarımları %93 makro F1-skoru ve %92 doğruluk olarak bulunmuştur. Yapılan çalışma diğer çalışmalarla karşılaştırıldığında, önerilen modelin başarımları daha yüksektir. Tek kanallı EEG sinyali ile uyku evrelemesi yapılabilmesi sonucunda kişisel kullanıma uygun ve giyilebilir uyku izleme sistemlerinin geliştirilmesinin mümkün olacağı değerlendirilmektedir.

Referanslar

- [1] R. Sharma, R.B. Pachori, and A. Upadhyay, “Automatic Sleep Stages Classification Based on Iterative Filtering of Electroencephalogram Signals,” *Neural Computing and Applications*, vol. 28, no. 10, pp. 2959-2978, 2017.
- [2] J. A. Hobson, “Sleep: Biochemical Aspects,” *New England Journal of Medicine*, vol. 281, no. 26, pp. 1468-1470, 1969.
- [3] C. Iber, “Development of A New Manual for Characterizing Sleep,” *Journal of Sleep and Sleep Disorder Research*, vol. 27, no. 2, pp. 190-192, 2004.
- [4] B. J. Swihart, B. Caffo and K. Bandeen, “Characterizing Sleep Structure Using the Hypnogram,” *Journal of Clinical Sleep Medicine*, vol. 4, no.04, pp. 349-355, 2008.
- [5] H. Phan, F. Andreotti, N. Cooray and O. Y. Chen, “Joint Classification and Prediction CNN Framework for Automatic Sleep Stage Classification,” *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 5, pp.1285 – 1296, 2019.
- [6] A. Supratak, H. Dong, C. Wu and Y. Guo, “DeepSleepNet: A Model for Automatic Sleep Stage Scoring Based on Raw Single-Channel EEG,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 11, pp. 1998-2008, 2017.
- [7] S. J. Redmond and C. Heneghan, “Cardiorespiratory-based Sleep Staging in Subjects with Obstructive Sleep Apnea,” *IEEE Transactions on Bio-Medical Engineering*, vol. 53, no. 3, pp. 485-496, 2006.
- [8] B. Koley and D. Dey, “An Ensemble System for Automatic Sleep Stage Classification Using Single Channel EEG Signal,” *Computers in Biology and Medicine*, vol. 42, no.12, pp.1186-1195, 2012.
- [9] E. Alickovic, J. Kevric and A. Subasi, “Performance Evaluation of Empirical Mode Decomposition, Discrete Wavelet Transform, and Wavelet Packed Decomposition for Automated Epileptic Seizure Detection and Prediction,” *Biomedical Signal Processing and Control*, vol. 39, pp. 94-102, 2018.
- [10] N. Stephansen, E. Shane, C. Jessica, R. Jason. and R. David, “Survey of Machine Learning Techniques in Drug Discovery,” *Current Drug Metabolism*, vol. 20, no. 3, pp.185-193, 2019.
- [11] K. O. Shea and R. Nash, 2015. “An Introduction to Convolutional Neural Networks,” *Arxiv:1511.08458v2*, 2015.

- [12] M. Sokolovsky, F. Guerrero, S. Paisarnsrisomsuk and S. A. Alvarez, "Deep Learning for Automated Feature Discovery and Classification of Sleep Stages," *Transactions of the Institute of Measurement and Control*, vol. 38, pp. 435-451, 2018.
- [13] H. Ghimatgar, K. Kazemi, M. S. Helfroush. and A. Aarabi, "An Automatic Single-Channel EEG-Based Sleep Stage Scoring Method Based on Hidden Markov Model," *Journal of Neuroscience Methods*," vol. 324, pp. 270-285, 2019.
- [14] O. Tsinalis, P. M. Matthews and Y. Guo, "Automatic Sleep Stage Scoring Using Time-Frequency Analysis and Stacked Sparse Autoencoders," *Annals of Biomedical Engineering*, vol. 44, pp. 1587-1597, 2016.
- [15] J. Zhang, R. Yao, W. Ge and J. Gao, "Orthogonal Convolutional Neural Networks for Automatic Sleep Stage Classification Based on Single-Channel EEG," *Transactions of the Institute of Measurement and Control*, vol. 38, pp. 435-451, 2019.
- [16] A. Sors, S. Bonnet, S. Mirek, L. Vercueil and J. Payen, "A Convolutional Neural Network for Sleep Stage Scoring from Raw Single-Channel EEG," *Biomed. Signal Proc. Control*, vol. 42, pp. 107-114, 2018.
- [17] physionet, "MIT-BIH Polysomnographic Database Files," 1999. [Online]. Available: <https://physionet.org/content/slpdb/1.0.0/>. [Accessed: 19-Oct-2019].
- [18] physionet, "MIT-BIH Polysomnographic Database Information," 1999. [Online]. Available: <https://physionet.org/content/slpdb/1.0.0/slpdb.shtml>. [Accessed: 21-Oct-2019].
- [19] D. Pedamonti, "Comparison of Non-Linear Activation Functions for Deep Neural Networks on Mnist Classification Task," *Arxiv:1804.02763v1*, 2018.