Nonlinear Feature Analysis for EEG-Based Biometric Authentication via Machine-learning Makine Öğrenmesi Yoluyla EEG Tabanlı Biyometrik Kimlik Doğrulama için Doğrusal Olmayan Özellik Analizi

Mustafa Yasir ÖZDEMİR[1*], Yalçın İŞLER[2]
[1]Department of Biomedical Technologies, Izmir Katip Celebi University, Izmir, Turkey
[2]Department of Biomedical Engineering, Izmir Katip Celebi University, Izmir, Turkey
ORCIDs: 0009-0009-6984-0386, 0000-0002-2150-4756
E-mails: musyasoz@gmail.com, islerya@yahoo.com
*Corresponding author.

***Abstract*—Among biometric recognition systems, a system using brain waves via EEG will hold a special place. The EEG signal, with its nonlinear structure, is unique to the individual and nearly impossible to replicate. In designing such a system, various signal processing and classification methods are considered. In this study, nonlinear features such as Fractal Dimension, Second Order Sample Entropy, Quantities Graph, and Visibility Graph were used, allowing the examination of the EEG signal independently of the amplitude scale. To reduce computational load, the resting state alpha waves, which are prominent features of the EEG, were focused on, and a low number (8) of electrodes were used. The obtained features were analyzed separately for each electrode, aiming to identify the most distinctive feature and electrode. The classification was performed using five different machine-learning methods. The highest accuracy was achieved by the Random Forest algorithm. The most distinctive electrode and features were identified as the Fractal Dimension of the F5 electrode and the Fractal Dimension of the Oz electrode.**

***Keywords*—EEG; biometric; Fractal dimension; Quadratic Sample Entropy; Quantiles Graph; Visibility Graph**

***Özetçe*—Biyometrik tanıma sistemleri arasında EEG ile beyin dalgalarının kullanıldığı bir sistem özel bir yer teşkil edecektir. EEG sinyali nonlinear yapısıyla kişiye özgüdür ve taklit edilmesi neredeyse imkansızdır. Böyle bir sistemin tasarlanması aşamasında farklı sinyal işleme ve sınıflama yöntemleri ele alınmaktadır. Bu çalışmada EEG sinyalini genlik skalasından bağımsız bir şekilde incelemeye fırsat verecek doğrusal olmayan Fraktal Boyut, İkinci Dereceden Örnek Entropisi, Nicelikler Grafiği, Görünürlük Grafiği özellikleri kullanılmıştır. Hesaplama yükünü azaltmak için EEG'nin belirgin özelliklerinden olan resting state alfa dalgaları odağa alınmış ve düşük sayıda elektrot (8) kullanılmıştır. Elde edilen özellikler her bir elektrot için ayrı ayrı ele alınmış, sonuçta en büyük ayırt ediciliği gösteren özellik ve elektrotun saptanması amaçlanmıştır. 5 farklı makine öğrenmesi metoduyla sınıflama gerçekleştirilmiştir. En yüksek doğruluk Rastgele Orman algoritmasına aittir. En ayırt edici elektrot ve özellikler F5 elektrotu Fraktal Boyutu ve Oz elektrotu Fractal Boyutu özelliği olarak bulunmuştur.**

***Anahtar Kelimeler*—EEG; biometrik; Fraktal boyut; İkinci dereceden örnek entropisi; Nicelikler Grafiği; Görünürlük Grafiği**

I. INTRODUCTION

The electrical activities obtained through electroencephalography (EEG) recordings contain certain biometric characteristics. Clinicians can identify patterns and characteristics in these recordings that serve as unique identifiers specific to each patient. While the level of biometric precision in EEG may not yet match that of fingerprints, its potential as a reliable method for human identification warrants further investigation.Exploring the EEG as a biometric tool could enhance the field of personal identification and contribute to advancements in security and medical diagnostics [1]–[4].

Each biometric method undergoes evaluation based on seven criteria: universality, uniqueness, permanence, measurability, performance, acceptability, and circumvention [5]. Notably, fingerprints are widely recognized for their superior performance in identification or verification. Nonetheless, instances have been documented where fingerprint-based authentication systems were compromised by counterfeit fingers [6]. This vulnerability arises from the exposure of fingerprints on the body's surface.

In terms of resistance to such circumvention, biometrics stored within the body, such as vein patterns, offer greater reliability. However, there have been instances where even vein-based authentication systems accepted anomalies during enrollment and verification processes. This issue stems from the absence of liveness detection, which determines whether an object is a component of a living organism. The implementation of a liveness detection scheme is imperative to safeguard biometric authentication systems against spoofing using counterfeit artifacts. Additionally, a study [7] observed

that biometric features did not change significantly with age in individuals followed over a period of 3 years.

In this paper, we examine the characteristics of the EEG as a biometrics by examining 4 nonlinear features; Fractal Dimension (FD), Quadratic Sample Entropy (QSE), Quantile Graph (QG) and Visibility Graph (VG). These features are used for complexity and similarity analysis while eliminating the need for signal normalization. They focus on extracting fundamental characteristics by rendering the one-dimensional signal in the time domain independent of its magnitude. Self scaling ability of these methods provide value more easy to processed by some machine learning methods.

Conventional approaches typically do not involve specific EEG recordings; all electrodes are used for calculation and acquisition times are long. Furthermore, certain approaches utilized autoregressive (AR) modeling [8] for feature extraction and neural networks as a learning algorithm for verification. Despite achieving verification rates exceeding 90%, their substantial computational demands rendered them unsuitable for practical use. To reduce computational load and identify features that can adapt to an online process, a combination of 8 channels(which are) and 4 nonlinear characteristics were tested. Thus, the discriminative power of individual channels and features can be evaluated.

II. Materials & Methods

A. EEG & Brain Wave

EEG signals measure post-synaptic brain activities through electrical potentials from the scalp using lightweight and non-invasive devices. Electrical changes detected macroscopically on the scalp using an electrode are defined as brain waves. These waves are illustrated in a graph called an Electroencephalogram, where the horizontal axis represents time and the vertical axis represents voltage. Generally brain waves are categorized into five bands: delta (0.5-3Hz), theta (4-7Hz), alpha (8-13Hz), beta (14-30Hz), and gamma (>30Hz). The alpha wave emerges when the eyes are closed and relaxed; this wave characteristic is considered an oscillator structure that provides information transfer. Due to the being individual alpha oscillations and their ease of detection, they have been used as a biomarker in various studies [9].

Due to its high temporal resolution and rich dynamics, it is considered one of the most promising biological signals for biometric applications, alongside others such as electromyography (EMG) and electrocardiography (ECG). Furthermore, EEG presents several peculiarities that make it more advantageous than traditional biometric modalities such as fingerprints and iris scans in terms of robustness against spoofing attacks, privacy compliance, and aliveness detection [10].

B. Dataset

The dataset used in the study was obtained from eyes closed resting state EEG recordings of 20 Alzheimer's patients. The individuals in the set were selected as 10 men and 10 women with similar ages and demographic characteristics. Recordings are realized with Ant Neuro sports EEG recording devices with 10-20 electrode systems. Common Average Referencing (CAR) was chosen to mitigate spatial biases, enhance the comparability of EEG signals across channels, and optimize the computation of various metrics such as fractal dimension, quadratic sample entropy, quantiles graph, and visibility graph. Original recordings has 500 Hz sampling rate for decreasing the computational load, downsampled to 200 Hz. 1-45 Hz filter applied on recordings.

The recordings were taken with 21 electrodes in addition to one reference electrode in the 10-20 electrode system. To reduce computational load, 8 of these electrodes were selected. The electrodes (Fpz, F3, F4, C3, C4, P3, P4, and Oz) were chosen from points close to the zentrum line.

C. Preprocessing

Artifact removal and external noise detection were performed via Wavelet enhanced Independent Component Analysis (wICA) method [11]. This method, known for its effectiveness in distinguishing neural signals from various forms of noise, was implemented using a specialized MATLAB toolbox. After wICA, a predefined function incorporating the Artifact Subspace Reconstruction (ASR) method [12] was used to refine the data further to eliminate noisy segments from the EEG traces. These methods were used in the EEGLAB toolbox [13] in MATLAB. Besides automatic cleaning, data is visually inspected for abnormalities and then epoched within 5 seconds (1000 samples). Epochs are chosen manually from the regions where the Alpha wave is at the original signal. There is no 8-12 Hz filter usage for it. Alpha waves, which can be observed when the eyes are closed, were chosen.

D. Nonlinear Features

The nonlinear methods used for feature extraction produce derived quantities based on the form of the signal, which are independent of the amplitude scale of the 1-D temporal EEG data. This approach helps to overcome the scaling problem of EEG data. It allows for the elimination of quantities that affect all channels, such as electromagnetic-field noise, without the need for filtering. Due to these characteristics, these methods are capable of addressing a broader range of data sets.

Fractal dimension (FD) is a derived quantity used to quantify the similarity and complexity within a signal, offering insights into its structural properties. In this study, we utilized the fractal dimension algorithm proposed by Katz [14]. According to this algorithm, $FD = 1.0$ for a straight line, $FD = 1.15$ for a random signal, and $FD = 1.5$ for a signal with the highest similarity. An increase in FD indicates an increase in the signal's internal similarity and complexity.

If n is the length of a time series x_i, the length of the waveform L is calculated with the Equation 1, the diameter of the waveform d is calculated with the Equation 2 and The Katz fractal dimension FD of the signal is calculated using the Equation 3.

$$L = \sum_{i=1}^{n-1} \sqrt{(x_{i+1} - x_i)^2 + 1} \tag{1}$$

$$d = \max(\sqrt{(x_i - x_1)^2 + (1 - i)^2}) \quad (2)$$

$$FD = \frac{\log(n)}{\log(n) + \log(d/L)} \quad (3)$$

In this study, we used entropy which serves as a crucial metric in quantifying the randomness and unpredictability of signals, with higher entropy indicating greater irregularity and uncertainty. In the context of physiological signals like EEG, methods such as approximate entropy (ApEn) have been developed to assess entropy from time series data. However, since ApEn is parameter-dependent, Richman and Moorman (2000) [15] recommended the Sample Entropy (SampEn) method and SampEn is preferred for processing physiological signals [16].

ApEn and SampEn are constrained by a limited range of parameters. To overcome this limitation, Lake and Moorman proposed the quadratic sample entropy (QSE) method [17]. By expanding the parameter scale, QSE enables the analysis of signals with a wider range of parameters, facilitating the examination of diverse physiological phenomena.

SampEn calculates the matches within signal segments of length m and tolerance r. Let B_i represent the number of matches at a distance of m for each segment, and A_i represent the number of matches at a distance of $m + 1$. "Detailed information about the equations can be found in the study [18]." SampEn is then calculated as the negative natural logarithm of the conditional probability of these two cases, normalized by the entire tolerance window as shown in Equation 4.

$$\text{SampEn} = -\ln\left(\frac{A^m(r)}{B^m(r)}\right) \quad (4)$$

To eliminate the dependency of SampEn on the tolerance value r, this probability function is normalized by dividing it by twice the tolerance window, resulting in the quadratic sample entropy (QSE) criterion as given in Equation 5.

$$\text{QSE} = \text{SampEn} + \ln(2r) \quad (5)$$

The quantile graph (QG) method is a time series analysis approach used to construct a network model based on the quantiles of the time series data. In this method, the time series is partitioned into Q quantiles, denoted as $q_1, q_2, ..., q_Q$. Each quantile serves as a node in the resulting network model, and transitions between quantiles are represented as edges.

Consequently, the complex network generated by the QG method can be represented as $g = \{M, R\}$, where M represents the number of nodes and R represents the number of edges. To construct the network, an adjacency matrix is created, summarizing the connections between nodes by counting the number of reachable nodes after a certain time k from a given node at time t in the time series. After applying suitable normalization, the adjacency matrix is transformed into a Markov transition matrix [19].

The topology and other characteristics of the resulting network model provide insights into the time series [20]. In the study the average jump length serves as a complexity measure, offering insights into the extent of transition variability between quantiles over a specified number of steps.

The average jump length ($\Delta(k)$) is calculated for various neighborhood coefficients (k). This metric quantifies the average change in transition between quantiles over k steps and is computed with the following Equation 6:

$$\Delta(k) = \frac{1}{Q} tr(PW_k^T) \quad (6)$$

Here, W_k represents the normalized adjacency matrix, W_k^T denotes the transpose of this matrix, P signifies the distance between quantiles, and tr represents the trace operation.

The last metric used in this study was the visibility graph (VG). Like the quantile graph approach, the VG method facilitates the transformation of time series data into a network structure, enabling the application of complex network theory principles. This algorithm, proposed by Lacasa, is particularly adept at converting fractal structures within the time series into scale-free graphs [21].

In the VG network structure, the points in the time series serve as nodes, and edges represent pairs of points that have a direct line of sight to each other. The visibility between two nodes i and j is determined using a trigonometric formula involving intervening points k is calculated with Equation 7.

$$x(k) \leqslant x(j) + (x(i) - x(j))\frac{j - k}{j - i} \quad (7)$$

This process results in an unweighted orthogonal adjacency matrix of size $T \times T$, where T is the length of the time series. The complexity of this matrix is utilized as a feature of the visibility graph, termed Graph Index Complexity (GIC) [22] which is calculated as in Equation 8.

$$\text{GIC} = 4c(1 - c) \quad (8)$$

Where c is computed with the formula in Equation 9.

$$c = \frac{\lambda_{\max} - 2\cos\left(\frac{\pi}{N+1}\right)}{N - 1 - 2\cos\left(\frac{\pi}{N+1}\right)} \quad (9)$$

Here, $\lambda_{\max}$ corresponds to the largest eigenvalue of the adjacency matrix, and N represents the number of nodes in the visibility graph. In this study, the Lanczos algorithm [23] was applied for eigenvalue computations to mitigate computational costs, ensuring efficient processing of large-scale EEG datasets.

E. Machine Learning Methods

Five different machine learning (ML) methods were tested for the classification of the extracted features. The efficiency of these methods will be evaluated based on specific metrics.

- Multinominal Naïve Bayes (MultiNB)
- K-nearest neighbor algorithm (KNN)

- Support Vector Machine (SVM)
- Decision Tree Classifier (DTC)
- Random Forest Classifier (RFC)

MultiNB is a generative model. MultiNB is preferred in the case where the variables are independent, i.e., we assume that the non-diagonal elements of the covariance matrix are 0. However, if we cannot measure the dependence between the variables, then we can apply this classification method by considering the variables as independent [24]. The model, which uses Bayes' Theorem, is based on conditional probability calculation with the help of probabilistic expressions obtained by using the statistical properties of the data. We think that, as if we first pick a class C at random by sampling from P(C), and then having fixed C, we pick an x by sampling from p(x|C). Bayes' rule inverts the generative direction and allows a diagnosis;

$$P(x) = \frac{P(C)P(C)}{P(x)} \tag{10}$$

In this case, we can use a test set to find the conditional probability of a feature for the class we are currently examining. In this way, the classification is performed by calculating and summing the relationship between each feature and the class in the context of conditional probability. It is a classifier with a high computational load because the calculations will be repeated for each variable and class.

K nearest neighbor algorithm is a nonparametric supervised learning algorithm. It uses proximity information to classify or make predictions about a data point. It is also used for regression along with classification. It is based on the assumption that similar points will be close to each other. Classification is carried out by label assignment, called majority voting. While doing this, the distances of the existing data with its neighbors are calculated and the classes of the neighbors are kept. If the available data gets enough votes in the majority vote, it is considered to belong to that class. Usually, the Euclidian distance is calculated as a measure of distance.

$$d(x, y) = \sqrt{\sum_{i=1}^{n} (y_i - x_i)} \tag{11}$$

The K value represents the neighborhood degrees. Considering the value K = 1, the current point is only assigned to the class label with the nearest neighbor. During the training, the distance between the current point and all other points is calculated and the classification is performed according to the result. In short, KNN is a supervised learning algorithm that simply stores labeled training examples during the training phase.

Support vector machines (SVM) are one of the most robust prediction methods within the statistical learning framework. The decision support method is basically a function estimation problem [25]. If we consider y as a high-dimensional output vector and z as a high-dimensional feature vector;

$$y = (w^* z) + b \tag{12}$$

we can separate binary and larger data classes with a linear function. It tries to make this separation process in a way that maximizes the gap between classes. In the case of when given a binary training set, the decision support machine creates a non-probabilistic binary linear model in which newly entered data is assigned to one class or another. In SVCs, the classes of newly entered data are predicted based on which side of the separator space falls.

Since SVMs provide separation with linear methods, they have problems with datasets that cannot be linearly separable. At this point, kernel tricks are used and some modifications are applied to the data. These modifications aim to make the data linearly separable without corrupting the separable structure of the class.

Decision trees are the structures that provide the most information to the user about the operation of the algorithm among machine learning algorithms. The algorithm is apparent due to the progression structures similar to flow charts. Information such as the progress of the algorithm, the level of progress, the stage where the error is high, and the branching structure can be easily obtained. The structure of the branching can be changed depending on preferences. Decision trees are one of the most popular machine learning methods due to their simple and intelligible structure and informative properties [26].

Forest classifiers, as the name suggests, are ensemble methods in which more than one decision tree classifier is used. Decision tree structures are preferred in many areas where large data are processed, such as data mining, due to their interpretability. RFC is a homogeneous ensemble method using only the decision tree structure and provides a solution for the overfitting problem encountered in decision trees; increases accuracy. Tree structures that reach very deep structures tend to learn irregular patterns, i.e. they have low bias and high variance. The random forest structure aims to reduce the variance by averaging multiple deep trees trained from different parts of the training set [27]. Although this causes a small increase in bias and a decrease in the interpretability of the model, it generally improves the performance of the model.

The random forest consists of a collection of tree-structured classifiers $\{h(x, \Theta_k), k = 1, \ldots\}$. Here $\{\Theta_k\}$ are independent identically distributed random vectors, and each tree votes one unit for the most popular class in entry x [28].

In this way, the random forest classifier increases the accuracy by highlighting the best of its methods, and the increase in accuracy depends on the decision trees used in the algorithm. Features are allocated to tree structures, which in a way allows the features that are effective in class selection to come to the fore. In order to avoid the overfitting problem in random forest structures, methods are used such as limiting the depth of the trees and the number of branches. In addition, giving the inputs randomly to the trees reduces overfitting.

F. Evaluations Metrics

There are many metrics besides accuracy in evaluating the performance of machine learning methods. These aim to

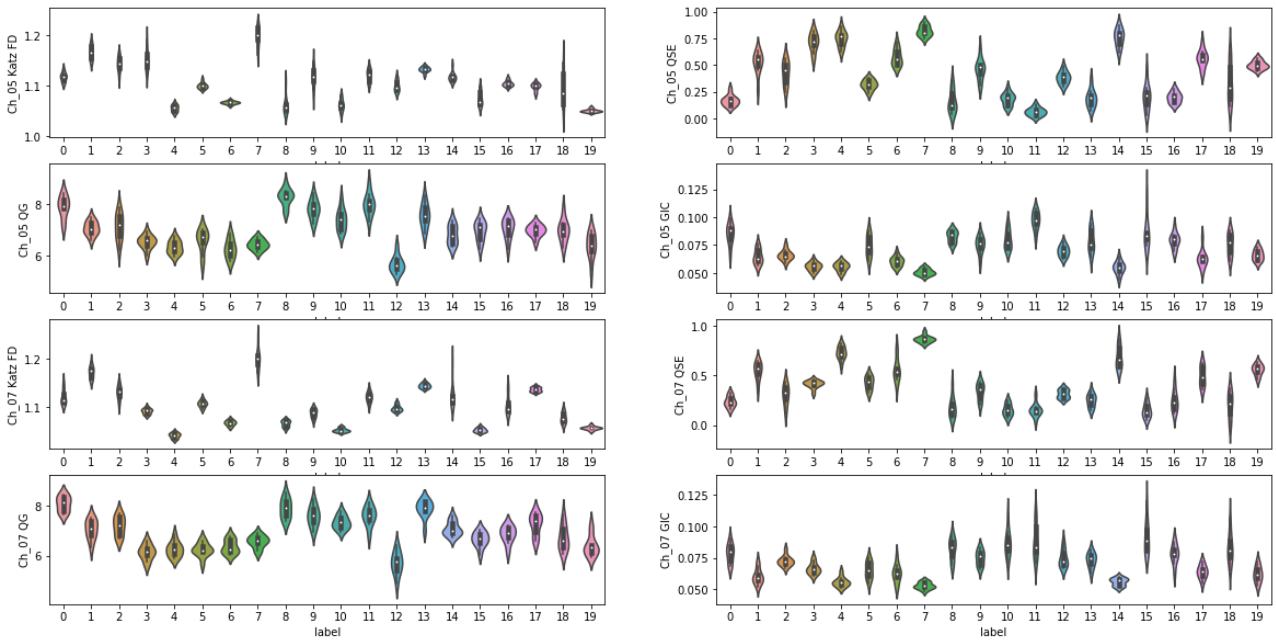

Figure 1: Person-based distribution of features for 2 electrodes

measure different features of the method such as selectivity, generalization ability, and sensitivity.

$$\text{Accuracy} = \frac{\text{True Numbers}}{\text{Total Size of Prediction}} \tag{13}$$

Hamming loss can be used as the opposite of the accuracy value. Hamming loss is the ratio of false labels to the total number of labels.

Apart from these, although the accuracy of the method is high, it is useful to look at metrics such as precision and recall. The data set used may not always have a balanced distribution. In this case, although the accuracy is high, this may be due to the abundance of instances in the data that are more likely to be correct. Precision is a measure of how well the method, on a per class basis, separate that class from the others. It measures how precise the method is when viewed for all features. This score is included in the study by taking the average for all classes.

$$\text{Precision} = \frac{TP}{(TP + FP)} \tag{14}$$

Another metric in this context is the Recall value. This metric, more commonly called sensitivity in the medical field, is a measure of how accurately the class label currently under investigation is estimated [29]. This score is included in the study by taking the average for all classes.

$$\text{Recall} = \text{Sensitivity} = \frac{TP}{(TP + FN)} \tag{15}$$

Another metric frequently used in the medical field is the specificity [12]. This metric is a measure of how well the classes other than the current class are predicted. In a way, this metric measures how much false alarm the method gives. It is mostly used in binary applications and is the 1's complement of the Recall metric. For this reason, in this study, the Recall metric is included instead of specificity.

$$\text{Specificity} = \frac{TN}{(TN + FP)} \tag{16}$$

Some combined methods are used for performance measurement. One of these metrics, the f1 score, consists of a combination of Precision and Recall values. There is an inverse relationship between Precision and Recall, one may need to be sacrificed to increase the other one. F1 score extracts a criterion from the conflict of these two values.

$$F_1 = \frac{2 \cdot \text{ Precision } \cdot \text{ Recall}}{\text{Precision } + \text{ Recall}} \tag{17}$$

This score is included in the study by taking the average for all classes.

III. Results

The distributions of the obtained nonlinear features for 4 electrodes are shown in Fig. 1. As seen in the distribution graphs, locational changes have less impact on the distribution compared to feature changes. The features show more varied distributions among individuals and provide clearer information for differentiation.

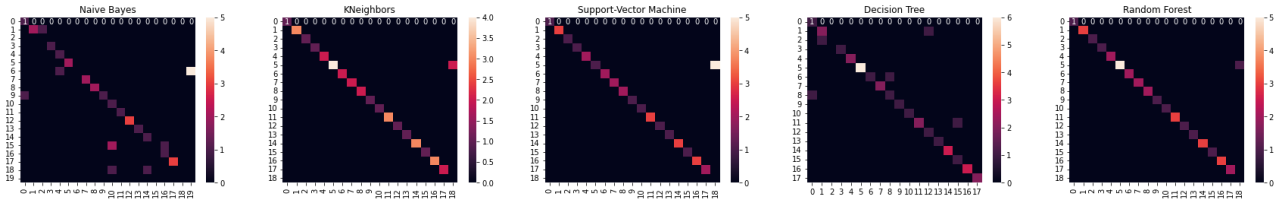

Figure 2: Confusion Matrices

	Methods				
	MultiNB	KNN	SVM	DTC	RFC
Accuracy	0.639	0.944	0.861	0.861	0.972
Hamming Loss	0.361	0.056	0.139	0.139	0.028
f1 score	0.627	0.937	0.91	0.813	0.943
Sensitivity	0.612	0.947	0.947	0.815	0.947
Selectivity	0.708	0.93	0.904	0.852	0.939

Table I: Evaluation metrics results

According to the evaluation metrics most successful method is the Random Forest algorithm. The algorithm, which uses a 100-tree structure with default values, uses the "gini" method as the evaluation criterion. By using different parameters, gives more low accuracy values. Therefore results of these parameters are not included the work.

After the Random Forest algorithm, the next most successful method was KNN with an accuracy of 0.94. In the KNN algorithm, 5 neighborhoods are looked at, and the method to be used in determining the most neighbors is automatically determined. Euclidean distance is used to calculate the distance between neighbors. With the current parameters, the classifier has high accuracy after RFC. This result gives the idea that class properties form clusters in multidimensional space.

Following them, SVM and Decision Tree algorithms come with slightly lower accuracy. RFC is an algorithm where a large number of Decision Trees are tried with different parameters. As a result, when the optimal parameter values obtained in RFC are applied to Decision Tree, the same result will be obtained. The high accuracy value of SVC may be due to the implementation of the kernel trick in the default parameters. When the kernel trick method is taken as linear, the accuracy of the algorithm decreased a lot.

MultiNB has very low accuracy. The overlap of feature distributions can make them appear dependent, which is not suitable for an algorithm based on Bayes' theorem. Also, obtaining poor results when the kernel trick is not applied in the SVM algorithm suggests that the features have low linear separability.

IV. Conclusion

In the authentication using the brain wave, we proposed nonlinear features with different electrodes. In doing so, we preferred epochs where we observed alpha oscillations. We demonstrated that these oscillations, which are thought to be related to information transmission, are unique to individuals and suitable for use as a biometric marker. Various methods were tried in processing the extracted features to select the one that could provide the best identification. According to the obtained results, the RFC method distinguished individuals with high accuracy. Due to the method's traceability advantage, the most important features were identified as the Fractal Dimension of the F5 electrode and the Fractal Dimension of the Oz electrode. The results are consistent with study [9] suggesting that biometric identification can be performed with a smaller number of electrodes. It suggests that a biometric system could be developed using only these two electrodes.

References

[1] E. Sayilgan, Y. K. Yuce, and Y. Isler, "Prediction of evoking frequency from steady-state visual evoked frequency," *Natural and Engineering Sciences*, vol. 4(3), pp. 91–99, 2019.

[2] E. Sayilgan, Y. K. Yuce, and Y. Isler, "Estimation of three distinct commands using fourier transform of steady-state visual-evoked potentials," *Duzce Universitesi Bilim ve Teknoloji Dergisi*, vol. 8(4), pp. 2337–2343, 2020.

[3] E. Sayilgan, Y. K. Yuce, and Y. Isler, "Frequency recognition from temporal and frequency depth of the brain-computer interface based on steady-state visual evoked potentials," *Journal of Intelligent Systems with Applications*, vol. 4(1), pp. 68–73, 2021.

[4] M. Degirmenci, Y. K. Yuce, and Y. Isler, "Motor imaginary task classification using statistically significant time domain and frequency domain eeg features," *Journal of Intelligent Systems with Applications*, vol. 5(1), pp. 49–54, 2022.

[5] A. K. Jain, R. Bolle, and S. Pankanti, *Biometrics: personal identification in networked society*, vol. 479. Springer Science & Business Media, 2006.

[6] T. Matsumoto, H. Matsumoto, K. Yamada, and S. Hoshino, "Impact of artificial" gummy" fingers on fingerprint systems," in *Optical security and counterfeit deterrence techniques IV*, vol. 4677, pp. 275–289, SPIE, 2002.

[7] E. Maiorana and P. Campisi, "Longitudinal evaluation of eeg-based biometric recognition," *IEEE transactions on Information Forensics and Security*, vol. 13, no. 5, pp. 1123–1138, 2017.

[8] R. Paranjape, J. Mahovsky, L. Benedicenti, and Z. Koles, "The electroencephalogram as a biometric," in *Canadian Conference on Electrical and Computer Engineering 2001. Conference Proceedings (Cat. No. 01TH8555)*, vol. 2, pp. 1363–1366, IEEE, 2001.

[9] I. Nakanishi, S. Baba, and C. Miyamoto, "Eeg based biometric authentication using new spectral features," in *2009 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*, pp. 651–654, IEEE, 2009.

[10] M. Wang, J. Hu, and H. A. Abbass, "Brainprint: Eeg biometric identification based on analyzing brain connectivity graphs," *Pattern Recognition*, vol. 105, p. 107381, 2020.

[11] N. P. Castellanos and V. A. Makarov, "Recovering eeg brain signals: Artifact suppression with wavelet enhanced independent component analysis," *Journal of neuroscience methods*, vol. 158, no. 2, pp. 300–312, 2006.

[12] C. A. Kothe and S. Makeig, "Bcilab: a platform for brain–computer interface development," *Journal of neural engineering*, vol. 10, no. 5, p. 056014, 2013.

[13] A. Delorme and S. Makeig, "Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis," *Journal of neuroscience methods*, vol. 134, no. 1, pp. 9–21, 2004.

[14] M. J. Katz, "Fractals and the analysis of waveforms," *Computers in biology and medicine*, vol. 18, no. 3, pp. 145–156, 1988.

[15] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *American journal of physiology-heart and circulatory physiology*, vol. 278, no. 6, pp. H2039–H2049, 2000.

[16] A. Delgado-Bonal and A. Marshak, "Approximate entropy and sample entropy: A comprehensive tutorial," *Entropy*, vol. 21, no. 6, p. 541, 2019.

[17] D. E. Lake and J. R. Moorman, "Accurate estimation of entropy in very short physiological time series: the problem of atrial fibrillation detection in implanted ventricular devices," *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 300, no. 1, pp. H319–H325, 2011.

[18] S. Simons, D. Abasolo, and J. Escudero, "Classification of alzheimer's disease from quadratic sample entropy of electroencephalogram," *Healthcare technology letters*, vol. 2, no. 3, pp. 70–73, 2015.

[19] A. S. Campanharo, E. Doescher, and F. M. Ramos, "Automated eeg signals analysis using quantile graphs," in *Advances in Computational Intelligence: 14th International Work-Conference on Artificial Neural Networks, IWANN 2017, Cadiz, Spain, June 14-16, 2017, Proceedings, Part II 14*, pp. 95–103, Springer, 2017.

[20] A. M. Pineda, F. M. Ramos, L. E. Betting, and A. S. Campanharo, "Quantile graphs for eeg-based diagnosis of alzheimer's disease," *Plos one*, vol. 15, no. 6, p. e0231169, 2020.

[21] L. Lacasa, B. Luque, F. Ballesteros, J. Luque, and J. C. Nuno, "From time series to complex networks: The visibility graph," *Proceedings of the National Academy of Sciences*, vol. 105, no. 13, pp. 4972–4975, 2008.

[22] M. Ahmadlou, H. Adeli, and A. Adeli, "New diagnostic eeg markers of the alzheimer's disease using visibility graph," *Journal of neural transmission*, vol. 117, pp. 1099–1109, 2010.

[23] C. Lanczos, "An iteration method for the solution of the eigenvalue problem of linear differential and integral operators," *Journal of Research of the National Bureau of Standards*, 1950.

[24] E. Alpaydin, *Introduction to machine learning*. MIT press, 2020.

[25] V. N. Vapnik, "The support vector method," in *International conference on artificial neural networks*, pp. 261–271, Springer, 1997.

[26] X. Wu, V. Kumar, J. Ross Quinlan, J. Ghosh, Q. Yang, H. Motoda, G. J. McLachlan, A. Ng, B. Liu, P. S. Yu, *et al.*, "Top 10 algorithms in data mining," *Knowledge and information systems*, vol. 14, pp. 1–37, 2008.

[27] T. Hastie, R. Tibshirani, J. H. Friedman, and J. H. Friedman, *The elements of statistical learning: data mining, inference, and prediction*, vol. 2. Springer, 2009.

[28] L. Breiman, "Random forests," *Machine learning*, vol. 45, pp. 5–32, 2001.

[29] N. B. Smith and A. Webb, *Introduction to medical imaging: physics, engineering and clinical applications*. Cambridge university press, 2010.