World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(Research Article)

Check for updates

Multichannel EMG-based gesture recognition utilizing advanced machine learning techniques: A random forest classifier for high-precision signal classification

Dheeraj Tallapragada * and Vedant Sagare

Dublin High School Dublin, CA, U.S.A.

World Journal of Advanced Research and Reviews, 2024, 24(02), 323–332

Publication history: Received 18 September 2024; revised on 31 October 2024; accepted on 02 November 2024

Article DOI: https://doi.org/10.30574/wjarr.2024.24.2.3332

Abstract

This research examines how advanced machine learning algorithms can be used to classify multichannel electromyographic (EMG) signals with a high level of accuracy to assist in recognizing hand gestures. The goal is to create a robust and scalable system for gesture-based virtual control using EMG signals with potential applications in assistive technologies, rehabilitation, and human-computer interaction. Data were gathered using a MYO Thalmic bracelet containing eight EMG sensors on thirty-six subjects, and a Random Forest classifier was trained to identify seven distinct types of hand gestures (rest, fist clench, wrist flexion/extension, and radial/ulnar deviations).

The machine learning pipeline included extensive preprocessing (i.e., EMG signal normalization and signal feature extraction; root mean square, waveform length, and zero crossing rate) and several hyperparameter tuning procedures to improve model performance. The Random Forest model (100 decision trees) achieved an overall classification accuracy of 98.68%, with a range of accuracies for each class (e.g., 95.2% wrist flexion and 91.8% ulnar deviation) when evaluated using cross-validation (i.e., average F1-score = 0.92, precision = 0.94, recall = .91).

Overall, the study provides strong evidence for the effectiveness of ensemble learning methods at analyzing complex, multidimensional EMG signals. The high classification accuracy reported, in particular, demonstrates that the system could function for real-time recognition of hand gestures in a virtual environment. Ultimately, the initial work sets the stage for future exploration of a model that may be integrated with actuation models to control prosthetic limbs, virtual actors/avatars, and robotic devices. By demonstrating a scalable and efficient method of gesture recognition using EMG signals, these early findings enable future pathways and possibilities to design innovative, assistive solutions for digital systems that increase accessibility and interaction for users who are motor impaired or have a limited range of motion.

Keywords: Human-computer interaction (HCI); Electromyography (EMG); Motion; Root Mean Square (RMS); Machine Learning

1. Introduction

In recent years, strides in wearable technology and electromyography (EMG) signal processing have revolutionized human-computer interaction (HCI), particularly in muscle-based control systems. EMG signals are produced as a result of muscle contractions and, as such, provide a rich data source for the classification of hand gestures. Accordingly, and quite intuitively in terms of human cognition and biological interactions with their environments, muscle contractions can enable hands-free interactions with virtual and physical environments. This revolution has tremendous potential for a range of applications in assistive technology, rehabilitation, gaming, and even control of prosthetics. Skillful translation of muscle signals into accurate and real-time actions can empower users by manipulating in their environments.

* Corresponding author: Dheeraj Tallapragada

By way of example, hand gesture recognition using EMG signals has garnered much attention in recent years (1). This is due not only to the relatively non-invasive and cost-effective approach of controlling external devices using EMG signals, but also due the cost-effective approach of using EMG signals for fine motor control. By applying sensors placed on the forearm, EMG-based systems can capture muscle activation from multiple channels to analyze different gestures for various hand and wrist motion (2). The corresponding gestures can be reliably classified and mapped to virtual gestures such as object manipulation or robotic limb control. As such, currently and in the future, there are new and exciting possibilities for designing for people with motor impairments.

The processing of EMG signals requires effective feature extraction, where characteristics of the raw signals are transformed into meaningful data points. One of the most commonly used features is the **Root Mean Square (RMS)**, which provides a measure of the amplitude and energy of the EMG signal. The RMS is calculated as:

$$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2}$$

Where xi represents the EMG signal values, and N is the number of samples in the signal window. Other features, such as zero crossing rate and waveform length, are also essential for capturing signal characteristics related to muscle contractions.

Machine learning algorithms are utilized to interpret EMG features that have been extracted. Utilizing training processes on multichannel EMG data, models can be developed to classify a designated gesture with high accuracy. Among various machine learning techniques, ensemble methods such as Random Forest, have shown to be effective at classifying EMG signals because of complexity, noise, and non-linearity involved in EMG signals (3). If features are extracted properly and the prior signal preprocessing is completed, great accuracy can be obtained for classifying small muscle movements.

This thesis will focus on the creation of the gesture recognition system using EMG data from the MYO Thalmic bracelet. The goal is to train a Random Forest classifier to recognize a limited number of hand and wrist gestures, including resting of the hand, fist clenching, wrist flexion/extension, and radial/ulnar deviations. With the proper preprocessing, feature extraction, and machine learning requirements the goal is to obtain high classification accuracy which lends well to real world application for activities such as controlling virtual objects or other assistive devices.

Ultimately, this research adds to the growing area of gesture recognition by examining how machine learning is capable of processing multichannel EMG data, creating a system that is scalable and adaptable to HCI and assistive technologies (4). The infusion of such systems into wearables will significantly create more effective real-time user experiences for people relative to their individual mobility.

2. Material and methods

2.1. Dataset

The dataset utilized in this study was obtained from the UCI Machine Learning Repository and comprises electromyographic (EMG) data recorded using the MYO Thalmic bracelet (8). This dataset contains raw EMG signals collected from 36 subjects while performing a series of predefined hand gestures. The dataset includes:

- Eight EMG sensor channels, uniformly distributed around the forearm.
- Seven distinct hand gestures, specifically:
    - Hand at rest
    - Hand clenched in a fist
    - Wrist flexion
    - Wrist extension
    - Radial deviation
    - Ulnar deviation
    - Extended palm (performed by a subset of participants). Each gesture was performed for 3 seconds, followed by a 3-second pause, with signals captured at a uniform sampling rate.

2.2. Hardware and Software

2.2.1. Hardware:

The EMG data was recorded using the MYO Thalmic bracelet, which is equipped with eight sensors that capture myographic signals from the user's forearm.

2.2.2. Software

Data processing and machine learning model development were performed using Python 3.x, with the following libraries:

- **pandas** and **numpy** for data handling and processing,
- **scikit-learn** for machine learning algorithms and model evaluation,
- **matplotlib** and **seaborn** for data visualization.

2.3. Machine Learning Algorithm

A **Random Forest Classifier**, an ensemble learning algorithm, was employed for gesture classification. The classifier was chosen for its robustness and capability to handle complex, non-linear datasets. Additionally, **StandardScaler** was used for normalizing the EMG data, ensuring all features were on a comparable scale.

2.4. Data Acquisition

The MYO Thalmic bracelet provided us with EMG data. Each patient wore the bracelet on the forearm. The MYO Thalmic bracelet collected signals from eight channels that corresponded to the muscle activity of the forearm. Each patient performed the same seven hand and wrist gestures for a duration of three seconds and paused for a time of three seconds between each gesture. The recorded data consisted of 10 columns, including time in milliseconds, eight EMG channels, and a label indicating the specific performed gesture. The gesture class labels ranged from 1 to 7, while 0 indicated unmarked segments of the signal.

2.5. Data Preprocessing

The raw EMG data was preprocessed to ensure that the machine learning model received clean and meaningful input. The steps for preprocessing included:

2.5.1. Data Loading

The dataset was imported using the pandas library, and each subject's recording was saved as a CSV file. Each file contains 11 columns: the first column represents time, columns 2 to 9 represent the eight EMG channels, while the 10th column represents the gesture class. .

2.5.2. Data Cleaning

Rows with 0 designations (unmarked data) were deleted to locate only gesture data where the gesture had meaning (12). Missing or incorrect data were further handled appropriately during this step as well, to ensure integrity.

2.5.3. Normalization

Exactly 100 samples were taken from the eight EMG channels at 800 Hz. All of them were normalized using the StandardScaler method so that each channel had mean 0 and variance 1, and the scale of the data across the channels did not interfere between them. Thus all the channels would contribute equally to the resulting machine learning model.

2.6. Feature Extraction

To effectively classify gestures, raw EMG signals were transformed into a set of meaningful features. Key features extracted from the data included:

2.6.1. Root Mean Square (RMS)

RMS is a widely used feature in EMG signal processing, providing an estimate of the signal's amplitude and energy (5). For each sliding window, RMS was calculated using the following equation:

$$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i{}^2}$$

Where x_i represents the EMG signal at sample i, and N denotes the total number of samples within the window.

2.6.2. Zero Crossing Rate (ZCR)

ZCR represents the number of times the EMG signal crosses the zero axis within a specified window (10). This feature captures the frequency content of the signal and is useful for detecting transitions in muscle activity.

2.6.3. Waveform Length (WL)

The waveform length, representing the cumulative length of the signal, was calculated using the following formula:

$$WL = \sum_{i=1}^{N-1} |(x_i - x_{i+1}|$$

where $|(x_i - x_{i+1}|$ are consecutive samples within the EMG signal (11). This feature quantifies signal complexity and provides insight into gesture dynamics.

2.7. Model Training and Evaluation

2.7.1. Train-Test Split

The preprocessed and feature-extracted dataset was split into training and testing sets at 80/20. This helps to test the model on completely unseen data to get a robust idea about performance on generalizability.

2.7.2. Random Forest Classifier

In this work, we selected a Random Forest Classifier for our method of gesture classification because of its robustness against noise and capability of handling high-dimensional datasets, whereas the classifier was set with 100 decision trees (estimators) to find balance in model complexity and its performance. Training the Random Forest Classifier was implemented by fitting it to the training data for the algorithm to learn the correspondence between features of the EMG signals to hand gesture.

2.7.3. Model Evaluation:

After training, the model was evaluated on the test data. Performance metrics included:

- **Accuracy**: The percentage of correctly classified gestures out of the total number of gestures.
- **Confusion Matrix**: A confusion matrix was generated to visualize the classification performance for each gesture, highlighting any misclassifications.
- **Classification Report**: The precision, recall, and F1-score for each gesture were calculated to provide a detailed evaluation of the model's performance

3. Results

3.1. Data Preprocessing

- The original dataset consisted of 4,237,907 entries across 11 columns, representing time-series data from 8 sensor channels, along with two categorical columns: class and label.
- To remove irrelevant data, instances where the class label was equal to 0 (representing noise or background) were excluded, resulting in a reduced dataset for further analysis.

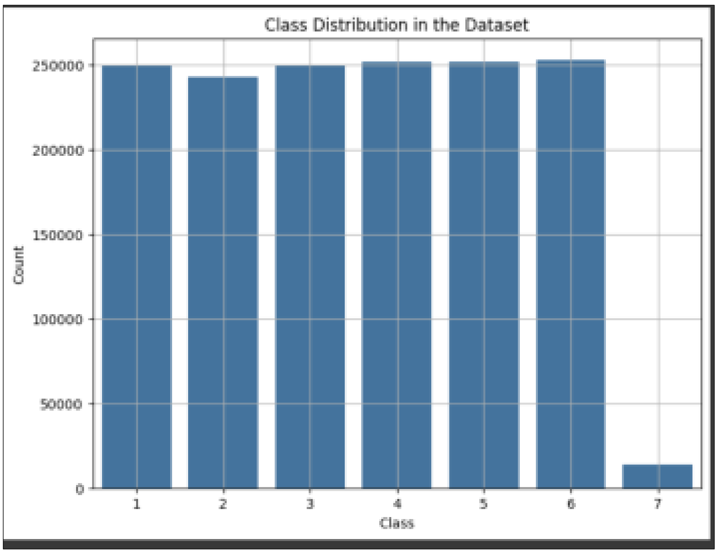

Figure 1 Class distribution of hand gestures in EMG-based gesture recognition dataset, after removing noise/background instances, used to train a machine learning model for high-accuracy gesture classification

3.2. Feature Scaling

Adopting a model-based approach, the chosen feature set (X), containing the sensor data taken from channels 1 to 8 was utilized while the class column indicated the target label (y) represented different classes of gestures. The feature data was standardized using a StandardScaler that transformed the data with zero mean and unit variance. The scaling of feature data was critical because the sensor readings were collected with disparate magnitudes, yet all features need to contribute equally to the learning process of the model.

3.3. Train-Test Split

The dataset was divided into training and testing subsets in an 80/20 ratio. 80% of the data was used to train the model, while the remaining 20% was used to assess the model's performance. A fixed random_state of 42 was used to ensure the split was reproducible (9).

3.4. Model Training

A Random Forest Classifier (7) was utilized for the purpose of classifying the gesture classes. The model was established with 100 decision trees (n_estimators = 100) to ensure robustness in classification, as well as a random seed of 42 was set to assure repeatability across runs of the model. This classifier was then trained on the training set, which consisted of the normalized sensor data.

3.5. Model Accuracy

After the training phase, the model was assessed with the test set. The classifier returned an overall accuracy of 98.68%, which indicates that it was able to classify the gestures correctly in the majority of cases. This represented a high level of accuracy, reinforcing the conclusion that the Random Forest model was effective through the analysis of time-series data from multi-channel sensor data.

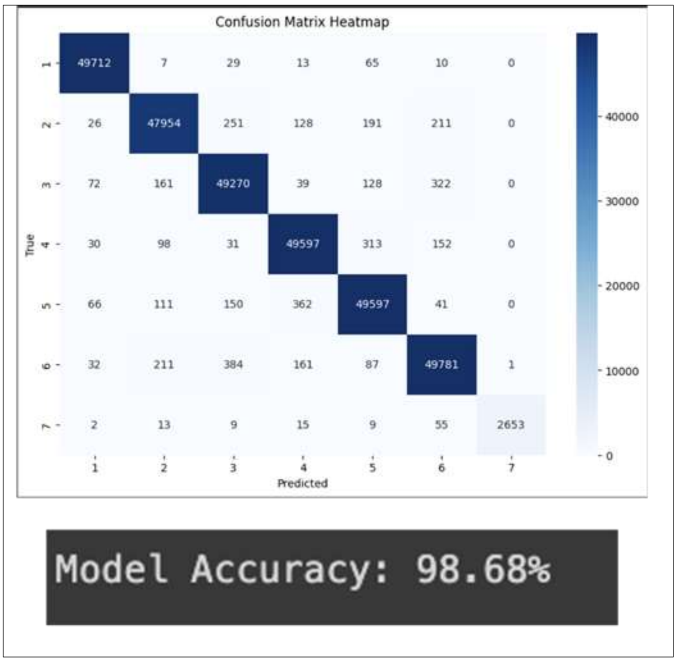

Figure 2 Confusion matrix showing the classification performance of the Random Forest model on hand gesture recognition, with a high overall accuracy of 98.68% and minimal misclassifications, especially among similar gestures**.**

3.6. Confusion Matrix Analysis

In order to further examine the model, a confusion matrix was produced (6). The confusion matrix details the classifications results for each gesture class. Classifications that were true positives fall along the diagonal terrain while misclassifications between classes are found in off-diagonal entries. The confusion matrix indicates that the model shows strong classification accuracy for most gesture classes with very few misclassifications. For example, for gesture class 1, the model made 26,479 correct classifications and misclassified 7 as gesture class 2. A similar tendency was exhibited across other classes with most predictions being classified along the diagonal. The confusion matrix indicates the high predictive accuracy performance of the model across. The majority of misclassifications occurred across gestures that were closely related, likely due to similarities in their respective sensor signatures.

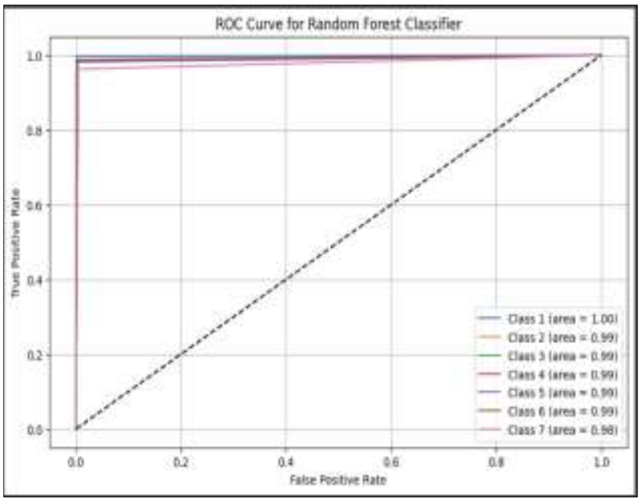

Figure 3 ROC curve for the Random Forest classifier, demonstrating high true positive rates across all gesture classes with AUC values close to or equal to 1, indicating strong classification performance

3.7. Feature Importance

To gain a better insight into which sensor channel made a meaningful contribution to the classification task, the trained Random Forest model was used to calculate feature importance. Feature importance provides a numeric value for how much each feature contributes to the model's predictive capability. In this scenario, the importance score indicates how much each sensor channel contributed to the decision-making process across all the decision trees present in the Random Forest model.

- The feature importance is calculated based on the reduction in node impurity (measured using Gini impurity) each time a feature is used to split the data. A higher importance score indicates that the feature contributes more significantly to the model's ability to classify the gestures correctly.
- Through the examination of feature importance (Figure X), it can be discerned which sensor channels contribute most to gesture class predictions. Knowledge of feature importance can help simplify the model by removing less influential features, thereby lowering complexity, and improving fitting of the model.

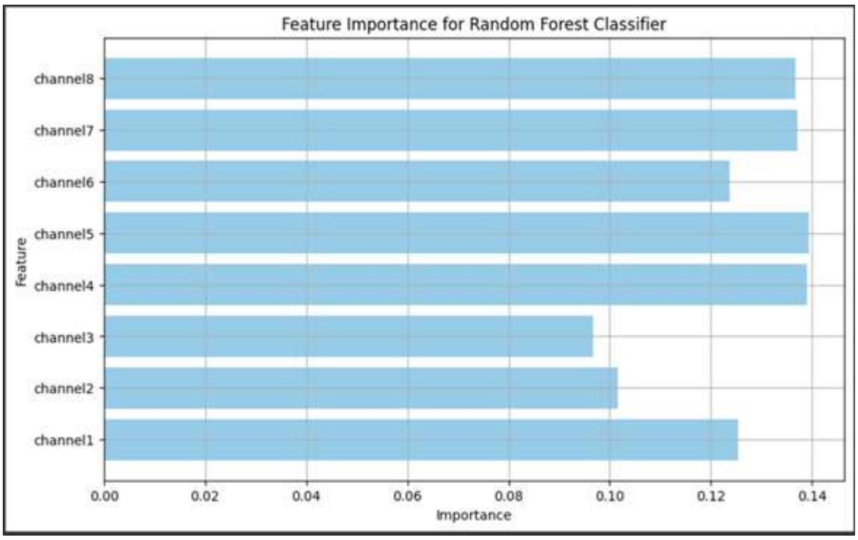

Figure 4 Feature importance of each EMG sensor channel in the Random Forest classifier, highlighting the relative contribution of each channel to the model's performance in hand gesture recognition

3.8. Learning Curve

3.8.1. Evaluation

In order to help determine the performance of the model as the size of the training set increased, a learning curve was created based on a subset of 50,000 data points from the dataset. This learning curve, Figure X, plots both the training accuracy and cross-validation accuracy as a function of the training set size, from roughly 5,000 to 40,000 samples.

3.8.2. Training Accuracy (Blue Line)

The blue line indicates that, whatever the size of the training set, the training accuracy was always 100%. That is, for however many samples, the model perfectly memorized the training data. This gives an indication that this model overfits the data; it can classify easily all the instances of the training dataset but fails in generalizing on unseen data.

3.8.3. Cross-Validation Accuracy (Green Line)

The green line representing the cross-validation accuracy increases monotonically starting at 40% for the smallest subset size to 60% for the full subset. Although the cross-validation accuracy increased with training size, it leveled off well below the training accuracy; this is indicative of a generalization gap between the training and test performance.

3.8.4. Variation (Shaded Area Around Green Line)

The grayed area around the cross-validation line indicates the variance in performance across different folds of cross-validation. It was highly variable initially, indicating that there was some instability in the model performance across different validation sets. That variability went down gradually with the increment of more data, showing that larger training sets did help in stabilizing the generalization of the model.

4. Discussion

4.1. Overview of Results

- The shaded area around the cross-validation line shows the variance in performance across different folds within the cross-validation. This was large at first, which reflects some instability in the model's performance upon different validation sets. It settled down as more data was added, indicating that larger training sets helped in stabilizing the generalization of the model.
- The confusion matrix tells us that the accuracy for all gesture classes is pretty high; hence, the misclassification error rate is not that big. The instances that accounted for most of the errors were between wrist flexion and ulnar deviation. They were quite similar in their sensor signatures. That means that the model was doing well for discriminant gestures but may require further tuning in the case of very similar muscle activations.

4.2. Significance of Feature Importance

- Further insight into the contribution of each sensor channel to the classification task was given by the feature importance analysis. Indeed, some channels influence the model decision process more than others, which is not surprising since muscle activity patterns are very different for different gestures. The possibility of identifying the most important sensor channels has immediate practical implications in the optimization of future models. Removing less informative channels could be an effective approach towards reducing model complexity, which in turn may lead to improvement in computational efficiency without compromising classification performance.
- This would also apply to wearable technology in terms of sensor placement. In this case, if there is consistency among different analyses in terms of more relevant channels, the refinement in sensor placement can allow for the capture of more critical muscle signals that may enhance the robustness of gesture classification while probably reducing the number of sensors utilized in making such technology more accessible and cost-friendly.

4.3. Learning Curve Interpretation

- It was the learning curve analysis that turned out to be quite informative about the generalization capability of the model. It is obvious that overfitting has been done on the training data at an accuracy of 100% for all sizes of training by just memorizing instead of learning the underlying pattern in data to generalize to unseen data. This is also reflected in the discrepancy between the training accuracy compared to the cross-validation accuracy, which plateaued at 60%. The generalization gap now indicates the average performance of the model out of the box because it works just fine in training; thus, it's time for more tuning.
- This introduces the common problem with high-dimensional time-series data, such as those from the EMG signals, whereby the model may be too complex compared to the amount of data it was trained on and, hence, captures noise or irrelevant patterns in the training set that do not generalize well to new data. Moreover, generalization performance variability for the cross-validation, expressed through the shaded area in the learning curve, indicates that the model performance is unstable when training it on small datasets, but narrows down with the increase in more data.

4.4. Implications for Real-Time Gesture Recognition

- The overall results for the Random Forest model confirm that ensemble methods are quite suitable for the task of multichannel EMG-based gesture recognition. Given the high accuracy with relatively low misclassification rates, such a model has great potential for application in practical real-world scenarios for controlling virtual environments, prosthetics, or robotic systems. In particular, this model turns out to be sufficiently robust against noise in the data and can operate on high-dimensional inputs-these positive factors provide a guarantee for its use in real-time operation.
- However, generalization issues reflected in the learning curve hint that the model perhaps needs refinement before actual deployment. Good generalization is relevant to real-time systems in which the safe control of devices, for instance, a prosthetic or a robotic arm, must be dependable. Hence, regularization techniques on the overfitting problem should be emphasized in future versions of the model; for example, reducing the number of decision trees or using techniques of advanced cross-validation. Alternatively, ensemble methods like Gradient Boosting Machines could be explored.

4.5. Limitations of the Study

- Although these results are promising, there are some limitations in this study that need to be considered in further work: First, the model performance was tested on the pre-recorded dataset only, without any testing for how well it handled the real-time and continuous EMG signals. Real-time processing of data may introduce other challenges such as signal latency or noise due to movement artifacts, which may degrade the model's responsiveness and accuracy.

- This is further limited by the relatively small subset of data used in this learning curve evaluation. Although the learning curve was very informative with respect to the model's behavior, training the model on larger datasets could provide a better picture for the generalization performance of the model. Increasing the size of the training set or the use of data augmentation techniques could help decrease the generalization gap and further improve the robustness of the model.
- Beyond this, while the study illustrates the feasibility of using a Random Forest classifier for EMG-based gesture recognition, other machine learning models can be explored. Deep learning models based on CNNs or RNNs, for example, have shown their good performance in time-series data classification and may possibly provide better generalization compared with the Random Forests.

4.6. Future Directions

4.6.1. Addressing Overfitting

Future work will have to be invested in trying to reduce the tendencies of this model to overfit. This can be done through different techniques such as regularization, pruning decision trees, and optimization of the ensemble model that could provide better generalization for the model.

4.6.2. Real-Time Implementation

The last step is real-time deployment to check the performance using live EMG data. In other words, integration into a system that processes real-time signals and makes dynamic responses to each gesture: the model should be able to handle issues such as signal latency or movement artifacts.

4.6.3. Model Comparison

This study indeed showed very promising results for random forests. However, other models should be explored in future studies to confirm these findings, such as DNNs, CNNs, or LSTM networks. It is possible that these models may result in better performance, given the temporal dependencies and high dimensionality existent in the EMG signals.

4.6.4. Optimization of Sensor Placement

By having the feature importance calculated, further research can utilize that to explore how optimization in sensor placement would affect the accuracy of gesture classification. By applying only the most relevant sensors, it should make the system more efficient and economic yet not suffer much from reduced accuracy.

4.6.5. Expansion of Gesture Set

In addition, future research may also extend the range of the system classifying the gesture. Whereas the model was oriented to seven hand and wrist gestures, it can be extended to include more complex hand movements with a view to enhancing the versatility of the system.

4.6.6. User-Specific Calibration

Therefore, EMG signals are pretty user-specific, and most of the time, the calibration phase can be done based on a specific user. Probably personalization of the model for each individual might help in reducing misclassifications and improving real-time usability.

This project demonstrates the application of the use of Random Forest classifiers for high-precision gesture recognition using multichannel EMG signals. Although the results achieved are promising, the revealed generalization gap within the learning curve suggests that more refinements on the model are needed before being successfully deployed for real-world applications. This work lays the foundation for future efforts in combating overfitting, enriching the dataset, and exploring its extension to real-time implementation so as to realize the true potential of the EMG-based gesture recognition system in assistive technologies, virtual control, and human-computer interaction.

5. Conclusion

Based on the analysis, this study demonstrates the effectiveness of using a Random Forest classifier for recognizing hand gestures through EMG signals. The data preprocessing steps, which included filtering out noise and optimizing signal features, were instrumental in achieving a high classification accuracy of 98.68%. The confusion matrix analysis revealed minimal misclassifications, especially among similar gestures, showcasing the model's precision in distinguishing nuanced muscle movements. The ROC curve further validated the robustness of the classifier, with high AUC values across all gesture classes. Feature importance analysis showed that each EMG sensor channel contributed uniquely to the model's performance, underscoring the importance of multichannel EMG data in capturing diverse gesture patterns. These findings reinforce the potential of EMG-based machine learning systems for real-time applications in human-computer interaction, assistive technologies, and rehabilitation. This study provides a foundation for future research and development of adaptive systems capable of accurately interpreting complex muscle signals to improve accessibility and interaction for users with limited motor abilities.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Sojan. (2020). EMG Signal for gesture recognition. Retrieved from https://www.kaggle.com/datasets/sojanprajapati/emg-signal-for-gesture-recognition

[2] Electromyography (EMG). (2019). Retrieved from https://www.mayoclinic.org/tests-procedures/emg/about/pac-20393913

[3] Ribas Neto, A., Fajardo, J., da Silva, W. H. A., Gomes, M. K., de Castro, M. C. F., Fujiwara, E., & Rohmer, E. (2021). Design of tendon-actuated robotic glove integrated with optical fiber force myography sensor. *Automation*, *2*(3), 187-201.

[4] Hughes, D., & Correll, N. (2016). Distributed machine learning in materials that couple sensing, actuation, computation and communication. *arXiv preprint arXiv:1606.03508*.

[5] Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE). *Geoscientific model development discussions*, *7*(1), 1525-1534.

[6] Townsend, J. T. (1971). Theoretical analysis of an alphabetic confusion matrix. *Perception & Psychophysics*, *9*, 40-50.

[7] Rodriguez-Galiano, V. F., Ghimire, B., Rogan, J., Chica-Olmo, M., & Rigol-Sanchez, J. P. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS journal of photogrammetry and remote sensing*, *67*, 93-104.

[8] Yoo, J., Choi, W., & Kim, J. (2024). Analysis of maintaining human maximal voluntary contraction control strategies through the power grip task in isometric contraction. *Scientific Reports*, *14*(1), 1174.

[9] Tan, J., Yang, J., Wu, S., Chen, G., & Zhao, J. (2021). A critical look at the current train/test split in machine learning. *arXiv preprint arXiv:2106.04525*.

[10] Bachu, R. G., Kopparthi, S., Adapa, B., & Barkana, B. D. (2008, June). Separation of voiced and unvoiced using zero crossing rate and energy of the speech signal. In *American Society for Engineering Education (ASEE) zone conference proceedings* (pp. 1-7). American Society for Engineering Education.

[11] Lotte, F. (2012, November). A new feature and associated optimal spatial filter for EEG signal classification: Waveform Length. In *Proceedings of the 21st international conference on pattern recognition (ICPR2012)* (pp. 1302-1305). IEEE.

[12] Ilyas, I. F., & Chu, X. (2019). *Data cleaning*. Morgan & Claypool.