

Original Article

## HAND GESTURE RECOGNITION FOR EMG SIGNAL BASED ON DIFFERENT MACHINE LEARNING METHODS

Batool Abd Alhade <sup>1\*</sup>

<sup>1</sup> College of Science, Al-Qasim Green University, Babylon, Iraq



### ABSTRACT

Muscle electrical activity is measured by electromyography (EMG) signals, which are typically expressed as frequency, amplitude, and phase as function of time. Applications for biosignals include the diagnosis of neuromuscular disorders, the control of assistive devices, like orthotic or prosthetic devices, the control of machinery, computers, robots, etc. For the purpose of improving hand gesture recognition, the presented study employed EMG signals, band pass filtering, and finite impulse response for removing artifacts as well as low-frequency noise (like baseline drift or body movement) and high-frequency noise (like electrical interference) impairing the performance of the system. In order to more precisely identify hand gestures, efficient methods have been employed for classifying gestures utilizing a variety of machine learning (ML) classifiers (SVM, Random Forest (RF), KNN, and Gradient Boosting).

The classification accuracy using Random Forest (99.9%) reached, accuracy using SVM (97.4%), accuracy using KNN (99.1%). while the classification accuracy using Gradient Boosting (99.7%)

**Keywords:** RF, SVM, KNN, GBM, FIR

### INTRODUCTION

Applications for EMG signals include the diagnosis regarding developing muscle-oriented exercise equipment, neuromuscular disorders, orthotic and prosthetic device control, human-machine interfaces, VR games, and more [Kumari and Ali \(2015\)](#). Recently, myoelectric control was utilized extensively, with hand-posture recognition depending upon EMG technology being the most common. The more complex and larger motions of wrist, hand, and arm—like forearm supination, forearm pronation, wrist flexion, wrist extension, wrist ulnar deviation, wrist radial deviation, and wrist external and internal rotation—have been the subject of the majority of earlier research [Shi et al. \(2018\)](#). The demand for more accurate and natural human-computer interaction systems has led to significant advancements in hand gesture detection systems during the past 20 years. Additionally, upper limb prosthesis is one where Human Computer Interfaces (HCI) are increasingly significant. Since the hands are among the most vital and useful parts of bodies, losing them could significantly lower a person's quality of life. This is the primary cause of the abundance of research in literature looking for the best methods to operate upper limb active prostheses. Despite the volume of publications, there are still a few significant obstacles when it comes to actual HCI methods for controlling upper limb prostheses [Toro-Ossaba et al. \(2022\)](#).

#### \*Corresponding Author:

Email address: Batool Abd Alhade ([Batool@uoqasim.edu.iq](mailto:Batool@uoqasim.edu.iq))

Received: 22 February 2026; Accepted: 15 April 2026; Published 09 May 2026

DOI: [10.29121/DigiSecForensics.v3.i1.2026.99](https://doi.org/10.29121/DigiSecForensics.v3.i1.2026.99)

Page Number: 27-35

Journal Title: Journal of Digital Security and Forensics

Journal Abbreviation: J. Digi. Sec. Forensics

Online ISSN: 3048-894X

Publisher: Granthaalayah Publications and Printers, India

Conflict of Interests: The authors declare that they have no competing interests.

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Authors' Contributions: Each author made an equal contribution to the conception and design of the study. All authors have reviewed and approved the final version of the manuscript for publication.

Transparency: The authors affirm that this manuscript presents an honest, accurate, and transparent account of the study. All essential aspects have been included, and any deviations from the original study plan have been clearly explained. The writing process strictly adhered to established ethical standards.

Copyright: © 2026 The Author(s). This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

With the license CC-BY, authors retain the copyright, allowing anyone to download, reuse, re-print, modify, distribute, and/or copy their contribution. The work must be properly attributed to its author.

A two-channel sEMG-based system was created in [Shi et al. \(2018\)](#) for controlling a bionic hand as well as identifying human hand postures. Four time-domain features, which are (ZC, MAV, SSC, and WL) have been derived from sEMG signals obtained from flexor digitorum superficialis as well as extensor digitorum muscles. Using K-nearest neighbors (KNN) classifier, a total of four distinct hand postures were recognized. Custom-built bionic hand's servo motors have been driven by Arduino controller that received classification outputs. The bionic hand effectively replicated the desired hand postures, and the experimental results demonstrated high online accuracy of 94%.

In [Zhang et al. \(2011\)](#), the researchers presented a model for hand-gesture recognition, combining information from multichannel EMG sensors as well as three-axis accelerometer. EMG intensity is used for automatically identifying gesture segments, and multistream hidden Markov models and a decision tree (DT) are used for obtaining the final classification. Experiments on 40 continuous sentences as well as 72 Chinese Sign Language words demonstrate the technique's efficacy as well as the complementary strengths regarding EMG and ACC signals. For showing gesture-based control, an 18-gesture virtual Rubik's cube game is created in real time. Tests that are both user-independent and user-specific verify that the framework facilitates natural and intelligent HCI.

In [Qi et al. \(2020\)](#), with the use of principle component analysis (PCA) as well as GRNN neural network, researchers created gesture recognition system that improves detection accuracy and efficiency by reducing dimensionality and redundancy in EMG signals. Four signal characteristics have been taken from arm sEMG data and applied to a total of 9 static gestures as samples. Following neural network training as well as dimension reduction, the system's overall recognition rate was 95.1%, with average recognition time of 0.19 seconds.

In [López et al. \(2024\)](#), Through comparing CNN-LSTM and CNN models and using post-processing technique for filtering out spurious predictions, the researchers were able to enhance hand gesture recognition (HGR) from EMG signals. With EMG-EPN612 dataset, which included 5 gestures from 612 people, post-processing improved CNN-LSTM and CNN accuracy by 24.77% and 41.86%, respectively. Accuracy increased by 3.29% with the addition of memory cells (CNN-LSTM), yet 53 times more parameters were needed. CNN-LSTM with post-processing produced a 90.55% mean recognition accuracy. The findings demonstrate the advantages regarding memory as well as post-processing in HGR and offer avenues for further investigation.

In [Aarotale and Rattani \(2024\)](#), the authors combined ML and DL models with fused time-domain, temporal-spatial, and wavelet-based features to establish benchmarks for novel feature extraction techniques. With the use of fused time-domain descriptors, 1D Dilated CNN obtained 97% accuracy on Grabmyo dataset. RF used temporal-spatial descriptors for obtaining 94.95% accuracy on FORS-EMG dataset.

The remaining parts of the study are arranged as follows: In the following section, we discuss the key topic of the presented study. In part (3), the suggested model is discussed. Experimental results are shown in Part (4). Lastly, part (5) contains the paper's conclusion

## MATERIALS AND METHODS

### SIGNAL ACQUISITION

MYO Thalmic bracelet, as depicted in [Figure 1](#), was worn on volunteer's forearm for recording patterns, and a computer that has a Bluetooth receiver. The bracelet's eight sensors, which are evenly spaced throughout the forearm, were designed to simultaneously record electromyography signals. The computer received the signals via Bluetooth

**Figure 1**



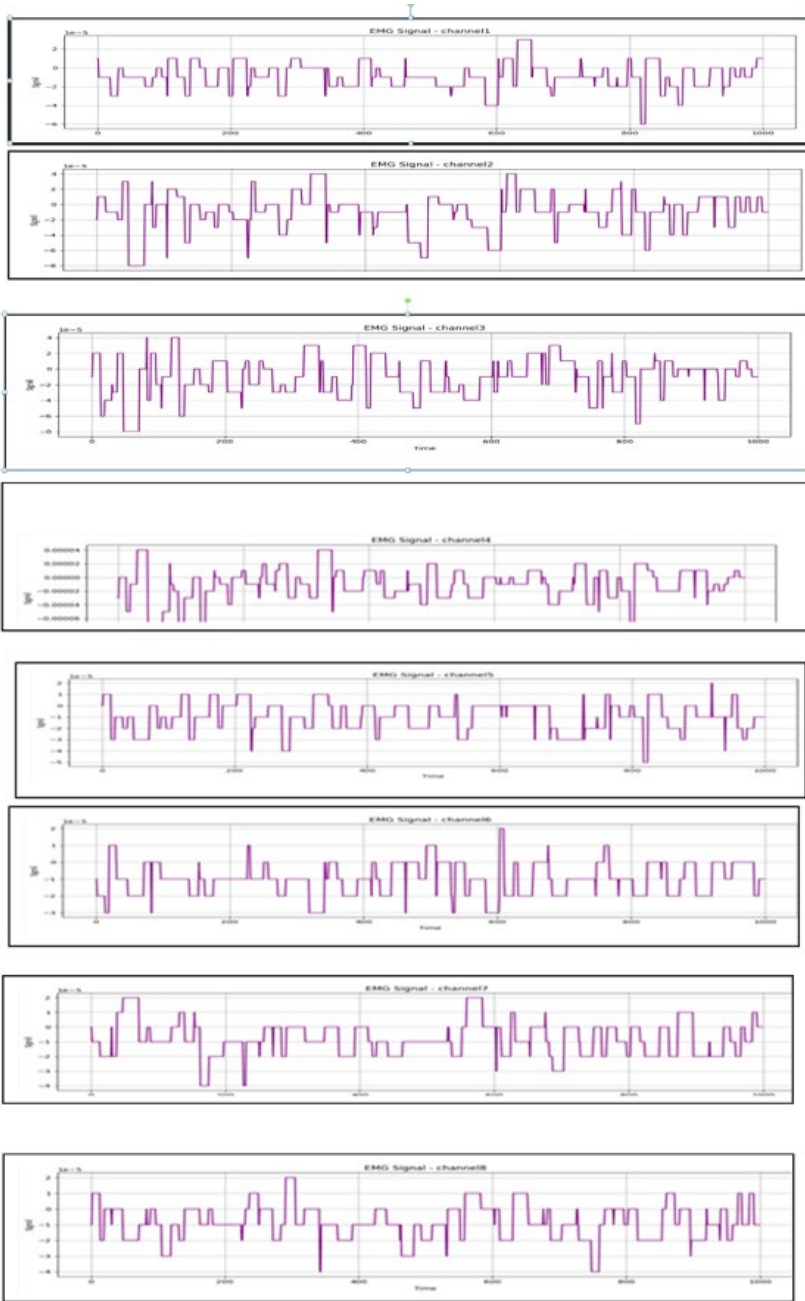
**Figure 1 Location of MYO Armband on the Forearm [Chung and Benalcázar \(2019\)](#)**

### RAW DATA

We recorded EMG signals from 36 distinct users using MYO Thalmic. They made a number of static hand gestures. Every individual made two sets of six or seven gestures, each lasting three seconds and interspersed with three-second pause. There are ten columns in each file. Whereas the first column shows time in ms, columns 2-9 show MYO Thalmic bracelet's EMG channels, and

column ten shows the row with gesture label, which is in the following sequence: (1) The unmarked data, (2) hand at rest, (3) hand clenched in a fist, (4) wrist flexion, (5), wrist extension, (6) the radial deviations, (7) the ulnar deviations, and (8) the expanded palm. Additionally, the individual who carried out the experiment is listed in column (11) "Label." Each of the 36 participants did seven gestures twice. In other words, there are eleven columns in the file. This is a readme file from the data-set itself. <https://archive.ics.uci.edu/ml/datasets/EMG+data+for+gestures>

**Figure 2**



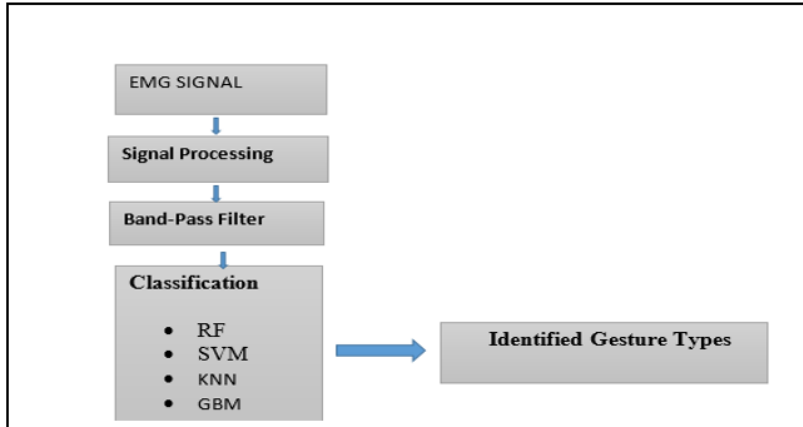
**Figure 2 EMG Signal**

**PROPOSED SYSTEM**

The suggested system consists of 2 stages: the first is signal processing using a finite impulse response (Band-Pass Filter) technique to focus only on important and relevant data and exclude unimportant data. The second is the use of different classifiers

(RF, SVM, KNN, GBM) and comparison of performance accuracy to classify smart electrical signal patterns and identify gesture types. All steps were implemented using Python. Figure 3 shows the suggested system

**Figure 3**



**Figure 3 The Proposed System**

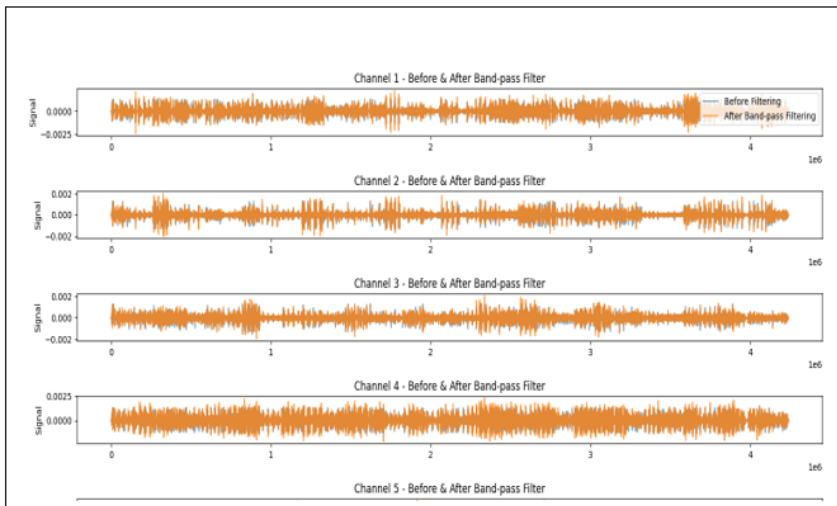
### SIGNAL PROCESSING

The extracted EMG features (represented by Filtered) underwent preprocessing using standardization prior to classifier training. In particular, each feature has been scaled to have zero mean as well as unit variance using the Standard Scaler. For algorithms like SVM and KNN, which are sensitive to feature magnitudes, this step guarantees that each feature contributes equally to the process of classification.

#### 1) Band-pass filter

A band-pass filter was applied to retain frequencies in the range of [typically 20–450 Hz], which correspond to the main EMG signal components, while removing low-frequency motion artifacts and high-frequency noise.

**Figure 4**



**Figure 4 Before and After Band-Pass Filter**

### SIGNAL CLASSIFICATION

Four commonly used classifiers for biosignal analysis were employed:

## RANDOM FOREST (RF)

Since it has first been introduced by L. Breiman in 2001, RF showed to be a quite effective general-purpose classification as well as regression method [Biau and Scornet \(2016\)](#). It attracts scholars from variety of backgrounds because of its inherent interdisciplinary nature [Akar and Güngör \(2012\)](#). An ensemble of DTs that efficiently manages feature interactions and noise. Each one of the tree classifiers in RF classifier casts a unit vote for most popular class for classifying an input vector, and every one of the classifiers is created with the use of random vector selected separately from input vector [Al Sayaydeha and Mohammad \(2019\)](#). The value utilized commonly as square root of total number of variables is considered rather robust to the process of number selection in RF. Also, RF trees are constructed without pruning. According to the value of the threshold that the user had selected or by majority of votes from classification trees in forest, the identification of predictions for test samples will be carried out. It has shown sufficient performance when utilized for tasks of pattern recognition [Ali et al. \(2022\)](#)

Figure 5

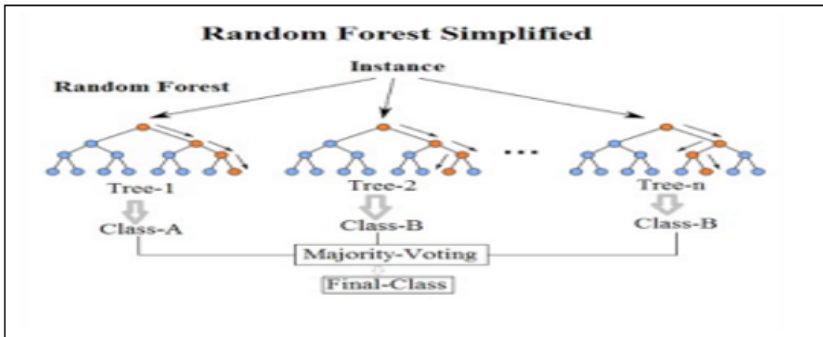


Figure 5 Typical Random Forest Classification [Abdulla et al. \(2023\)](#)

## SUPPORT VECTOR MACHINE (SVM)

At AT & T Bell Laboratories, Vladimir Vapnik and his colleagues invented SVM. It is recognized as one of the most common algorithms in supervised ML and is distinguished from other well-known DM techniques by its reliability, strength, and high accuracy. Put simply, each example in given set of training samples is assigned into one of 2 groups. [Obeas et al. \(2024\)](#). SVMs are classified with the use of hyperplane or set of hyperplanes that, with the help of support vectors, separate data points. The idea regarding a maximum margin hyperplane is illustrated in [Figure 6](#), which divides positive examples (green squares) from the negative examples (red circles); the red circles as well as darker green squares stand in for related support vectors. The hyperplane with the largest functional margin achieves a total separation from the nearest training data point for any class.

Figure 6

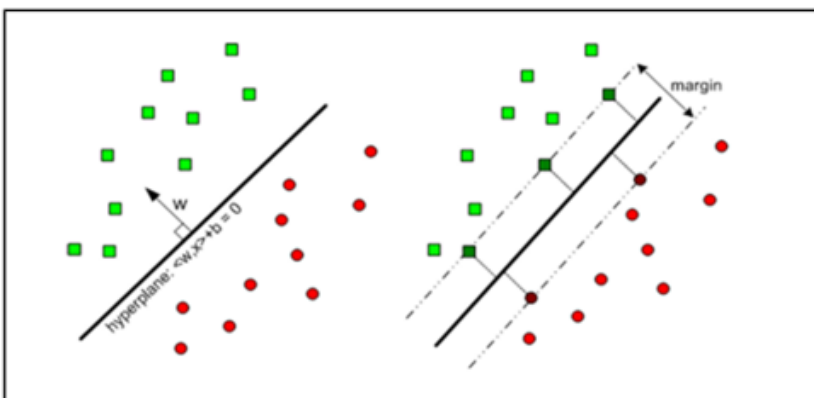


Figure 6 Maximum Margin

## 2) K-NEAREST NEIGHBORS (KNN)

K-means clustering, usually referred to as "Forgy's algorithm," is the most popular as well as majorly utilized data-segmented clustering technique. Its primary goal is processing a large amount of high-dimensional data for identifying representative data. Cluster centers are another name for such representative data. Large volumes of data could be compressed and classified using

such cluster centers. In the case when employing K-means clustering, the number of clusters must be established, and following repeated iterative computing, the errors in each cluster must be steadily reduced until they don't change and converge to final results of clustering. Figure 7. demonstrates how K-means algorithm works. Prior to calculating every data point's distance from cluster center as well as allocating it to the closest cluster center, it calculates cluster number K and establishes cluster center based on K's value. Following distribution, a new center of the cluster is calculated for distribution to the point where the clustering process is finished and distance from the new cluster center and the data meets the end condition.

Figure 7

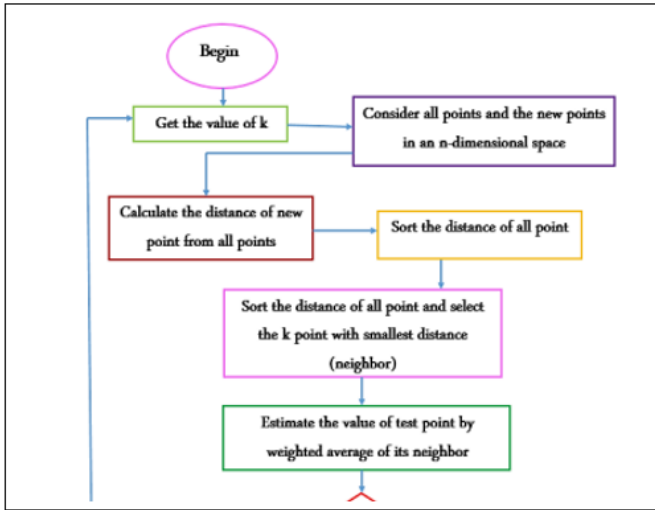


Figure 7 Flowchart of KNN Algorithm Ahmed et al. (2022).

### 3) Gradient Boosting (GB)

An advanced ensemble algorithm which increases accuracy by building weak learners sequentially. A family of potent ML approaches that are referred to as gradient boosting machines has shown notable performance in various practical scenarios Natekin and Knoll (2013). GBMs are the names given to gradient-descent based formulation of boosting techniques as well as the associated models. GBMs build base learners iteratively through reweighting misclassified observations, just as AdaBoost. In contrast to AdaBoost, GBMs use negative partial derivatives regarding loss function at every one of the training observations to calculate weights. The partial derivatives are referred to as pseudo-residuals, and they're utilized for the iterative expansion of an ensemble. Which is why, feature space is divided into groups based on related pseudo-residuals. Scalable variants of GBMs are required for much bigger datasets, even though they could be effective for fairly small datasets. To meet such need, tree-based scalable GBMs called LightGBM, XGBoost, and CatBoost were created lately. We refer to the scalable variants of GBMs as gradient boosted decision classifiers (GBDCs) in order to differentiate them from original GBMs that use DTs as base learners. LightGBM, XGBoost, and CatBoost are scalable GBDT systems that we compare performance metrics with using GBDCs as a baseline in the presented study Dev and Eden (2019).

## TRAINING AND EVALUATIONS

Every one of the models has been trained on the training set then evaluated on testing set.

Performance metrics included Vujović (2021), Najjar et al. (2025):

- **Accuracy**

Accuracy calculation involves the division of total number of data-sets (P + N) by summation of two accurate predictions (TP + TN). 1.0 represents maximum accuracy and 0.00 is minimum.

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

- **Precision**

Precision is calculated by dividing total number of positive prediction cases (TP + FP) by number of true positive prediction cases (TP). 1.0 is the optimal level of accuracy while 0.0 represents the worst.

$$\text{Precision} = \frac{TP}{TP+FP} \tag{2}$$

- **Recall**

Recall measures the classifier’s ability to find all positive instances. It is also known as sensitivity or completeness, showing how many actual positive cases were identified correctly.

$$\text{Recall} = \frac{TP}{TP+FN} \tag{3}$$

- **F1-score**

F-score, which is also referred to as F-Measure is a test’s accuracy measure. It is calculated, based on precision and reminders, by the formula

$$F\_Score = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \tag{4}$$

## RESULTS AND DISCUSSION

This section discusses the proposed methodological results, as outlined below. The data is processed and classified using the proposed machine learning classifiers, and the results from both are compared.

**Table 1**

Table 1 Shows the Classification Performance of Four Machine Learning						
Signal Processing	Model	Accuracy	Precision	Recall	F1-Score	
	RF	99.9%	100%	100%	100%	
EMG SIGNAL	SVM	97.4%	98%	97%	97%	
	BPF	KNN	99.1%	99%	99%	99%
		GBM	99.7%	100%	100%	100%

Table 1. Figure 8 shows the classification performance of four machine learning algorithm models applied in the electromyography (EMG) recognition task. The RF and GBM algorithms achieved the highest overall accuracy, at 99.97% and 99.79%, respectively. Both models also achieved perfect values in precision, recall, and F1-score metrics. This demonstrates their high ability to distinguish between different categories of muscle movements with minimal errors in classification.

The K-Nearest Neighbors (KNN) model also showed outstanding performance, achieving 99.11% accuracy with balanced values for both accuracy and recall (0.99%), indicating its high efficiency in classifying EMG signal patterns.

In contrast, the SVM algorithm achieved a relatively lower accuracy of 97.49%, but still maintained high performance. This reflects its effectiveness in dealing with linearly separable data, but its relatively limited ability to represent complex nonlinear relationships compared to ensemble algorithms (Ensemble Methods).

Overall, these results demonstrate that cluster learning models (RF and GBM) outperform traditional algorithms in EMG signal classification tasks, due to their ability to integrate multiple decision trees and effectively reduce variance and overfitting. These results confirm the suitability of these models for bio signal analysis and gesture recognition applications requiring high accuracy and reliability.

Figure 8

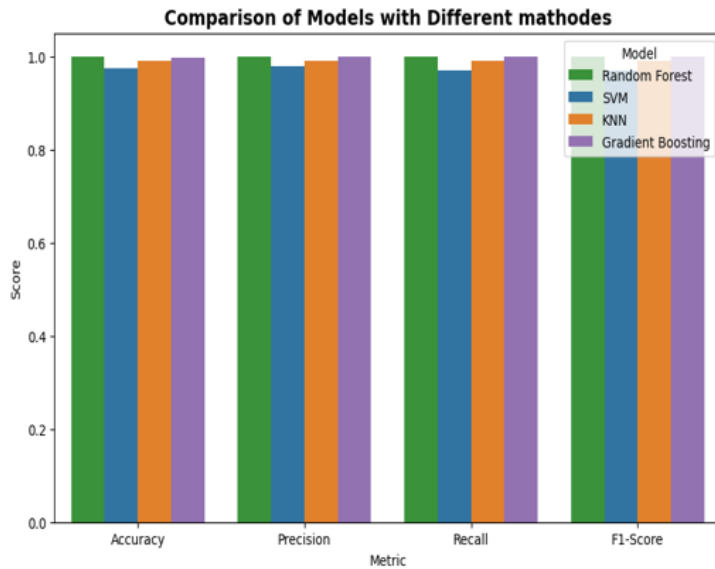


Figure 8 Comparison Between Four Models of Machine Learning Algorithms

## CONCLUSION

In this study, electromyography (EMG) signals were processed and analyzed to recognize hand gestures with high reliability. An FIR band-pass filter was applied to remove motion artifacts, baseline drift, and electrical noise, resulting in cleaner and more representative EMG features. Several machine learning classifiers were evaluated, including RF, SVM, KNN, and Gradient Boosting. The results showed excellent performance across all models, with Random Forest achieving the highest accuracy of **99.9%**, followed by Gradient Boosting (**99.7%**), KNN (**99.1%**), and SVM (**97.4%**). These findings confirm that EMG signals, when properly filtered and processed, can be used effectively for gesture recognition tasks. Additionally, the strong performance of ensemble models highlights their suitability for

## ACKNOWLEDGMENTS

None.

## REFERENCES

- Aarotale, P. N., and Rattani, A. (2024). Machine learning-based SEMG signal classification for hand gesture recognition. in 2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) (6319–6326). IEEE. <https://doi.org/10.1109/BIBM62325.2024.10822133>
- Abdulla, A., Baryannis, G., and Badi, I. (2023). An Integrated Machine Learning and MARCOS Methods for Supplier Evaluation and Selection. *Decision Analytics Journal*, 9, 100342. <https://doi.org/10.1016/j.dajour.2023.100342>
- Ahmed, A. S., Obeas, Z. K., Alhade, B. A., and Jaleel, R. A. (2022). Improving Prediction of Plant Disease Using k-Efficient Clustering and Classification Algorithms. *IAES International Journal of Artificial Intelligence*, 11 (3), 939–948. <https://doi.org/10.11591/ijai.v11.i3.pp939-948>
- Akar, Ö., and Güngör, O. (2012). Classification of Multispectral Images Using Random Forest Algorithm. *Journal of Geodesy and Geoinformation*, 1 (2), 105–112. <https://doi.org/10.9733/jgg.241212.1>
- Al Sayaydeha, O. N., and Mohammad, M. F. (2019). Diagnosis of Parkinson Disease Using Enhanced Fuzzy Min-Max Neural Network and OneR Attribute Evaluation Method. In 2019 International Conference on Advanced Science and Engineering (ICOASE) (64–69). IEEE. <https://doi.org/10.1109/ICOASE.2019.8723870>
- Ali, A. J. M., Hasan, T. M., and Mohammed, S. D. (2022). Digital Modulation Classification Based on Chicken Swarm Optimization and Random Forest. *Journal of Engineering Science and Technology*, 17, 2095–2103.
- Biau, G., and Scornet, E. (2016). A Random Forest Guided Tour. *TEST*, 25(2), 197–227. <https://doi.org/10.1007/s11749-016-0481-7>
- Chung, E. A., and Benalcázar, M. E. (2019). Real-Time Hand Gesture Recognition Model Using Deep Learning Techniques and EMG Signals. in 2019 27th European Signal Processing Conference (EUSIPCO) (1–5). IEEE. <https://doi.org/10.23919/EUSIPCO.2019.8903136>

- Dev, V. A., and Eden, M. R. (2019). Formation lithology Classification Using Scalable Gradient Boosted Decision Trees. *Computers and Chemical Engineering*, 128, 392–404. <https://doi.org/10.1016/j.compchemeng.2019.06.001>
- Kumari, M., and Ali, I. (2015). An Efficient Algorithm For Gender Detection Using Voice Samples. In *2015 Communication, Control and Intelligent Systems (CCIS)* (221–226). IEEE. <https://doi.org/10.1109/CCIntelS.2015.7437912>
- López, L. I. B., Ferri, F. M., Zea, J., Caraguay, Á. L. V., and Benalcázar, M. E. (2024). CNN-LSTM and post-processing for EMG-based hand gesture recognition. *Intelligent Systems with Applications*, 22, 200352. <https://doi.org/10.1016/j.iswa.2024.200352>
- Najjar, E., Alkhaykane, N. A., and Breesam, A. M. (2025). Classification of Arabic Documents with Five Classifier Models Using Machine Learning. *JOIV: International Journal on Informatics Visualization*, 9(1), 365–370. <https://doi.org/10.62527/joiv.9.1.2539>
- Natekin, A., and Knoll, A. (2013). Gradient Boosting Machines, a Tutorial. *Frontiers in Neurorobotics*, 7, 21. <https://doi.org/10.3389/fnbot.2013.00021>
- Obeas, Z. K., Najjar, E., Obaid, A. J., and Rasappan, S. (2024). Estimate Human Emotion Using Machine Learning Based on Voice Signals. In *International Conference on Emerging Trends in AI and Computational Technologies* (pp. 398–410). Springer. [https://doi.org/10.1007/978-3-031-89960-7\\_30](https://doi.org/10.1007/978-3-031-89960-7_30)
- Qi, J., Jiang, G., Li, G., Sun, Y., and Tao, B. (2020). Surface EMG Hand Gesture Recognition System Based on PCA and GRNN. *Neural Computing and Applications*, 32(10), 6343–6351. <https://doi.org/10.1007/s00521-019-04142-8>
- Shi, W.-T., Lyu, Z.-J., Tang, S.-T., Chia, T.-L., and Yang, C.-Y. (2018). A Bionic Hand Controlled by Hand Gesture Recognition Based on surface EMG signals: A preliminary study. *Biocybernetics and Biomedical Engineering*, 38(1), 126–135. <https://doi.org/10.1016/j.bbe.2017.11.001>
- Toro-Ossaba, A., Jaramillo-Tigeros, J., Tejada, J. C., Peña, A., López-González, A., and Castanho, R. A. (2022). LSTM Recurrent Neural Network for Hand Gesture Recognition Using EMG signals. *Applied Sciences*, 12(19), 9700. <https://doi.org/10.3390/app12199700>
- Vujović, Ž. (2021). Classification Model Evaluation Metrics. *International Journal of Advanced Computer Science and Applications*, 12(6), 599–606. <https://doi.org/10.14569/IJACSA.2021.0120670>
- Zhang, X., Chen, X., Li, Y., Lantz, V., Wang, K., and Yang, J. (2011). A framework for Hand Gesture Recognition Based on Accelerometer and EMG sensors. *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans*, 41(6), 1064–1076. <https://doi.org/10.1109/TSMCA.2011.2116004>