Proc. of the 3rd International Conference on Electrical, Communication and Computer Engineering (ICECCE)
12-13 June 2021, Kuala Lumpur, Malaysia

2021 International Conference on Electrical, Communication, and Computer Engineering (ICECCE) | 978-1-6654-3897-1/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICECCE52056.2021.9514134

Classification and Feature Extraction of Different Hand Movements from EMG Signal using Machine Leaning based Algorithms

Bipasha Kundu
Electrical Engineering
University of Minnesota Duluth
Duluth,United States
kundu020@d.umn.edu

Dr. Desineni Subarram Naidu
Electrical Engineering
University of Minnesota Duluth
Duluth, United States
dsnaidu@d.umn.edu

***Abstract*—Prosthetic plays an important role for the amputees to improve the ability and mobility of their regular activities. Electromyography(EMG) has been used for decades in the control of the motorized upper-limb prosthesis. Processed EMG can imitate human movements. Mayo armband is a wireless sensor of low power, Bluetooth, and small interference which provides a good quality EMG signal. The Myo armband measures the EMG from the upper-limb. In this paper, the statistical time-domain features have been considered to classify different hand movements. The classification and comparison have been performed by 4 different Machine Learning based algorithms i.e. Support Vector Machine(SVM), Naïve Bayes(NB), Random Forest(RF), and K-Nearest Neighbor(KNN). The data has been collected from subjects (males and females) of different ages. The classifier model has used 80% data as a training set and the remaining 20% of data as the test set. The result shows that Random Forest and SVM outperform the other two algorithms with an accuracy of 98%. Referring to the accuracy here, this classification model serves as a promising candidate for the input of prosthetic hand control systems.**

Keywords*—*prosthesis, electromyography, myo armband, time domain feature extraction, machine learning, Support Vector Machine(SVM), Naïve Bayes(NB), K Nearest Neighbor(KNN), and Random Forest(RF).

I. INTRODUCTION

Upper-limb amputations are traumatic occurrences for individuals. In the United States, overall, approximately 1.7 million people or approximately 1 of every 200 people living with a limb loss [1]. According to the National Center for Health Statistics, every year 50,000 new amputation cases are added. Among them, the most common is partial hand amputations with loss of one or more fingers [2]. There are several causes for imputations. The most common causes are poor circulation because of damage or narrowing of the arteries, called peripheral arterial disease. The body's cells cannot get the oxygen and nutrients they need without adequate blood flow. The affected tissue begins to die if they don't get proper oxygen and several infections may set in [3]. The other causes are severe injuries from road accidents, war, serious burns, fireworks, cancerous tumors in the muscle or bone of the limb, infections that do not get better with treatment, and many more.

Human Bio-electric signals are extensively studied and applied in various clinical and psychophysiological researches. The Bioelectrical signal means an electrical signal obtained from any organ that exhibits a physical variable of interest. This signal is commonly a function of time and is definable in terms of its amplitude, frequency, and phase[4]. An Electromyography (EMG) signal is a biomedical signal to measure muscle responses or electrical activity produced by skeletal muscles. The nerves control the muscles by electrical signals called an impulse, these impulses can be measured and analyzed with the help of EMG sensors [5]. The attributes (i.e., amplitude and spectrum) of an EMG depend on several factors including thickness and temperature of the skin, the thickness of the fat between the muscle and the skin, the velocity of the blood flow, and location of the sensors. Factors like fatigue, aging, and neuromuscular diseases degrade muscle performance as well as EMG patterns [6].

There are two types of EMG depending on the type of sensors. One of them is the surface and another one is intramuscular. In surface EMG, non-invasive surface sensors are placed on the skin to record the electrical activity of the muscles under it [2, 3]. In intramuscular EMG, an invasive sensor (i.e., needle) is introduced into the muscle [6]. Here, we have used surface EMG.

In recent years, EMG and gesture classification has become a very popular topic. Many researchers have done research on feature extraction, classification of gestures/movements at offline and real-time, comparison of time domain, and frequency domain features, and others. According to [6], five different hand gestures were classified using KNN and dynamic wrapping algorithms. Then accuracy was compared with the accuracy of the MYO system. They found that their classification model(accuracy 86%) performs better than a MYO system(accuracy 83%). Nikita anil at [7], used a signal processing technique wavelet decomposition In [8], feature extraction and classification of EMG signals were performed. They used Principal Component Analysis(PCA) and uncorrelated linear discriminant analysis (ULDA) for feature reduction purposes and applies SVM to recognize different gestures in real-time. The authors in [9] extracted five eigenvalues in a time domain and applied Neural Network(NN) to classify six gestures. In [10], 11 movements has been considered to train Deep Neural Network and SVM. They acquired a result with 97.30% accuracy for Hand movements which was significantly greater than the Neural Network. The authors in [11] determined and compared the efficiency of different Neural network based machine learning algorithms for hand motion recognitions. They achieved an accuracy higher than 98%.

However, this paper aims to collect the EMG signal from Myo armband sensor, extract four time-domain features i.e. Mean absolute value(MAV), Root Mean Square(RMS), variance(VA), Simple Square Integral(SSI) from raw EMG signal. Finally use the extracted features to train the four different machine learning algorithms(SVM, NB, KNN, RF) to classify five different hand movements.

II. MATERIALS & DATA COLLECTION

In this section, the main features and characteristics of MYO armband and the acquisition of EMG signals have been discussed.

A. Materials

In this paper, the data have been collected from the MYO armband. The Myo armband in Fig. 1, is a gesture recognition wireless Sensor developed by Thalmic Labs. It is a wearable gesture and motion control device that uses a set of 8 sensors, combined with IMU((Inertial Measurement Unit) sensors, including gyroscope, accelerometer, and magnetometer, to recognize gestures [12].

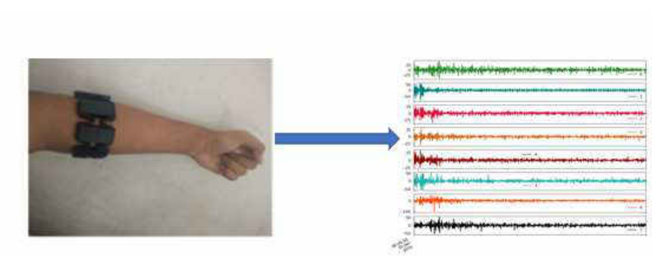

Fig. 1. The myo armband worn in hand and the signal from 8 sensors

These 8 sensors acquire the data from the forearm. The Myo armband has a signal frequency of 200Hz. This device requires the user to wear it and synchronize it with the hand movements before it can be used. The data is generated as a Bluetooth packet. Myo transmits the data through a Bluetooth Dongle connected to a PC [7]. The generated packet has two types of data. One is EMG data value and the other is IMU data values. However, only the EMG data has been used in this paper.

B. Data Collection

The EMG signal was collected from six abled-body subjects(three females/ three males, ages: 22-30 yrs.) and they had no accident history on their dominant hand. After wearing the MYO armband into their forearm, the subjects performed 5 hand movements including thumb flexion(TF), index finger down(ID), middle finger(MD), ring finger down(RD), little finger down(LD). Each movement was maintained for 5s and 1000 EMG samples were collected for every 8 sensors and each movement. The hand movement has been recorded using an acquisition software developed in python [11]. . Later, the EMG signal was converted into a CSV file format for feature extraction and feeding to the Machine Learning Algorithm for classification and recognition.

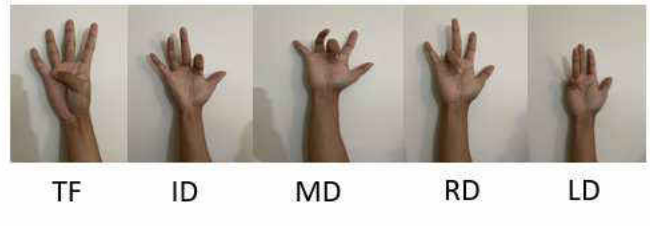

Fig. 2. Five Hand Movements

III. METHODS

This section describes the feature extraction of the raw EMG signal and a brief about the classification model used for training the data and recognition. The flow chart of this study has been shown in Fig. 3.

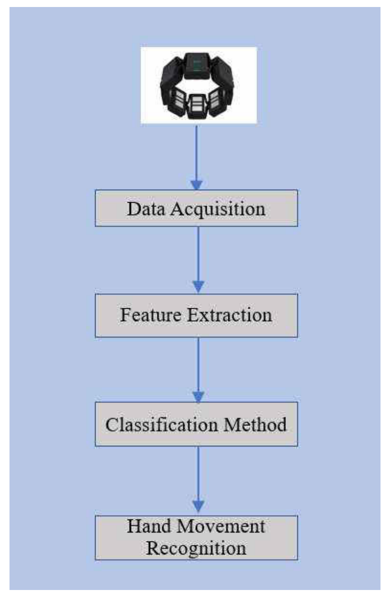

Fig. 3. Work flow chart

A. Feature Extraction

Feature extraction is the method of taking the features from the EMG signal. Mean absolute value(MAV), Root means square(RMS), Variance(VA), Simple Square Integral(SSI) has been analyzed for each movement. They are called time-domain(TD) features as the EMG signal is represented in time. The value of this feature extraction will be the input for the classification model.

a)Mean Absolute Value(MAV): MAV is defined as the average of absolute of the EMG signal It is calculated as follows

$$\mathrm{MAV} = \frac{1}{N} \sum_{k=1}^{N} |X_k| \quad (1)$$

Where N is the total length of the Signal and X_k represents the EMG Signal.

b) Root Mean Square(RMS): RMS defines the square root of the average power of the EMG signal. It is related to the constant force and non-fatiguing contraction. It is expressed as follows

$$\text{RMS} = \sqrt{\frac{1}{N}} \sum_{k=1}^{N} X_k^2 \quad (2)$$

c) Variance: EMG Signal Variance is the measure of the power density of the signal. The value of Emg variance can be zero because EMG signals are handed based on a white Gaussian Noise. It can be calulaed as

$$\text{VAR} = \frac{1}{N} \sum_{k=1}^{N} (X_k - \bar{X})^2 \quad (3)$$

d) Simple Square Integral(SSI): SSI gives a measure of the energy of the EMG signal. It is defined as follows

$$\text{SSI} = \sum_{k=1}^{N} (|X_k^2|) \quad (4)$$

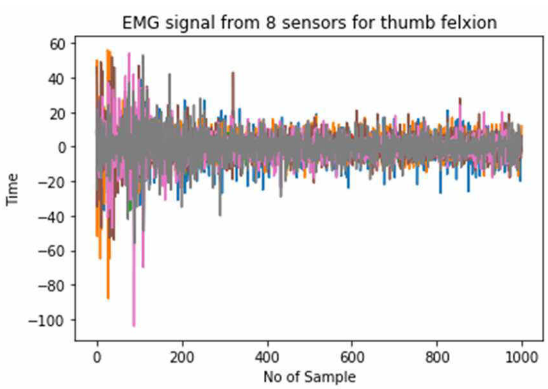

Fig. 3. EMG signal amplitude for one movement

B. Classification Models

1) Support Vector Machine: SVM is a supervised machine learning(ML) algorithm that analyzes data for classification, regression, and sometimes outlier detection. Each data item in SVM is plotted as a point in a m dimensional space(m is denoted as the total number of features). A particular co-ordinate uses each feature's value. Then SVM finds a hyperplane that works as a decision boundary to recognize two different classes.

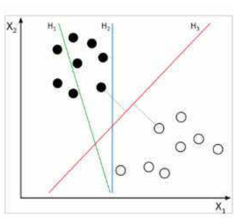

Fig. 4. SVM Classification [7]

Here in Fig. 4, the black and white circles are two different classes and H3 refers to the decision boundary in the co-ordinate.

1) Naïve Bayes: NB is one of the classification models of supervised ML algorithms. It uses Bayes' theorem with the "naive" assumption, which means features are independent of one another. NB simulates that a certain feature is supreme to any other class present in that data. Therefore, Bayes' theorem can be defined as

$$P_{(y|x_1,\ldots\ldots,x_n)} = \frac{(P_y)P_{(x_1,\ldots\ldots,x_n|y)}}{P_{(x_1,\ldots\ldots x_n)}} \quad (5)$$

Where y is the class variable and x is the feature vector of the data.

2) K-Nearest Neighbor: KNN is another supervised ML technique. It depends on the features of its neighboring data point to classify a new data point. It classifies a new class by the majority polling of its k neighbors. KNN is familiar as a non-parametric technique because of its of not making any assumptions of data points. This makes KNN more effective compared to other models. It is one of the simplest techniques among all the classification model. It is used for both classification and regression problems.

3) Random Forest: RF is another flexible supervised ML technique. It produces a great accuracy without any hyper-parameter tunning. It's consisted of a large number of decision trees that work as an ensemble. It is trained with the "bagging" method. The bagging method implies that the combination of learning models improves the overall result.

IV. Result Analysis

The sample raw EMG signal from 8 sensors has been shown in Fig.3. The figure illustrates that the amplitude is different for each sensor and it lies in between 1-100 mV. Similarly, amplitude and shape are different for different movements and subjects. In this study, the raw EMG signal has been used for feature extraction. The small interference and Bluetooth of Myo provides a good quality of the signal. So the feature has been extracted from the raw EMG signal. The time-domain features extracted here are MAV, RMS, VA, and SSI. We used a sliding window of step 5s for each of the 8 sensors. Therefore, for 1000 samples, we get 200 data points for each hand movement and 32features, Which are then fed into four different algorithms for classification. The classifier model used 80% of the data as the training set and 20% of the data as the test set. After passing these processed data, the following average accuracy was achieved. In Table 1, the comparison of average accuracy for each feature and classification technique has been shown.

TABLE 1 COMPARISON OF THE PERCENTAGE OF AVERAGE ACCURACY OF 5 HAND MOVEMENTS

Algorithm	**Features**			
	MAV	***RMS***	***VA***	***SSI***
SVM	98%	97%	76%	94%
NB	89%	94%	92%	88%
KNN	95%	94%	89%	87%
RF	96%	98%	94%	97%

From Table 1, shows the average accuracy for all movements. The classifier output shows that SVM and RF have achieved the highest accuracy 98% whereas the NB and KNN have a lower performance. They have an accuracy of 94% and 95%. A bar chart of the accuracy has been shown in Fig. 5.

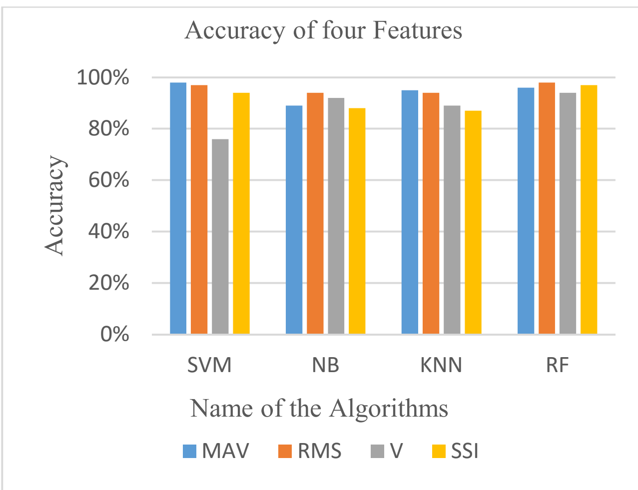

Fig. 5. Classification accuracy

Fig. 5 illustrates that SVM obtains the highest accuracy of 98% for MAV, NB has the highest accuracy of 97% for RMS and RF has the accuracy of 98% RMS. However, KNN has poor performance compared to other models. It has an accuracy of 95%. Among all features which have been used as the input for classification, MAV and RMS have provided a better accuracy.

The F-score, also called the F1-score, is a measure of a model's accuracy on a dataset i.e. measure of test accuracy. The F1-score for each movements and each features have been shown on Fig. 6 to Fig. 9.

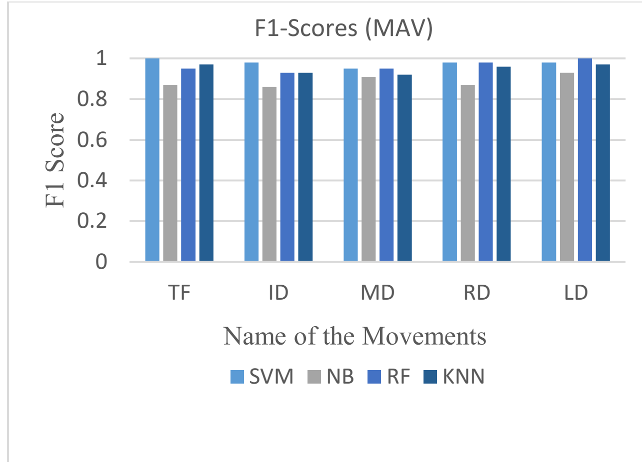

Fig. 6. F1-Score for MAV

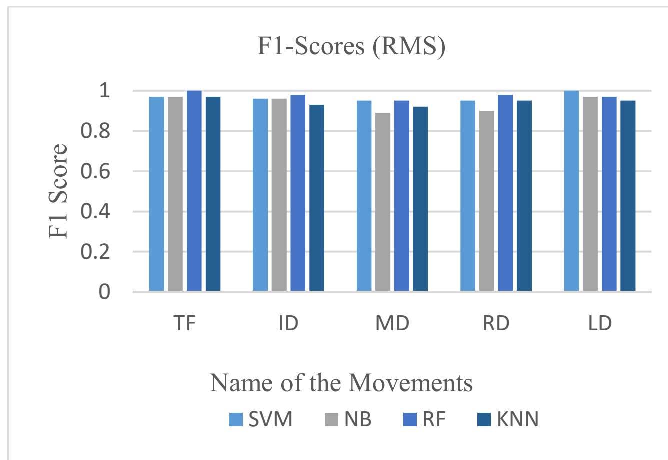

Fig. 7. F1-Score for RMS

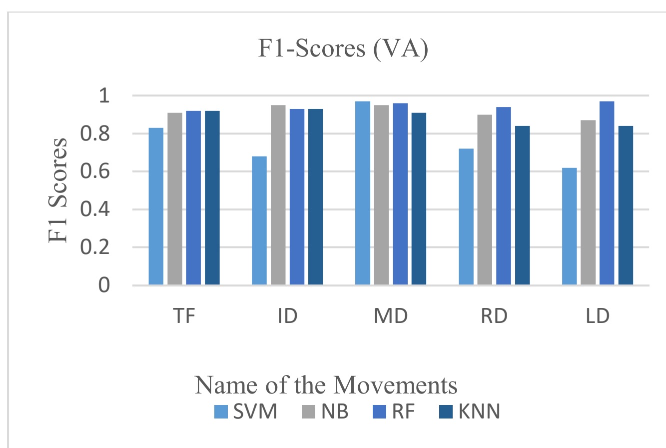

Fig. 8. F1-Score for VA

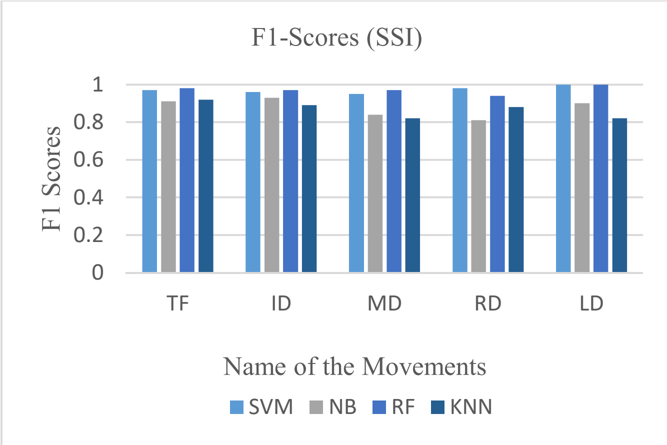

Fig. 9. F1-Score for SSI

From Fig. 6 to Fig. 9, it is noticed that TF, RD, and LD have a better F1 score for MAV and RMS in comparison to other two movements. Among all the movements, TF has the best F1 scores compared to the rest of the algorithms. Overall, referring to the analysis and results, these classification model can be the best approach as the input for the prosthetic hand.

V. Conclusion

This work aims to classify five different hand movements from the EMG signal which has been collected from the human body. A comparison of average accuracy has been shown for five movements. This offline classification uses the raw EMG signal to extract the time-domain features. Among the four features, MAV and RMS outperform VA, SSI. The SVM and RF perform best in terms of accuracy. The other two algorithms' performance is also noticeable but a bit lower than these two. For future work, this classification model and result can be used for real-time movement classification. This result will be a promising candidate for real-time classification of hand movements.

Acknowledgment

The first author would like to acknowledge her advisor Dr. Desineni Subbaram Naidu, for his patience and encouragement to work with Machine Learning and the people who voluntarily gave the EMG data for this work. Also thanks are due to EE department at University of Minnesota, Duluth (UMD) for supporting the first author's graduate study.

References

[1] Diane W. Braza, Jennifer N. Yacub Martin, "Upper Limb Amputations," Essentials of Physical Medicine and Rehabilitation(Fourth Edition),Elsevier,2020,page 651-657

[2] "https://u.osu.edu/wheelbarrows/upper-limb-injury-statistics/#"

[3] "https://www.webmd.com/a-to-z-guides/definition-amputation#1.

[4] Wan Mohd Bukhari, Wan Daud, Abu Bakar Yahya, Chong Shin Horng, Mohamad Fani Sulaima, Rubita Sudirman, "Features Extraction of Electromyography Signals in Time Domain on Biceps Brachii Muscle", International Journal of Modeling and Optimization, Vol. 3, No. 6, December 2013.

[5] Tony Chau "Machine Learning for Gesture Recognition with Electromyography,"Norwegian University of Science and Technology,Dept. of CS, June 2017,unpublished.

[6] Marco E. Benalcazar, Andres G. Jaramillo, Jonathan A. Zea, and Andres Paez, "Hand Gesture Recognition Using Machine Learning and the Myo Armband", 25th European Signal Processing Conference (EUSIPCO), 2017.

[7] Nikitha Anil, Sreeletha S.H "EMG based Gesture Recognition using Machine Learning," International Conference on Intelligent Computing and Control Systems (ICICCS 2018).

[8] Sachin Negi, Yatindra Kumar, V. M. Mishra,"Feature Extraction and Classification for EMG Signals using Linear Discriminant Analysis ," IEEE Trans, 2016.

[9] Shunzhan He, Chenguang Yang, Min Wang, Long Cheng, Zedong Hu, "Hand Gesture Recognition using MYO Armband", IEEE Trans, 2017.

[10] S Mary Vasanthi, T Jayasree, "Performance evaluation of pattern recognition networks using electromyography signal and time-domain features for the classification of hand gestures", Sage Journals, Vol. 234, issue 6, page 639-648, March 23, 2020.

[11] Yang Zhou, Chaoyang Chen, Juan Ni, Guoxin Ni, Min Li, Guanghua Xu, John Cavanaugh, Mark Cheng, Stephen Lemos, "EMG Signal Processing for Hand Motion Pattern Recognition Using Machine Learning Algorithms", Arch Orthop. 2020.

[12] Thalmic. Myo Armband. https://www.myo.com/.

[13] "https://github.com/NiklasRosenstein/myo-python/tree/master/myo"