POLİTEKNİK DERGİSİ

JOURNAL of POLYTECHNIC

ISSN: 1302-0900 (PRINT), ISSN: 2147-9429 (ONLINE)
URL: http://dergipark.org.tr/politeknik

Electromyography based hand movement classification and feature extraction using machine learning algorithms

***Yazar(lar) (Author(s))**: Ekin EKİNCİ[1], Zeynep GARİP[2], Kasım SERBEST[3]*

ORCID[1]: 0000-0003-0658-592X

ORCID[2]: 0000-0002-0420-8541

ORCID[3]: 0000-0002-0064-4020

To cite to this article: Ekinci E., Garip Z. and Serbest K., "Electromyography based hand movement classification and feature extraction using machine learning algorithms", ***Journal of Polytechnic***, 26(4): 1621-1633, (2023).

Bu makaleye şu şekilde atıfta bulunabilirsiniz: Ekinci E., Garip Z. ve Serbest K., "Electromyography based hand movement classification and feature extraction using machine learning algorithms", ***Politeknik Dergisi***, 26(4): 1621-1633, (2023).

Erişim linki (To link to this article): http://dergipark.org.tr/politeknik/archive

DOI: 10.2339/politeknik.1348121

Electromyography Based Hand Movement Classification and Feature Extraction Using Machine Learning Algorithms

Highlights

- *Gold-standard EMG measurement*
- *Six time-domain features*
- *Evaluation metrics (Confusion matrix, F1-score, cross-validation)*

Graphical Abstract

This study utilized electromyography data collected from the extensor digitorum and flexor carpi radialis muscles to classify open and closed hand positions using various machine learning algorithms.

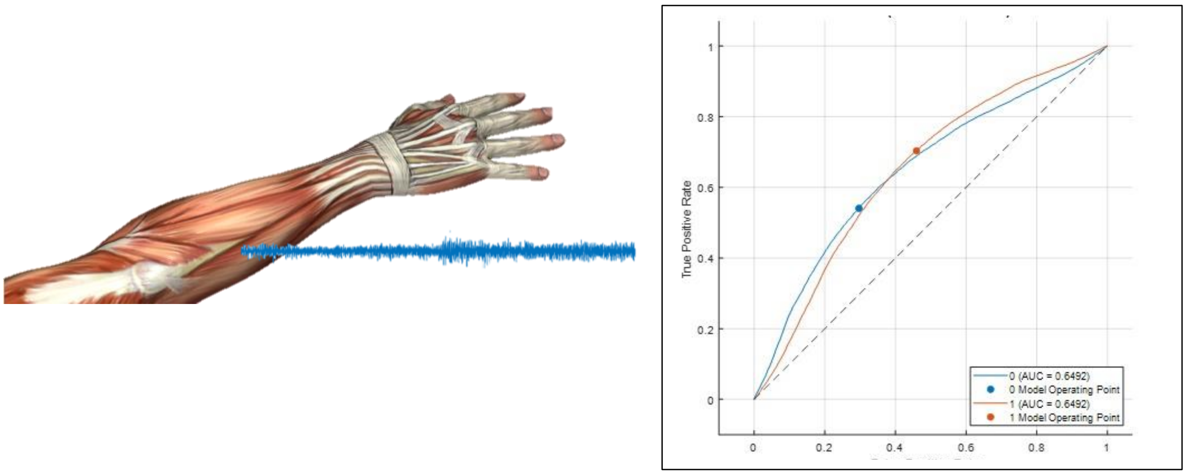

Figure. Representation of the study

Aim

The aim of the study to enhance the field of EMG-based hand movement classification by thoroughly investigating different methods for extracting features and utilizing machine learning algorithms.

Design & Methodology

A binary classification was carried out on humans by using EMG signals. Throughout the study two-channel EMG signals from four participants were obtained. Then, six different time-domain features were extracted from the raw signals. And classification algorithms were compared in terms of F1-score and area under curve (AUC).

Originality

A gold-standard EMG sensor is used and the classification performance of different ML algorithms is determined with additional features.

Findings

Among the tested algorithms, SVM achieved the highest success rate with a maximum accuracy of 73.1%, while KNN yielded the lowest success rate at a minimum accuracy of 55.9%.

Conclusion

This research highlights the influence of hand positioning on the challenge of predicting movements and emphasizes the importance of individual characteristics in determining how well algorithms perform.

Declaration of Ethical Standards

The author(s) of this article declare that the materials and methods used in this study do not require ethical committee permission and/or legal-special permission.

Electromyography Based Hand Movement Classification and Feature Extraction Using Machine Learning Algorithms

Research Article

Ekin EKİNCİ[1], Zeynep GARİP[1], Kasım SERBEST[2*]

[1]Sakarya University of Applied Sciences, Faculty of Technology, Computer Engineering Department, Türkiye
[2]Sakarya University of Applied Sciences, Faculty of Technology, Mechatronics Engineering Department, Türkiye

(Received : 22.08.2023 ; Accepted : 19.09.2023 ; Early View : 13.11.2023)

ABSTRACT

The categorization of hand gestures holds significant importance in controlling orthotic and prosthetic devices, enabling human-machine interaction, and facilitating telerehabilitation applications. For many years, methods of motion analysis based on image processing techniques have been employed to detect hand motions. However, recent research has focused on utilizing muscle contraction for detecting hand movements. Specifically, there has been an increase in studies that classify hand movements using surface electromyography (sEMG) data from the muscles of the hand and arm. In our study, we estimated the open (extension of the fingers) and closed (flexion of the fingers) positions of the hand by analyzing EMG data obtained from 4 volunteer participants' extensor digitorum and flexor carpi radialis muscles. In order to accurately discriminate EMG signals, various statistical measures such as variance, standard deviation, root mean square, average energy, minimum and maximum features were utilized. The dataset containing these additional features was then subjected to classification algorithms including Support Vector Machines (SVM), K Nearest Neighbour (KNN), Decision Tree (DT), and Gaussian Naive Bayes (GNB) for the purpose of classifying hand positions into open or closed states. Among the tested algorithms, SVM achieved the highest success rate with a maximum accuracy of 73.1%, while KNN yielded the lowest success rate at a minimum accuracy of 55.9%. To further enhance prediction accuracy in future studies, it is suggested that data from a larger set of muscles be collected.

Keywords: EMG, hand muscle, binary classification, support vector machines (SVM)

1. INTRODUCTION

Electromyography (EMG) is a technique used to measure the electrical activity of muscles [1]. EMG signals are obtained by recording the electrical potentials that occur during contraction and relaxation of muscle fibers. These signals are a reflection of the electrical activity of muscles in the process of producing movement. The main purpose of EMG is to study the way muscles work and their level of activity, to understand the nature of muscle activity and to assess muscle health [2].

In recent years, it has been a growing acknowledgement of the significance of electromyography as an indispensable tool for comprehending and examining human movements. EMG data is extensively employed in various fields, including studying muscular activity, movement control analysis, and evaluating muscle dysfunctions [3]. The broad range of applications demonstrates its immense value in multiple domains such as sports performance analysis, detection of muscle fatigue indications, rehabilitation management processes and prosthetic joint control strategies which carry significant implications. The application of EMG technology to sports performance analysis facilitates a detailed understanding of the neuromuscular aspects associated with athletic movements.

By monitoring electrical signals generated by muscles during physical activities or exercises using surface electrodes placed on the skin above specific muscles or motor units connections – researchers are able to assess patterns that may indicate factors affecting both capacity and efficiency within sporting performances [4]. Moreover, detecting muscle fatigue through EMG enables identifying gradual declines in muscular abilities over time due to exhaustive exercise or repetitive motions. This information could be crucial not only for athletes but also for patients undergoing rehabilitation programs from injuries where managing excessive exertion can play an essential role [5]. Regarding these scenarios, clinical practitioners employ pattern recognition algorithms based upon continuous measurements derived from real-time recordings carried out during varied tasks.

The utilization of electromyography data from hand muscles for movement classification is a topic that has garnered significant attention in recent years. Hand muscles are pivotal in facilitating various daily activities, thus it becomes imperative to accurately and efficiently classify different hand movements [6]. EMG data holds valuable insights into the functioning of hand muscles, enabling differentiation between distinct types of movements. The precise categorization of this data carries profound implications for an array of applications such as optimizing muscle control, enhancing athletic performance, improving rehabilitation procedures, and augmenting the functionality of prostheses and orthoses. By leveraging accurate movement classification derived from EMG data analysis, numerous advancements can be made with regard to human-machine interaction.

*Sorumlu yazar (Corresponding Author)
e-mail : kserbest@subu.edu.tr

Machine Learning (ML) algorithms successfully handle real-world problems that are challenging to model directly. Due to the complexity of the domain, EMG classification applications are therefore particularly ideal for successfully employing ML approaches. In the literature, ML algorithms are successfully applied in classification of EMG signals. The effectiveness of the classification was evaluated using algorithms such as Convolutional Neural Networks (CNN) [7], Bidirectional Long Short-Term Memory (Bi-LSTM), Artificial Neural Network (ANN), SVM, XGBoost, Multi-layer Perceptron (MLP) [8], hybrid algorithms.

Sahu et al. developed a new algorithm based on the artificial bee colony (ABC) to classify 17 different prosthetic hand movements. This proposed algorithm performed an effective EMG feature selection by giving an average classification of 94.13% and a maximum classification accuracy of 97.06% [9]. In the study conducted by Fu et al., objective measurements of the data obtained from surface EMG measurements placed according to the phenomenon of resting in an office chair in a supine sitting position were obtained [7]. These measurements were analyzed using linear regression and CNN networks. According to the experimental results, they provided classification with 0.99 accuracy with CNN model. Karnam et al. proposed a hybrid architecture called EMGHandNet based on CNN and Bi-LSTM. Using the proposed architecture, they performed hand activity classification of NinaPro DB1, NinaPro DB2, NinaPro DB4, BioPatRec DB2 and UCI Gesture EMG datasets [8]. Based on the findings of the experiment, it was observed that the proposed model achieved an average classification accuracy of 94.582%. In a study conducted by Fajardo et al., three different approaches for EMG movement classification were presented: handcrafted features obtained through time-spectral analysis, CNN, and a combined approach using both handcrafted and deep features. The results consistently showed that the combined strategy performed better than the other two scenarios [10]. Additionally, Tepe and Demir [11] carried out a study on real-time and non-real-time classification of EMG signals obtained from an armband using SVM algorithm. The experimental outcomes indicated that for non-real-time classification, the highest accuracy recorded was 96.38%, whereas for real-time classification, it reached 99.05%.

Our study endeavors to enhance the field of EMG-based hand movement classification by thoroughly investigating different methods for extracting features and utilizing machine learning algorithms. Expanding upon prior research findings, we aim to tackle the difficulties associated with accurately classifying movements amidst diverse EMG signal characteristics and noise levels. By evaluating and comparing a range of machine learning techniques, including Support Vector Machines (SVM), K Nearest Neighbour (KNN), and other architectures, we aim to contribute insights into the effectiveness of different approaches across diverse datasets. Additionally, our research aims to introduce feature extraction techniques that improve the reliability and flexibility of classification models. This study seeks to provide valuable guidance for researchers and practitioners involved in classifying hand movements based on EMG signals, while also highlighting the current opportunities and challenges in this ever-evolving interdisciplinary field.

In this study, a binary classification was carried out on humans by using dominant hand produced EMG signals; the hand open and hand closed positions were taken into account. Throughout the study two-channel EMG signals from four participants were obtained. Then, six different time-domain features were extracted from the raw signals. And classification algorithms namely, SVM, KNN, Decision Tree (DT) and Gaussian Naive Bayes (GNB) were compared in terms of F1-score and area under curve (AUC).

The rest of the study is organized as follows. Section 2 provides a detailed explanation of the experimental EMG data, feature extraction techniques, and classification algorithms utilized in our study. A comprehensive description of the experimental setup, evaluation metrics employed, and results obtained is presented in Section 3. Lastly, Section 4 concludes with a discussion on the findings.

2. METHODS AND IMPLEMENTATION

This section provides information on the setup used for EMG measurements, methods employed for feature extraction, as well as four distinct machine learning algorithms (SVM, KNN, DT and GNB) that were utilized for classification purposes.

2.1. EMG Data

In order to investigate muscle activity, electromyography measurement was conducted on a group of four participants who volunteered for the study. The group consisted of three male and one female individuals (age: 33.25±2.36 years). Prior to commencing the study, all participants were duly informed about its nature and purpose, and their consent was obtained in accordance with ethical guidelines. To ensure consistency in data collection, it is important to note that none of the participants had any pre-existing health conditions or disorders that could potentially affect their hand movements during EMG measurement sessions. EMG measurement was performed on the dominant hand of the participants.

Surface EMG measurements (Figure 1) were performed on the extensor digitorum (ED) muscle on the dorsal surface of the forearm and the flexor carpi radialis (FCR) muscle on the palmar surface of the forearm to determine the open (extension of the fingers) and closed (flexion of the fingers) positions of the hand [12]. In the measurement process, one Delsys DE surface EMG sensor (Bagnoli Desktop EMG Systems, Delsys Inc., USA) was placed on the ED and FCR [13]. To ensure secure attachment of the sensors to the skin, each sensor was affixed using a 2-slot adhesive skin interface. The reference electrode was positioned at the posterior aspect of the elbow joint on the distal end of the humerus bone. Signals were collected with a Delsys Bagnolli amplifier (8 channels, total gain of 1000).

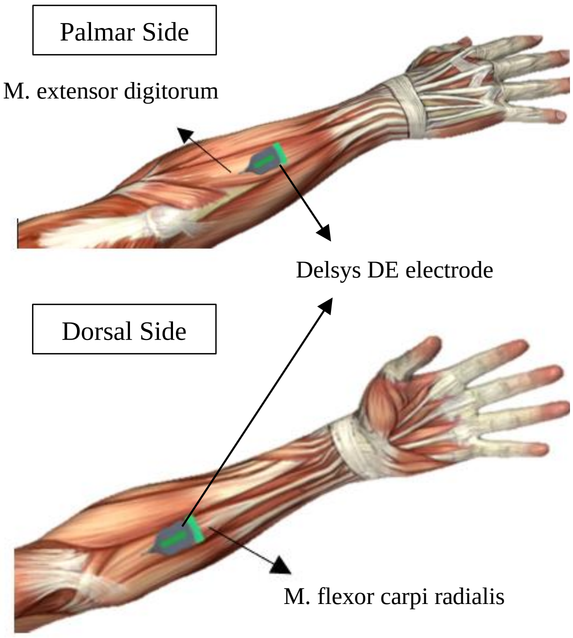

Figure 1. Electrode placement on the forearm for EMG measurement

During the EMG assessment, participants were instructed to maintain their hands in an open position for a duration of 10 seconds, after which they were asked to switch to a closed position for another 10 seconds. This specific protocol was performed thrice with each participant. A resting period of 5 minutes was allocated between each measurement session to minimize any potential fatigue or influence from prior measurements. To ensure unbiased results, participants were not allowed to view the measurement screen during data collection. The raw EMG signals recorded during extension and flexion of the fingers are shown in Figure 2.

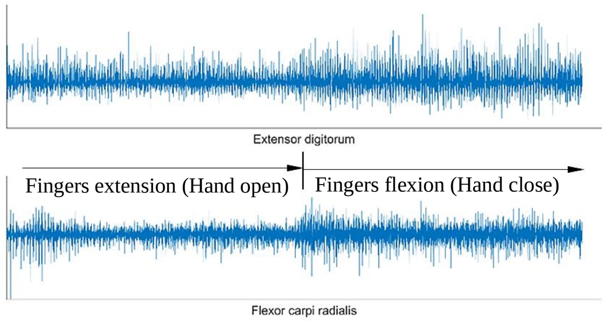

Figure 2. A sample of raw EMG data during the experiments

2.2. Feature Extraction

Since our device has only two channels, the original dataset consists of only two features. Two features may be insufficient for representing data. Therefore, in this paper, we attempt to extract new features to build a reliable classification models and improve the system's level of accuracy.

Time-domain features have a wide application in biomedical research because of direct clinical interpretation ability [14]. They are simple to extract, do not require any type of transformation and high computational burden [15]. We computed six time-domain features that have been successful in the literature in discriminating the EMG signals, namely, variance (var), standard deviation (std), root mean square (rms), average energy (ae), minimum (min) and maximum (max) [16-18]. Table 1 presents the equations of time-domain features.

Table 1. Equations of the features (x_i sample of the signal; N: length of the signal)

Feature	Definition	Equation
VAR	The definition of variance is the average of the square values of the variable's deviation.	$\frac{1}{N-1}\sum_{i=1}^{N} x_i^2$
STD	The distribution of the EMG signal's value is represented by the standard deviation.	$\sqrt{\frac{1}{N-1}\sum_{i=1}^{N} \lvert x_i-\mu\rvert^2}$
RMS	The RMS is the square root of the average EMG signal power over a specific amount of time.	$\sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2}$
AE	The Average Energy (AE) of EMG signal for measuring energy distribution.	$\frac{1}{N}\sum_{i=1}^{N} x_i^2$
MIN	Min is referred to as the EMG signal minimum value	$\min(x_1,\dots x_i,\dots,x_N)$
MAX	Max is referred to as the EMG signal maximum value	$\max(x_1,\dots x_i,\dots,x_N)$

2.3. Machine Learning Algorithm

Machine learning is a statistical learning method that utilizes data to solve complex problems. It has gained popularity in various fields, including information extraction and decision making, due to its ability to process large datasets. This section focuses on the application of machine learning algorithms for classifying hand gestures based on EMG signals. The selection and implementation of these algorithms are influenced by the complexity and diversity of EMG signals. Therefore, this section provides an explanation of how these underlying machine learning algorithms function and why they are advantageous in this paper.

2.3.1. Support Vector Machines (SVM)

SVM is a powerful machine learning (ML) algorithm used for classification problems developed by Vapnik and based on statistical learning theory [19]. It is used to distinguish between classes in a binary or multi-class dataset. SVM works by representing data points in a space and trying to find the best hyperplane for classification [20, 21]. For example, the n data samples are assigned a class label of + 1 and − 1 for given a training dataset (x1, …, xn). The hyperplane is determined to maximize the distance between the classes + 1 and - 1. When finding this hyperplane, SVM uses key data points, called support vectors, located at or near the boundary of the data distribution. The graphical representation of working principle of SVM is shown in Figure 3. Although originally designed for linear problems, SVM can be used where data is not linearly separable. For such data, the kernel method is used. The kernel method is used to move data into a higher dimensional space and make it linearly separable there. In particular, the Radial Based Function (RBF) kernel, Gaussian kernel and the polynomial kernel are commonly used kernel functions in SVM. The SVM algorithm is often used to analyze biomedical data such as EMG, EEG and ECG [22].

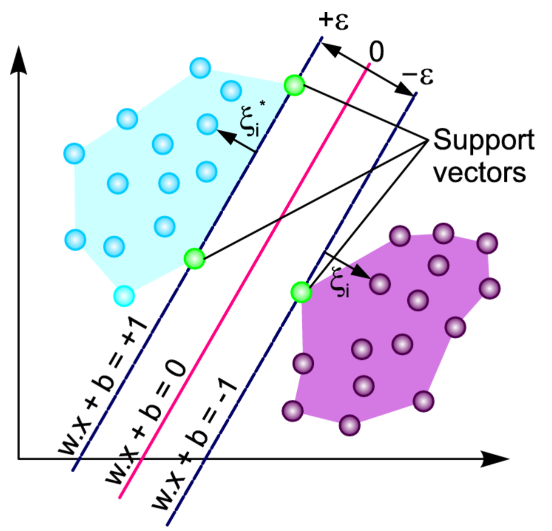

Figure 3. Graphical illustration of working principle of SVM [23]

2.3.2. K Nearest Neighbour (KNN)

KNN, a widely used and straightforward classification algorithm, can effectively classify data points into multiple classes [24]. Data points are represented by their position in a Euclidean space. Classification is done by finding the K nearest neighbour when a new data point arrives, and voting according to the number of classes of those neighbours. The Euclidean distance is used to define the nearest neighbour. The Euclidean distance between two data such as $X_1 = (x_{11}, x_{12}, \ldots, x_{1n})$ and $X_2 = (x_{21}, x_{22}, \ldots, x_{2n})$ is calculated according to equation below.

$$dist(X_1, X_2) = \sqrt{\sum_{i=1}^{n} (x_{1i} - x_{2i})^2} \tag{1}$$

For example, if K=3 and K = 5, the 3 and 5 nearest neighbours are evaluated respectively and the majority class of these neighbours is determined as the predictive class of the new data point as shown in Figure 4.

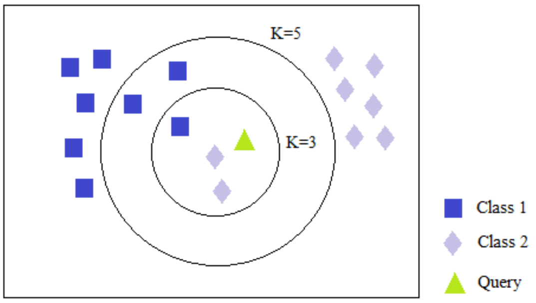

Figure 4. Representation of KNN algorithm for K=3 and K=5

2.3.3. Decision Tree (DT)

DT is a classification technique that is frequently used in data mining and ML, thanks to its simplicity. This algorithm analyzes the data using a set of decision rules and generates a tree structure as a result [25]. Each internal node branching from the root node represents a decision rule that tests the value of a feature, and the leaf nodes represent a classification result. Node selection in the decision tree takes place step by step, starting from the root node of the tree. At each node of the tree, the dataset needs to be split by the value of a particular attribute. This division aims to divide the dataset into homogeneous subgroups and obtain the best classification results. Therefore, it is important to choose a feature that can split the dataset at a node well. To do this, different division criteria such as gini index, information gain, classification error and twoing criteria are used. The structure of a DT model is given in the Figure 5.

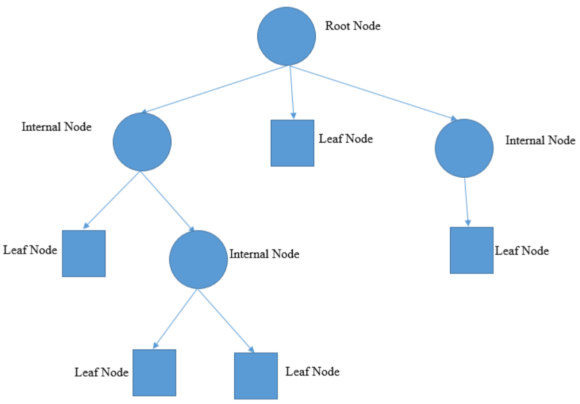

Figure 5. Structure of a DT model

2.3.4. Gaussian Naive Bayes (GNB)

GNB is a Naive Bayes derivative for continuous features that uses Bayes' theorem and the normal distribution (Gaussian distribution). This algorithm is based on the distribution of data to solve classification problems and performs calculations to distinguish classes from each other in feature space. The basic assumption of Gaussian Naive Bayes is that the features have a normal distribution within the classes. Therefore, it is important to calculate the mean and variance of each class for data features. The assumption that the features have a normal distribution works well in most practice, and so Gaussian Naive Bayes is widely used in classification problems.

The following equation is used to determine P(X|Yi), where X (x_1, x_2, x_3,...,xn) and Yi (i=1,2,...,m) represent the features of the data sample and classes, respectively.

$$P(X|Y_i) = \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-(X-\mu)^2}{2\sigma^2}} \quad (2)$$

In Equation 2, μ is the mean , σ is the standard deviation of the i^{th} class.

3. Experimental Setup and Evaluation Metrics

All of the experiments were conducted on MATLAB R2022b 64bit Intel(R) Core(TM) i5-6500 CPU @ 3.20GHz and 8.0 GB of RAM. First, the parameters of the algorithms are defined separately. The impact of various kernel functions on SVM varies greatly. Due to its superior performance, the Gaussian kernel is used in this paper. The trade-off between margins and classification errors is managed with penalty value (C). 1, 10, 100, and 1,000 are the C values that are most frequently used. With the use of grid search, we chose the proper C value. When the C value is 1, the SVM improved the classification accuracy. It is important to maximize the performance of the KNN by paying attention to selection of the K value. By using elbow function, we set K as 3. For, decision tree gini index is used as division criteria and maximum number of splits is selected as 100.

The classification performances of the algorithms are compared for each subject based on precision, recall, accuracy and F1-score obtained from confusion matrix. A confusion matrix given in Figure 6. provides details about the number of correctly and incorrectly classified samples. In the confusion matrix, true positive (TP) shows the number of correctly classified hand closed position, false positive (FP) represents the number of hand open position classified as hand closed position. False negative (FN) is the number of hand closed position classified as hand open position, True negative (TN) represents the number of correctly classified hand open position. In terms of TP, FP, FN, and TN, the formulas of performance metrics are given below.

ACTUAL

PREDICTED	Hand Closed Position 0	Hand Open Position 1
Hand Closed Position 0	TP	FP
Hand Open Position 1	FP	TN

Figure 6. Confusion Matrix

$$Acuuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (3)$$

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

$$F - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (6)$$

The Receiver Operating Characteristic (ROC) curve is a performance metric used to assess classification tasks at various threshold levels. While the ROC is a probability curve, AUC is the area under the curve. It demonstrates how well the model can differentiate across classes. The greater the AUC, the more accurately the model predicts class 0 as 0 and class 1 as 1 and the better the model distinguishes between hand closed position and hand open position. The ROC curve is plotted as True Positive Rate (TPR) versus False Positive Rate (FPR), with TPR on the y-axis and FPR on the x-axis.

Cross-validation (CV) is used to assess model performance on unseen data that was not used during training. The basic principle of CV is to exclude some of the data, build a model from the rest, and then estimate the excluded samples. The data is divided into k fold and the excluded samples are considered as the test sample in each iteration, for a total of k iterations. At the end, the average of the k performances obtained is accepted as the success of the k-fold CV. In this study, we applied 5-fold CV as shown in Figure 7.

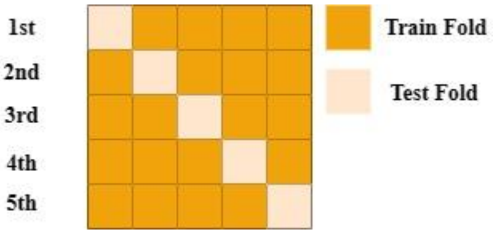

Figure 7. 5-fold cross validation

4. RESULTS AND DISCUSSION

In our study, the classification of EMG signals recorded by hand closed position and hand open position the human palm was carried out. Records from a total of 4 subjects were evaluated separately by performing 5-fold cross validation over SVM, KNN, DT and GNB classifiers. The average F1-score is calculated over all 5-fold for each subject and each algorithm. The results are given in Table 2.

Table 2. Classification results in terms of average F1-scores for each subject and each algorithm

	Algorithm			
Subject	SVM	KNN	DT	GNB
S1	**62.4**	55.9	61.9	60.7
S2	**68.3**	63.0	67.7	68.1
S3	**68.3**	63.4	67.7	68.1
S4	**73.1**	67.5	72.7	71.5

When examining the algorithms as a whole, it becomes clear that Subject4 consistently achieved higher F1-scores across all scenarios. On the other hand, Subjects 1, 2, and 3 exhibited more variability in performance among different algorithms. This disparity implies that certain algorithms may be more suitable for specific subjects due to variations in muscle activity patterns, noise levels, or other individual characteristics.

The results indicate that classification with SVM always achieves better performance than others in terms of mean F1-score for each subject in our study. Biomedical signals are not always absolutely reproducible and can sometimes even be contradictory. SVM deals with this by mapping the data points into high-dimensional space. So, SVM is useful in the processing and classification of biological signals. In addition, all the algorithms show better classification results for subject4, the worst classification results for subject1. This is because EMG data is heavily influenced by individual differences. To improve our understanding, confusion matrices are given in Figures 8-11 for each subject and each algorithm.

From confusion matrices above we discover that it is easier to predict hand open position than hand closed position. Because there is a clear difference in EMG amplitude between the ED muscle and the FCR muscle when the hand is open. But when the hand is closed, the amplitude difference between the two muscles decreases. So it may be easier to predict the open position of the hand.

In Figures 12-15 the estimated ROC curves and AUC values are given for each subject and for each algorithm. The area under the curve varies between 0.58 and 0.66 for subject1, 0.69 and 0.75 for subject2, 0.67 and 0.75 for subject3 and 0.72 and 0.79 for subject4. It has been seen that the results obtained with the ROC curve are compatible with the F1-scores and confusion matrices. According to the results we can say that algorithms show better statistical quality for subject4.

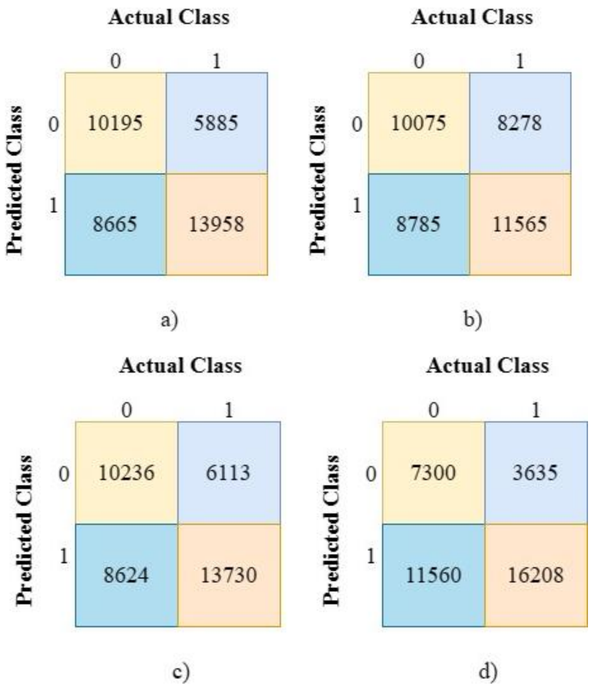

Figure 8. Confusion matrices for subject1 a) SVM b) KNN c) DT d) GNB

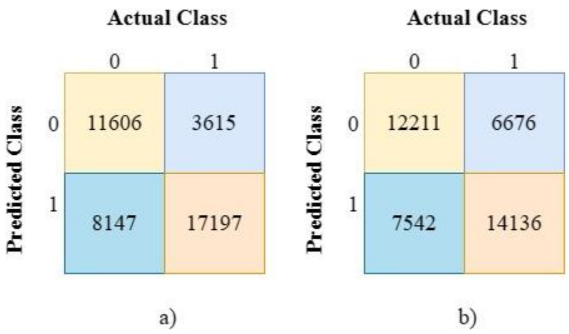

Figure 9. Confusion matrices for subject2 a) SVM b) KNN c) DT d) GNB

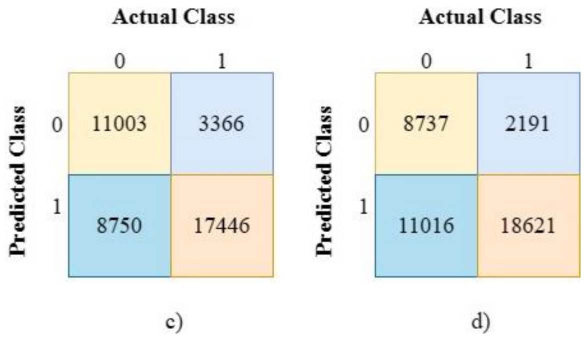

Figure 9. Contunie

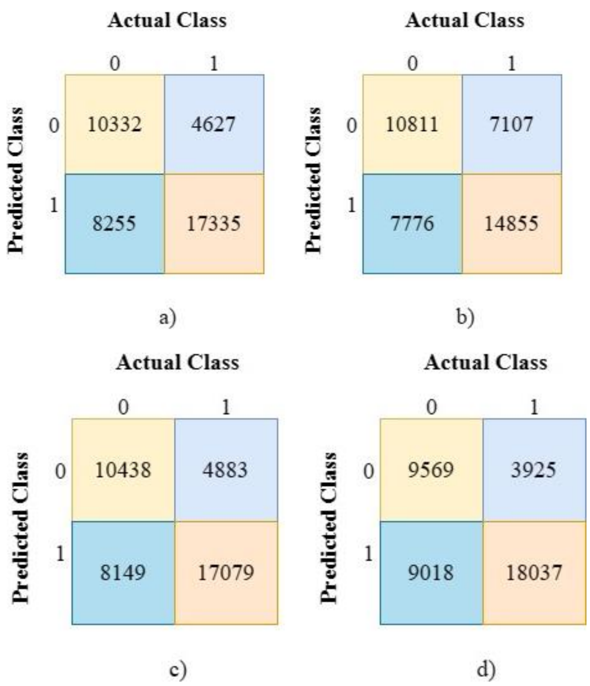

Figure 10. Confusion matrices for subject3 a) SVM b) KNN c) DT d) GNB

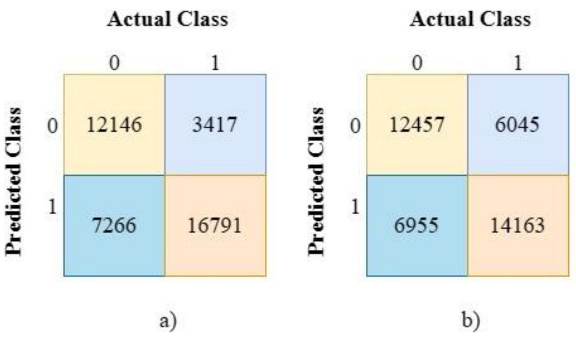

Figure 11. Confusion matrices for subject4 a) SVM b) KNN c) DT d) GNB

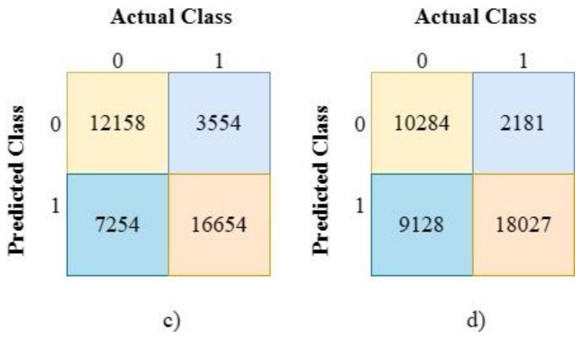

Figure 11. Continue

The results provide valuable insights into the accuracy of predicting hand positions based on EMG signals recorded from the ED and FCR muscles. The variations in EMG amplitude between these two muscles during different hand positions contribute to the differences in prediction difficulty observed. Notably, our results show that predicting a hand open position is generally easier due to distinct EMG amplitude variations compared to when the hand is closed. This difference leads to decreased prediction accuracy for the closed position. These findings support previous research indicating that muscle activation patterns during different hand positions can impact the discriminative ability of EMG signals [26].

Moreover, the diverse statistical characteristics found among various subjects present interesting implications. Subject4 consistently demonstrated superior outcomes across multiple measures, indicating that the algorithms possess stronger predictive abilities for this specific subject. This occurrence may be attributed to individual-specific muscle activation patterns, unique EMG signal attributes, or other physiological factors influencing the classification procedure. These results highlight the significance of customizing machine learning methods for each subject individually in order to potentially enhance accuracy and dependability of hand movement classification systems based on EMG data.

One limitation of the study is that EMG measurements were conducted on a limited sample size of four individuals, which restricts the dataset. However, to compensate for this limitation, additional features were introduced to enhance the comprehensiveness of the dataset. Furthermore, since the classification process already incorporates personalized techniques, the small number of participants does not hinder the success of classification.

In related literature, there have been studies on classifying EMG data using deep learning algorithms [7, 8]. These studies also utilized a small number of participants [7], indicating that our study's dataset can be effectively employed in deep learning methods as well.

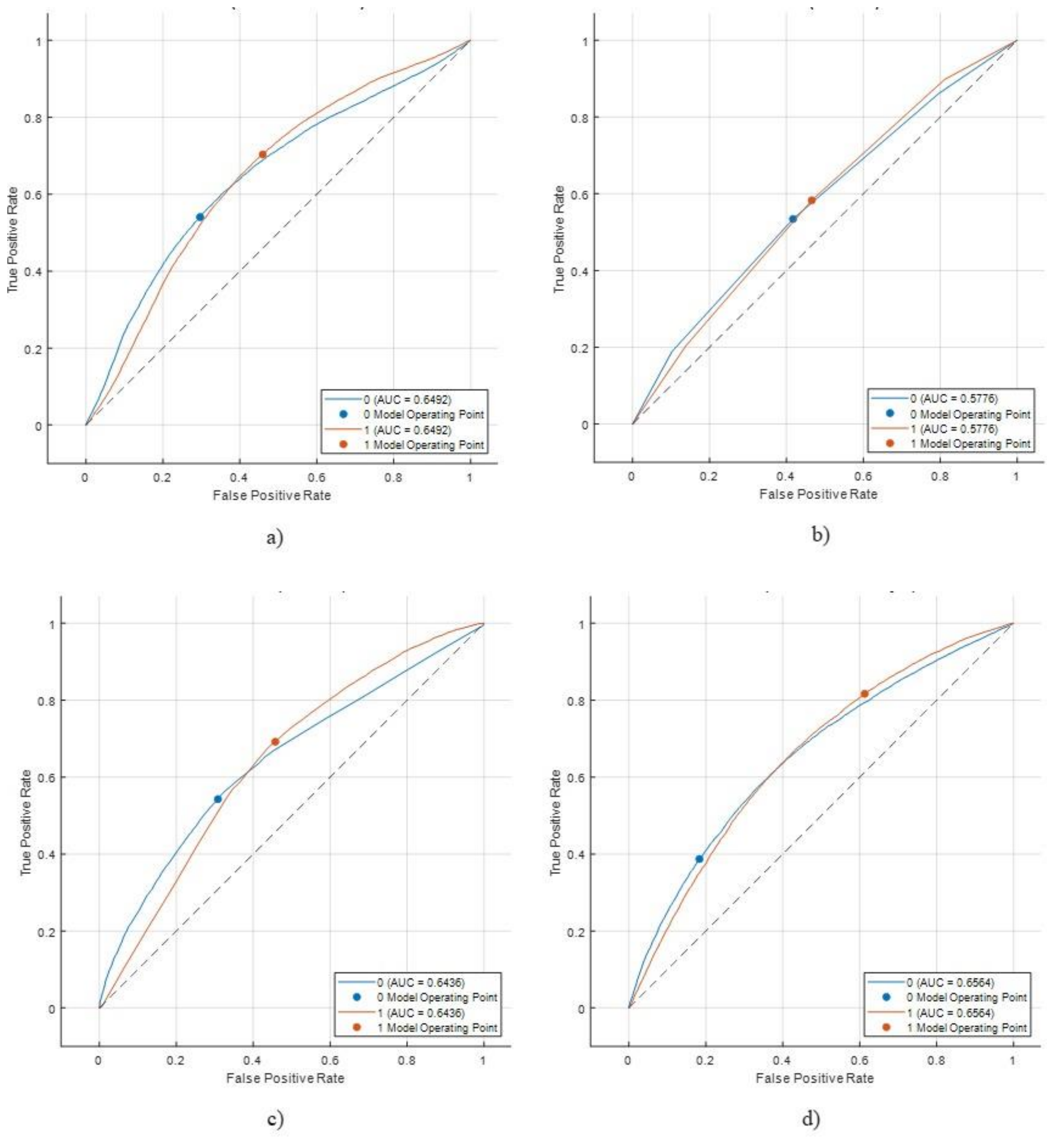

Figure 12. ROC graphs for subject1 a) SVM b) KNN c) DT d) GNB

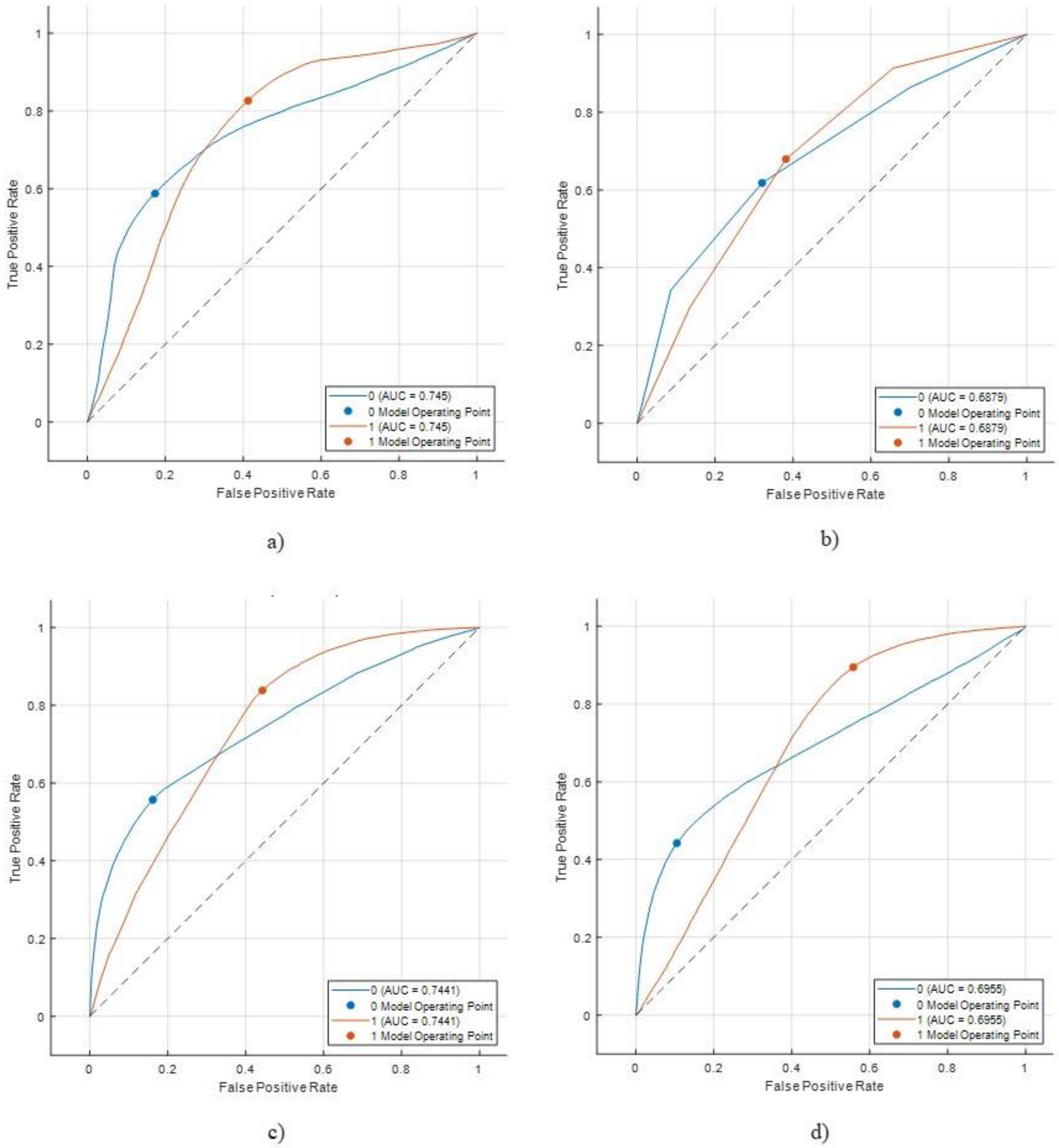

Figure 13. ROC graphs for subject2 a) SVM b) KNN c) DT d) GNB

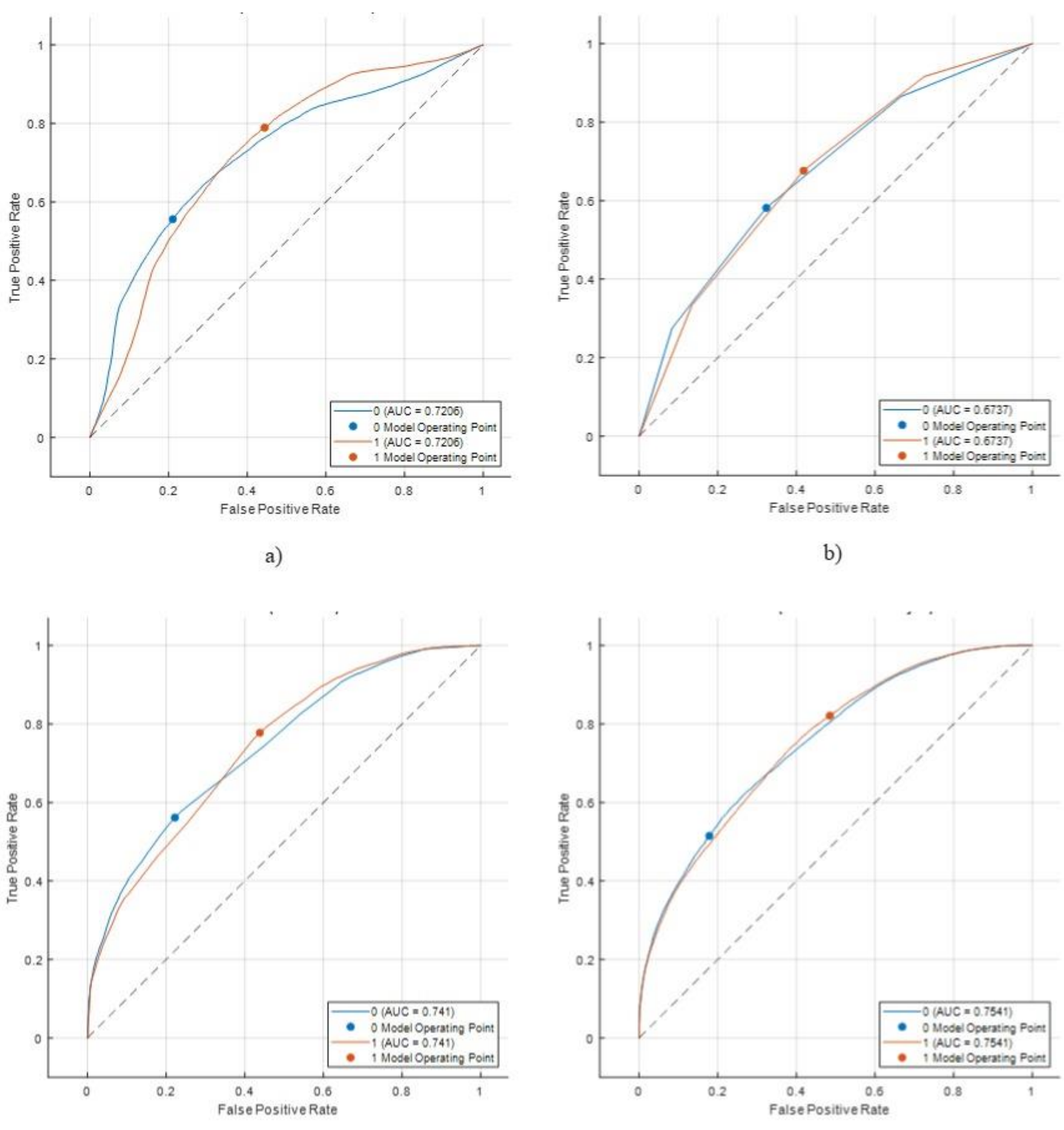

Figure 14. ROC graphs for subject3 a) SVM b) KNN c) DT d) GNB

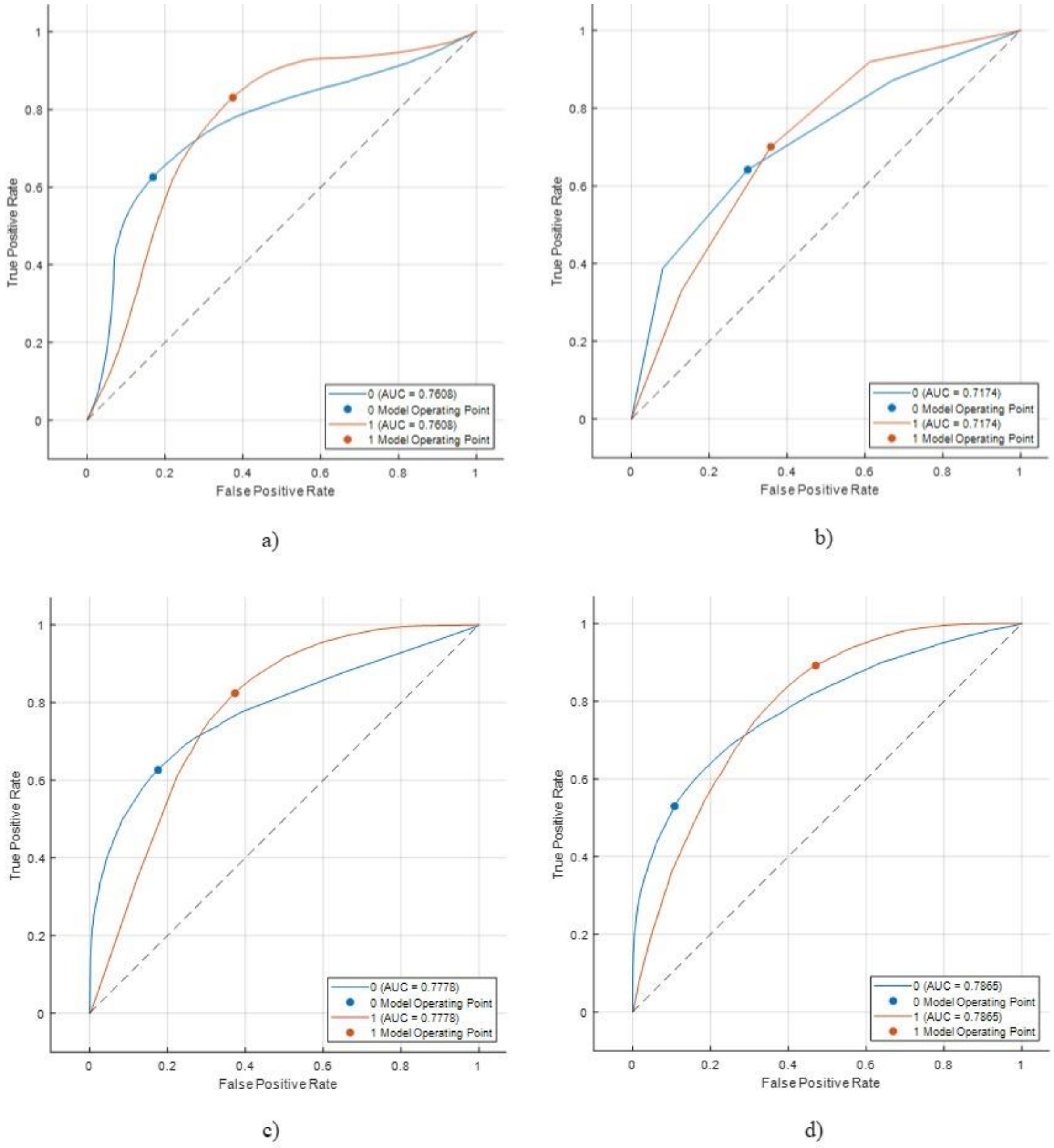

Figure 15. ROC graphs for subject4 a) SVM b) KNN c) DT d) GNB

5. CONCLUSIONS

This study mainly focused on the classification of hand open and hand closed position based on EMG data. In accordance with this purpose, we comparatively examined the effect of ML-based classification algorithms on EMG classification performance. A data set of EMG signals recorded from ED and FCR muscles of dominant hand of four healthy subjects. Feature engineering were realized and from original datasets time-domain features namely var, std, rms, ae, min and max were extracted to strengthen the classification performances. Experiments were performed under the same conditions for each subject. The SVM outperformed KNN, DT, and GNB in a variety of subjects which was proven through experimental results. The main findings of this study show that suggested algorithms are common and can be used for several biomedical classification tasks when there are datasets obtained from various subjects.

In conclusion, our research highlights the influence of hand positioning on the challenge of predicting movements and emphasizes the importance of individual characteristics in determining how well algorithms perform. These results have significant implications for personalized applications that rely on EMG technology, such as prosthetics and rehabilitation, where accurate classification of hand movements is critical.

Future work will be devoted to exploring datasets with more channels and optimizing feature extraction and selection in this context.

DECLARATION OF ETHICAL STANDARDS

The authors of this article declare that the materials and methods used in this study do not require ethical committee permission and/or legal-special permission.

AUTHORS' CONTRIBUTIONS

Ekin EKİNCİ: The development of a research idea, analysis of experimental data, implementation of ML algorithms, interpretation of findings, manuscript writing.

Zeynep GARİP: The development of a research idea, analysis of experimental data, implementation of ML algorithms, interpretation of findings, manuscript writing.

Kasım SERBEST: Corresponding of the study, the development of a research idea, collection and analysis of experimental data, interpretation of findings, manuscript writing, and general revision.

CONFLICT OF INTEREST

There is no conflict of interest in this study.

REFERENCES

[1] Kang S., Kim H., Park C., Sim Y., Lee S. and Jung Y., "sEMG-Based Hand Gesture Recognition Using Binarized Neural Network", ***Sensors***, 23(3): 1436, (2023).

[2] Jarosz J., Gołaś A., Krzysztofik M., Matykiewicz P., Strońska K., Zając A. and Maszczyk A., "Changes in muscle pattern activity during the asymmetric flat bench press (offset training)", ***International Journal of Environmental Research and Public Health***, 17(11): 3912, (2020).

[3] Borysiuk Z., Blaszczyszyn M., Piechota K., Konieczny M. and Cynarski W.J., "Correlations between the EMG Structure of Movement Patterns and Activity of Postural Muscles in Able-Bodied and Wheelchair Fencers", ***Sensors***, 23(1): 135, (2022).

[4] Xie Z., "Fatigue Monitoring and Recognition During Basketball Sports via Physiological Signal Analysis", ***International Journal of Information System Modeling and Design***, 13(2): 1-11, (2022).

[5] Phillips D.A., Del Vecchio A.R., Carroll K. and Matthews E.L., "Developing a Practical Application of the Isometric Squat and Surface Electromyography", ***Biomechanics***, 1(1): 145-151, (2021).

[6] Jochumsen M., Niazi I.K., Zia ur Rehman M., Amjad I., Shafique M., Gilani S.O., and Waris A., "Decoding attempted hand movements in stroke patients using surface electromyography", ***Sensors***, 20(23): 6763, (2020).

[7] Fu Y.L., Liang K.C., Song W. and Huang J., "A hybrid approach to product prototype usability testing based on surface EMG images and convolutional neural network classification", ***Computer Methods and Programs in Biomedicine***, 221: 106870, (2022).

[8] Karnam N.K., Dubey S.R., Turlapaty A.C., and Gokaraju B., "EMGHandNet: A hybrid CNN and Bi-LSTM architecture for hand activity classification using surface EMG signals", ***Biocybernetics and biomedical engineering***, 42(1): 325-340, (2022).

[9] Sahu P., Singh B. K. and Nirala N., "An improved feature selection approach using global best guided Gaussian artificial bee colony for EMG classification", ***Biomedical Signal Processing and Control***, 80: 104399, (2023).

[10] Fajardo J.M., Gomez O. and Prieto F., "EMG hand gesture classification using handcrafted and deep features", ***Biomedical Signal Processing and Control***, 63: 102210, (2021).

[11] Tepe C. and Demir M.C., "Real-Time Classification of EMG Myo Armband Data Using Support Vector Machine" ***IRBM***, 43(4): 300-308, (2022).

[12] Klein Breteler M.D., Simura K.J. and Flanders M., "Timing of muscle activation in a hand movement sequence", ***Cerebral Cortex***, 17(4): 803-815, (2007).

[13] Khushaba R.N, Kodagoda S., Takruri M. and Dissanayake G., "Toward improved control of prosthetic using surface electromyogram (EMG) signals" ***Expert Systems with Applications*** 39:10731-10738, (2012).

[14] Roldan-Vasco S., Orozco-Duque A. and Orozco-Arroyave J.R., "Swallowing disorders analysis using surface EMG biomarkers and classification models", ***Digital Signal Processing***, 133: 103815, (2023).

[15] Torres-Castillo J.R., Lopez-Lopez C.O. and Padilla-Castaneda M.A., "Neuromuscular disorders detection through time-frequency analysis and classification of multi-muscular EMG signals using Hilbert-Huang transform", ***Biomedical Signal Processing and Control***, 71: 103037, (2022).

[16] Putra D.S. and Weru Y.U.W., "Pattern recognition of electromyography (EMG) signal for wrist movement using learning vector quantization (LVQ)", ***In IOP Conference Series: Materials Science and Engineering***, 50681):, p. 012020, (2019).

[17] López L.I.B., Caraguay Á.L.V., Vimos V.H., Zea J.A., Vásconez J.P., Álvarez M. and Benalcázar M.E., "An energy-based method for orientation correction of EMG bracelet sensors in hand gesture recognition systems", ***Sensors***, 20(21): 6327, (2020).

[18] Kısa D.H., Özdemir M.A., Güren O. and Alaybeyoğlu, A., "A decision-making mechanism based on EMG signals and adaptive neural fuzzy inference system (ANFIS) for hand gesture prediction", ***Journal of the Faculty of Engineering and Architecture of Gazi University***, 38(3): 1417-1430, (2023).

[19] Vapnik V.N., "***The nature of statistical learning theory"***, Springer science & business media, (2000).

[20] Yiğit H., Köylü H. and Eken, S., "Estimation of road surface type from brake pressure pulses of ABS", ***Expert Systems with Applications***, 212: 118726, (2023).

[21] Ay Ş., Ekinci E. and Garip Z., "A comparative analysis of meta-heuristic optimization algorithms for feature selection on ML-based classification of heart-related diseases", ***The Journal of Supercomputing***, 1-30, (2023).

[22] Şengür D., "EEG, EMG and ECG based determination of psychosocial risk levels in teachers based on wavelet extreme learning machine autoencoders", ***Politeknik Dergisi***, 25(3): 985-989, (2022).

[23] Cica D., Sredanovic B., Tesic S. and Kramar, D., "Predictive modeling of turning operations under different cooling/lubricating conditions for sustainable manufacturing with machine learning techniques", ***Applied Computing and Informatics***, 1-19, (2020).

[24] Atban F., Ekinci E. and Garip Z., "Traditional machine learning algorithms for breast cancer image classification with optimized deep features", ***Biomedical Signal Processing and Control***, 81: 104534, (2023).

[25] Han J., "System optimization of talent Life cycle management platform based on decision tree model", ***Journal of Mathematics***, 2022: 1-12, (2022).

[26] Latash M.L., "Muscle coactivation: definitions, mechanisms, and functions", ***Journal of neurophysiology***, 120(1): 88-104, (2018).