

EMG Signal Analysis and Real-Time Muscle Activation Classification using Random Forest

Sayand K K¹, Jishnu Vijayan², M Jithin Vas³, Vidhya K⁴, Hema P Menon⁵ and Davidson Devasia⁶

¹⁻⁵Department of Computer Science and Engineering, Sreepathy Institute of Management and Technology, Vavanoor, India
Email: {sayand.kk, jishnuvijayan, jithinvas.m, vidhya.k, hemapmenon}@simat.ac.in

⁶Department of Pediatric Neurology, Aster Medcity, Kochi, Ernakulam, Kerala, India
Email: davidson.devasia@asterhospital.com

Abstract—Electromyography (EMG) plays a crucial role in the detection and diagnosis of muscle activities. It was found that EMG signals captured from body muscles can be used for human-computer interfaces and prosthetic control. This paper presents Machine Learning based muscle state identification through classification of the EMG obtained at real-time using surface electrodes and Advancer Technologies Muscle Sensor v3. In this work, the signals have been captured from the bicep muscle and are interfaced using an Arduino UNO. The signals are filtered using low-pass filter, moving average filter and peak- to-peak measurement method prior to further processing. The system was found to achieve an accuracy of 99% using random forest, trained on an average of 3000 data values representing the bicep muscle activation and relaxation obtained from 10 individuals. A comparison has been done with the commonly used Support Vector Machine (SVM) method also. The developed classification model is further used for control of opening and closing of computer-based applications.

Index Terms— Electromyography signals (EMG), Muscle sensor, Machine learning, Random Forest, Support Vector Machine (SVM), Surface electrodes, Arduino.

I. INTRODUCTION

Electromyography (EMG) signals are electrical patterns generated by muscle activity, widely used in applications such as prosthetic control, personalized rehabilitation, muscle anomaly detection, and enhancing human-computer interaction. EMG is measured using either invasive electrodes inserted directly into the muscle or non-invasive surface electrodes (sEMG) placed on the skin. The signals arise from the voltage difference between the muscle and a reference point, often a bony area like the elbow. Proper electrode placement on the targeted muscle allows for the collection of EMG signal variations. External noise can affect the signals, so applying effective noise reduction techniques and selecting an appropriate machine learning model is vital for ensuring reliable results in applications that support amputees and patients.

II. RELATED WORKS

Electromyography (EMG) is a technique used to evaluate and record the electrical activity produced by human muscles, and a few studies have explored its integration as follows: S. Inam et al., [1] has presented a review on the strategies that are used for the classification of EMG signals with emphasis on gesture estimation.

The authors have also quoted the use of EMG signals for prosthetic hand control. The paper by Ayachi and Seddik [2], explores the methods available for analysis of EMG signals commonly used for prosthetic devices. They discuss techniques for preprocessing the signal, filtering out noise, and detecting the signal envelopes to improve accuracy. The paper also discussed effective methods for feature extraction and classification of EMG signals. Additionally, they present the design of a multifunctional prosthetic hand that is claimed to have achieved 98.3% classification accuracy. The EMG signal can be acquired from human body parts using MyoWare muscle sensors which are available as wearable devices. One such work has been reported by Ajithkumar L, Kumar P, Vijaya Krishna B, Pradeep Kumar [3]. In their work, the sensor was used for measuring the muscle power level of bodybuilders and weightlifters. A study on the location where the sensors are placed for acquiring the muscle signal has been discussed by R. Lopez and T. C. Davis [4]. Their paper focuses on the muscle activity during the regular walking of a person. Following this, the survey reported by K. S. Pranoti and S. Guruprasad [5] focuses on the performance of various techniques for the effective and efficient analysis and classification of EMG signals. Moving on, D. Minh Dao, P. Dang Phc, T. Xuan Tuy, and T. Thuy Le [6], in their work, have developed a system for identifying the quadriceps muscles' response during knee movement. The system has been reported to be developed using an Arduino UNO controller. Further research on the use of surface EMG for identifying both surface and deep layer muscle activity has been done by many researchers over the decades. One such research has been done by T. Koshio, S. Sakura Zawa, M. Toda, J. Akita, K. Kondo, and Y. Nakamura [7], where the intersection of the muscles was used for placing the electrodes. The authors reported that the muscles were found to act independently for different kinds of body movements. The signals acquired using surface EMG contain noise which needs to be filtered prior to any further processing and analysis. Research over the period has experimented using different types of filters for obtaining noise-free signals.

A work using a Butterworth IIR filter for surface EMG signal filtering has been reported by Z. Xiao, J. Ye, H. Shen, S. Deng, H. Zhu, and H. Xiao [8].

There are various methods available for the classification of data. In recent years, artificial intelligence and machine learning based approaches are gaining a lot of prominence. Many researchers have reported using machine learning algorithms for hand movement classification. The work focuses on controlling a prosthetic hand using EMG-based machine learning done by M. Karuna and S. R. Guntur [9]. The work also presents a comparison of LDA, SVM, KNN, and DT algorithms for hand movement classification. J. R. Cardenas-Valdes, D. Valdes-Luis, M. d. J. Garcia-Ortega, A. H. Corral Dominguez, and J. A. Galaviz Aguiler [10] have developed a system for the classification of EMG signals based on the flexor digitorum profundus muscle activities using an SVM classifier. The authors have claimed an accuracy of 93.3%. The utilization of MyoWare Muscle sensors for detecting muscle activity in forearm muscles, combined with an Arduino UNO microcontroller to collect EMG signals, which are then transformed using short-time Fourier transform (STFT) to analyze them in the time-frequency domain. This approach has been discussed by Jose Paredes, Malena Loza, and Diego Benítez [11]. S. Bhavani, K. Lokesh Krishna, R. B. Yogananda Reddy, and T. Geethika [12] have developed a low-cost bionic arm by collecting the EMG signal using a MyoWare sensor and Arduino for the movement of the fingers. The use of a MyoWare sensor has allowed the development of the bionic arm to realize the natural movements of the hand. Recognizing finger movements from EMG data is a key challenge in prosthetic finger development. The study by Mucahit Findik, Seyma Yilmaz, and Mehmet Koseoglu [13], focus on automatically generating and selecting EMG features. Their random forest classification method stating an accuracy of 97.5% and highlights important features for distinguishing finger movements. The authors Sakshi Taori and Sol Limused [14] used armband sensor to detect the muscle activity and used machine learning to identify the lifting and lowering movements, this paper compare three machine learning models, random forest, support vector machine and logistic regression, the random forest and SVM shows best accuracy.

Phuc Thanh Thien Nguyen, Shuh Feng Su and Chung Hsien Kuo [15], in their paper introduces a way to recognize the hand gestures using muscle signals with the help of neural network for controlling the prosthetic hand. The authors have claimed an accuracy of 89.56%. EMG sensors are quite expensive nowadays; to address this, Panagiotis Flogeras, Dimitrios Tsipianitis, and George Souliotis [16] have designed a cost-effective EMG circuit to acquire EMG signals.

III. METHODS AND MATERIALS

The overall working of the proposed system is given in Fig 1. The stages are data acquisition, data preprocessing, dataset modification, classification model selection and prediction of muscle activation. Fig 2. shows the Pseudocode

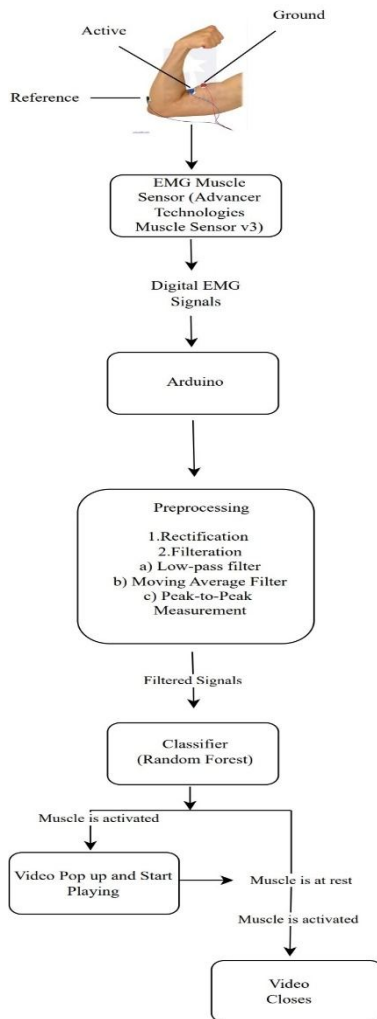


Fig. 1. System Architecture

Pseudocode:

```

1. Initialize EMG sensor
   EMG_Sensor = Advancer_Technologies_Muscle_Sensor_(V3)

2. Data Acquisition
   EMG_data = EMG_sensor.collect_data()
   Store EMG_data in CSV file with columns: Timestamp, Raw_EMG_Signal

3. Preprocessing
   Input: Raw_EMG_Signal
   Output: Preprocessed_EMG_Signal
   a) Apply Moving Average Filter
      Smoothed_signal = Moving_Average_Filter(Raw_EMG_Signal)
   b) Apply Low-Pass Filter
      Filtered_signal = Low_Pass_Filter(Smoothed_signal)
   c) Perform Peak-to-Peak Analysis
      PkPk_signal = Peak-to-Peak_Analysis(Filtered_signal)
   d) Determine Threshold
   e) Hand States
      For each signal:
        If signal > Threshold:
          Label = "Muscle is Activated"
        Else:
          Label = "Muscle is at relax"
   f) Store the Preprocessed Data
      Preprocessed_data = [Timestamp, Filtered_signal, Label]

4. Model creation:
   a) Train:
      SVM_Model = SVM(Preprocessed_data)
      Random_Forest_Model = RandomForest(Preprocessed_data)
   b) Test:
      SVM_Model.predict(Preprocessed_data - Label)
      Random_Forest_Model.predict(Preprocessed_data - Label)

5. Classification
   a) Input: Features, Label
   b) Output: Classified muscle state (Rest or Activated)
   c) Classify new EMG data
      For each new_signal:
        Predict(Random_Forest_model(new_signal))

6. Application:
   a) If 10 consecutive prediction == "Muscle is activated":
      Video opens and start playing
   b) If Subsequent 10 prediction == "Muscle is activated":
      Video stops and closes.
  
```

Fig. 2. Pseudocode

A. Data Acquisition

The data was collected using sEMG electrodes, Advancer technologies muscle sensor v3 and Arduino with a focus on activation and relaxation of the bicep muscle. For proper data acquisition subject was instructed to maintain the bicep muscle in both relaxed and activated states in different time intervals. The dataset includes subjects of the age group 18 - 25, which includes both male and female. The dataset comprises of timestamps and digitized EMG signal values. The digitized signals, initially containing noise, were processed in subsequent steps.

The hardware setup includes several key components that work together to capture EMG signals from muscle and classify the muscle activation and relaxation states.

The surface electrodes are specialized sensors that are placed on the skin surface above a specific muscle to capture EMG signals. Surface electrodes are commonly used for non-invasive EMG signal acquisition and provide a convenient method for capturing muscle activity during muscle activation. Three electrodes were used Active, Reference and Ground. The Active electrode is placed at the proximal end of the muscle, Ground electrode is placed at a distance from the Active electrode, and Reference is placed in a non-muscular region near to that muscle. The electrodes were tested and verified by ASTER MEDCITY, Kochi.

Advancer technologies muscle sensor v3 was used in order to collect EMG signal using surface EMG (sEMG) electrodes. The sEMG electrode is connected to muscle sensor using EMG cable. This component plays a crucial role by collecting and amplifying the EMG signals. These amplified signals are transmitted to Arduino UNO.

Arduino UNO collects the signals from the EMG muscle sensor and it transmit the signal to the computer. While preprocessing could be performed on the Arduino itself, in this project, the preprocessing was done on a

computer for simplicity, and as the focus was on a computer-based application. Here Arduino acts as the interface of the system.

In Fig 3, (a) shows the surface electrode used to acquire sEMG from muscles, (b) shows the Advancer Technologies Muscle sensor v3 used to collect sEMG signals from the electrodes, (c) shows the Arduino UNO which is used as an interface.

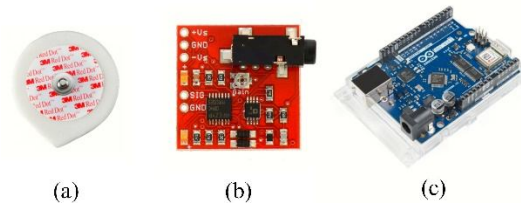


Fig. 3. Hardware components: (a) Surface electrode, (b) Advancer Technologies Muscle Sensor v3, (c) Arduino UNO

In Fig 4, (a) shows the muscle relaxed state and (b) shows the muscle activated state.

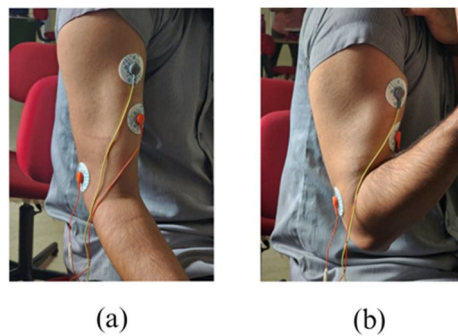


Fig. 4. Collection of EMG signals: (a) Muscle is at rest, (b) Muscle is activated

In Fig 5, (a) shows the plot of raw EMG, where time in milliseconds (ms) on the X-axis and amplitude on the Y-axis, (b) shows the digitized value of that signal.

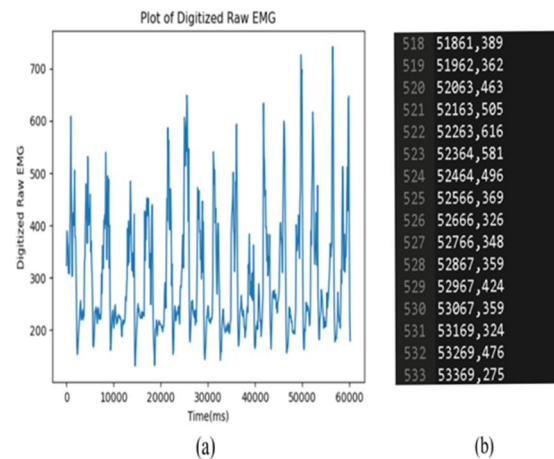


Fig. 5. (a) Signal waveform of acquired data, (b) Digitized raw data

B. Data Preprocessing

Preprocessing of signal is one of the important steps for better accurate model creation. Training a model with an incorrect dataset can lead to many anomalies. Here preprocessing of signal was done using:

Filtration: Three types of filtering techniques

- **Moving Average Filter:** This method calculates the moving average of the input data. This helps to smooth out short-term fluctuations and highlight long-term cycles. Equation 1 shows the Moving Average Filter, where N is the number of data points over which the average is calculated (window length) and $x(n)$ is the current data point.

$$MA(n) = \frac{1}{N} \sum_{i=0}^{N-1} x(n-i) \quad \text{—————} \quad (1)$$

- **Low-Pass Filter:** This method allows signals with a frequency lower than a certain cutoff frequency to pass through and neglects signals with a higher frequency than the cutoff frequency, which removes high-frequency noise from the EMG signal. Equation 2 shows the Low-Pass Filter, where f_s is the sample frequency, and f_c is the cutoff frequency.

$$LPF(n) = \frac{1}{\frac{0.125 \cdot f_s}{f_c}} \sum_{i=0}^{\frac{0.125 \cdot f_s}{f_c}} x(n-i) \quad \text{—————} \quad (2)$$

- **Peak-to-Peak Measurement:** This method measures the peak-to-peak amplitude of the signal within a specified frequency range. Equation 3 shows the Peak-to-Peak Measurement, where $\max(\text{data})$ is the maximum value in the window and $\min(\text{data})$ is the minimum value in the window.

$$PkPk(n) = \max(\text{data}) - \min(\text{data}) \quad \text{—————} \quad (3)$$

Rectification:

This step converts the signals to absolute values. This is a crucial step because rectification ensures that the signal represents the magnitude of muscle activity regardless of direction.

In Fig 6, (a) shows the plot of filtered EMG, where time in seconds (s) on the X-axis and amplitude on the Y-axis, (b) shows the digitized value of that signal.

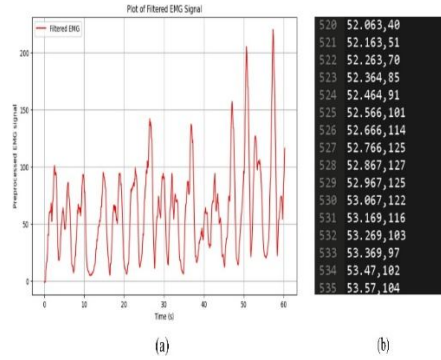


Fig. 6. (a) Signal waveform of filtered EMG, (b) Digitized filtered data

C. Classification models

For model creation, the dataset was modified by adding a muscle state column with 2 classes: 'Muscle is at relax' and 'Muscle is activated'. This column was generated based on a threshold value, determined by analyzing the pattern of the EMG signals. This dataset, stored in comma-separated values (CSV) format, was used to train two machine learning models. The first model was a Random Forest model which showed an accuracy of 99%. The second model was Support Vector Machine (SVM) and it's accuracy was 98%. These results indicate that both algorithms were highly efficient in accurately classifying the muscle state. Both models showed good results in real-time prediction also. Fig 7. shows the dataset used for training the classification models.

520	52.063,40,Muscle is at relax
521	52.163,51,Muscle is at relax
522	52.263,70,Muscle is at relax
523	52.364,85,Muscle is at relax
524	52.464,91,Muscle is at relax
525	52.566,101,Muscle is activated
526	52.666,114,Muscle is activated
527	52.766,125,Muscle is activated
528	52.867,127,Muscle is activated
529	52.967,125,Muscle is activated
530	53.067,122,Muscle is activated
531	53.169,116,Muscle is activated
532	53.269,103,Muscle is activated
533	53.369,97,Muscle is at relax
534	53.47,102,Muscle is activated
535	53.57,104,Muscle is activated

Fig. 7. Dataset

IV. RESULT AND ANALYSIS

An EMG signal analysis system was developed using machine learning models to classify muscle states, determining whether the muscle is activated or relaxed. This work showcases promising applications in various fields like prosthetics control in the medical field, various machine's on/off procedure control and so on. Fig 8 shows the experimental setup (a) Prediction of Muscle is at rest, (b) Prediction of muscle activated

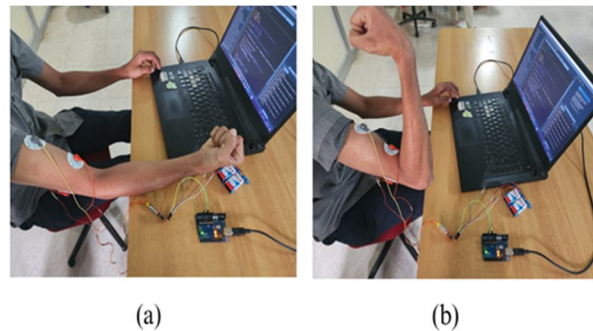


Fig. 8. Experimental setup: (a) Bicep muscle is relaxed, (b) Bicep muscle is activated

By doing this, it became evident that the output of these models can be used to control the on/off functions of various hardware and software. As a proof of this concept, a system was built that could recognize muscle activation, causing a video to pop up and start playing. A subsequent muscle activation would close the video. The Fig 9 shows the output of Fig 8 (b). A subsequent action in Fig 8 (b) will close that video.

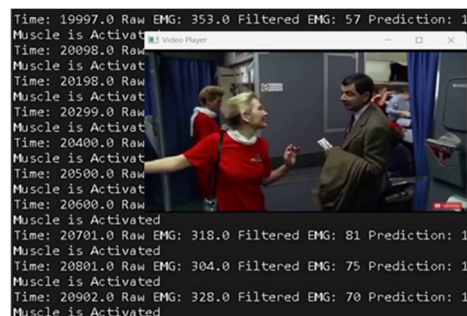


Fig. 9. Video Opens, when muscle is activated

In Fig 10, (a) shows the confusion matrix of Random Forest, (b) shows the confusion matrix of Support Vector Machine (SVM).

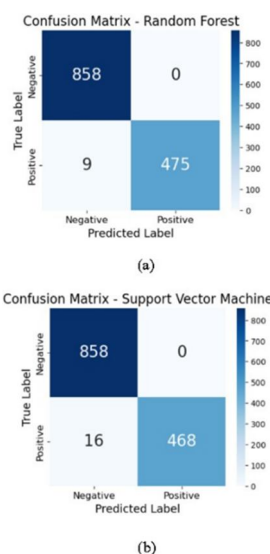


Fig. 10. Confusion matrices

TABLE I. PERFORMANCE METRICS OF RANDOM FOREST (RF) AND SUPPORT VECTOR MACHINE (SVM) ALGORITHMS

Algorithm	Accuracy	Muscle	Precision	Recall	F1 Score
Random Forest(RF)	0.99	At Relaxation	0.99	1.00	0.99
		At Activation	1.00	0.98	0.99
Support Vector Machine (SVM)	0.98	At Relaxation	0.98	1.00	0.99
		At Activation	1.00	0.97	0.98

V. CONCLUSION

In this work, a Random Forest-based classification system was developed for hand movement identification using EMG signals. The current work was carried out on a computer system connected to an Arduino UNO, which was utilized to acquire EMG signals through surface electrodes and a single Advancer Technologies muscle sensor v3. This work can be extended to identification of different muscle activities by using multiple sensors and taking the signals from muscles such as the Triceps brachii, Brachioradialis, Flexor carpi radialis. These findings can be then used for controlling prosthetic arm and also help with the rehabilitation activities. This work compares two classification algorithms, Random Forest and Support Vector Machine (SVM), which achieved accuracies of 99% and 98% respectively.

ACKNOWLEDGEMENT

The authors would like to thank Dr. K M Mathew, Dr. Dhinla and Mr. M C V Arun, Aster Medcity, Kochi for their expert advice and validation of the results throughout the work.

REFERENCES

- [1] Shehla Inam, Sana Al Harmain, Shehzaib Shafique, Mafia Afzal, Arfa Rabail, Faisal Amin, Muhammad Waqar, A Brief Review of Strategies Used for EMG Signal Classification, 2021 International Conference on Artificial Intelligence (ICAI), Islamabad, Pakistan, 2021, pp. 140-145.
- [2] Monaam Ayachi and Hassene Seddik, Overview of EMG Signal Pre- processing and Classification for Bionic Hand Control, 2022 IEEE Information Technologies and Smart Industrial Systems (ITSIS), Paris, France, 2022, pp. 1-6.
- [3] Ajithkumar L, Kumar P, Vijaya Krishna B, Pardeep Kumar, Biomechanical Aspects of Muscle Power Analysis by Using Myoware Muscle Sensor, 2022 International Conference on Smart and Sustainable Technologies in Energy and Power Sectors (SSTEPS), Mahendragarh, India, 2022, pp. 369-373.
- [4] Rhianne Lopez and T. Claire Davies, The effect of surface electromyography placement on muscle activation amplitudes and timing, 2016 IEEE EMBS International Student Conference (ISC), Ottawa, ON, Canada, 2016, pp. 1-4 1918- 1921.
- [5] K. S. Pranoti and S. Guruprasad, A Survey on Synthesis and Analysis of EMG Signal, 2023 International Conference on Recent Advances in Science and Engineering Technology (ICRASET), B G NAGARA, India, 2023, pp. 1-5.
- [6] Duc Minh Dao, Pham Dang Phuoc, Tran Xuan Tuy and Tram Thuy Le, Research on reading muscle signals from the EMG sensor during knee flexion Extension using the Arduino Uno controller, 2017 International Conference on Advanced Technologies for Communications (ATC), Quy Nhon, Vietnam, 2017, pp. 270-273.
- [7] T. Koshio, Shigeru Sakurazawa, Masashi Toda, Junichi Akita, Kazuaki Kondo and Yuichi Nakamura, Identification of surface and deep layer muscles activity by surface EMG, 2012 Proceedings of SICE Annual Conference (SICE), Akita, Japan, 2012, pp. 1816-1821.
- [8] Zeyu Xiao, Jiaxiang Ye, Hanxin Shen, Siyu Deng, Hao Zhu and Han Xiao, Analysis of Digital Filtering Design Based on Surface EMG Signals, 2023 IEEE 3rd International Conference on Electronic Communications, Internet of Things and Big Data (ICEIB), Taichung, Taiwan, 2023, pp. 542-546.
- [9] M. Karuna and S. R. Guntur, Classification of Hand Movements via EMG using Machine Learning Methods for Prosthesis, 2022 2nd International Conference on Artificial Intelligence and Signal Processing (AISP), Vijayawada, India, 2022, pp. 1-4.
- [10] Jose Ricardo Cardenas-Valdez, Daniela Valdez-Luis, Manuel de Jesus Garcia-Ortega, Angel Humberto Corral-Dominguez, Jose Alejandro Galaviz-Aguilar, SVM Classifier and evaluation of muscle power of EMG signals and Python implementation, 2020 International Conference on Mechatronics, Electronics and Automotive Engineering (ICMEAE), Cuernavaca, Mexico, 2020, pp. 20- 22.
- [11] Jose Paredes, Malena Loza and Diego Benitez, Towards Individual Finger Movement Detection Using EMG Sensors Placed on Forearm, 2023 IEEE CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON), Valdivia, Chile, 2023, pp. 1-5.

- [12] S. Bhavani, K. Lokesh Krishna, R. B. Yogananda Reddy and T. Geethika, A Low-Cost Bionic Arm Based on Electromyography Sensor, 2020 7th International Conference on Signal Processing and Integrated Networks (SPIN), Noida, India, 2020, pp. 896-901.
- [13] Mucahit Findik, Seyma Yilmaz and Mehmet Koseoglu, Random Forest Classification of Finger Movements using Electromyogram (EMG) Signals 2020 IEEE SENSORS, Rotterdam, Netherlands, 2020, pp. 1-4.
- [14] Sakshi Taori, Sol Lim, Use of a wearable electromyography armband to detect lift-lower tasks and classify hand loads, Applied Ergonomics, Volume 119, 2024.
- [15] Phuc Thanh Thien Nguyen, Shun Feng Su, Chung Hsien Kuo, A Frequency-Based Attention Neural Network and Subject-Adaptive Transfer Learning for sEMG Hand Gesture Classification, in IEEE Robotics and Automation Letters, vol. 9, no. 9, pp. 7835-7842, Sept, 2024.
- [16] Pana Flogeras, Dimitrios Tsipianitis and George Souliotis, Development of a Low-Cost Electromyogram (EMG) Circuit for EMG Tracking, 2023 14th International Conference on Information, Intelligence, Systems Applications (IISA), Volos, Greece, 2023, pp. 1-7.