Engineering and Technology Journal
Journal homepage: engtechjournal.org

Pattern Recognition of Composite Motions based on EMG Signal via Machine Learning

Noof T. Mahmood[a]*, Mahmuod H. Al-Muifraje [b], Sameer K.Salih[c], Thamir R. Saeed [d]

[a] Ph.D. Student, Electrical Eng. Dept. University of Technology, Baghdad, Iraq, noofthabit@gmail.com

[b] Associate Professor, Electrical Eng. Dept. University of Technology, Baghdad, Iraq, drmahmood6@gmail.com

[c] Doctorate, Ministry of Sciences and Technology, Baghdad, Iraq, sameerksalih@yahoo.com

[d] Professor, Electrical Eng. Dept. University of Technology, Baghdad, Iraq, 50257@uotechnology.edu.iq

*Corresponding author.

Submitted: 20/06/2020 Accepted: 10/08/2020 Published: 25/02/2021

K E Y W O R D S

Electromyography, Principle Component Analysis, Support Vector Machine, and K-Nearest Neighbors.

A B S T R A C T

In the past few years, physical therapy plays a crucial role during rehabilitation. Numerous efforts are made to demonstrate the effectiveness of medical/ clinical and human-machine interface (HMI) applications. One of the most common control methods is using electromyography (EMG) signals generated by muscle contractions to implement the prosthetic human body parts. This paper presents an EMG signal classification system based on the EMG signal. The data is collected from biceps and triceps muscles for six different motions, i.e., bowing, clapping, handshaking, hugging, jumping, and running using a Myo armband with eight electromyography sensors. The Root Mean Square, Difference Absolute Standard Deviation Value, and Principle Component Analysis are used to extract the raw signal data and enhance classification accuracy. The machine learning method is applied, i.e., Support Vector Machine and K-Nearest Neighbors are used for classification; the results show that the K-Nearest Neighbors method achieves a higher accuracy percentage than the SVM. Making high training accuracy for different physical actions helps implement human prosthetic parts to help the people who suffer from an amputee.

How to cite this article: : N. T. Mahmood, M. H. Al-Muifraje, S.K.Salih, T. R. Saeed, "Pattern Recognition of Composite Motions based on EMG Signal via Machine Learning," Engineering and Technology Journal, Vol. 39, Part A, No. 02, pp. 295-305, 2021.
DOI: https://doi.org/10.30684/etj.v39i2A.1743

1. Introduction

Physical therapy plays a significant role in the rehabilitation stage. So, many efforts are made to elucidate the leverage of medical/ clinical and human-machine interface (HMI) applications [1]. The dominant control methods are using electromyography (EMG) signals generated by muscle constriction to implement the prosthetic human body parts [2]. Therefore, robotic has the effectiveness to increase the independence of the individuals living lifestyle with their disabilities. The objective is to improve the life quality that happens by empowering people to achieve a wide range of daily responsibilities within a few times. Humanoid robots can be prepared robotic limbs to make the physical actions of individuals and the intelligent robotics industry. Nowadays, the most widely used are robotic hands and arms. The robotic hands would have the ability to achieve the primary skills like the transfer from one place to another and grasp of objects in the same nonamputees individuals do. Therefore, depending on the object shapes, the human hand actions must be trained and applied to the designed robotics hand. The intuitive approach creates an interface that brings the activity of the muscle and records it using EMG sensors [3].

EMG signals are a way for estimating the electrical signal that relatives to skeletal muscle; the electrical signal consists of some motor unit action potentials (MUAPs) [4]. The combination of muscle fiber action potentials from all the muscle fibers of a single motor unit is called the MUAP [5].

The pattern recognition (PR) system based on electromyography signals is divided into three main parts: data collection, feature extraction, and classification of the input data. There are many methods to collect the electromyography signals from the muscles; the first method by collecting data from the skin surface using different sensors such as the Myo armband, which is considered a commercial way. The second method is collecting the data using a needle that is regarded as a medical method and rarely used in research [6,7].

The second stage is extracting the features. Several techniques are applied in frequency domain analysis, time-domain analysis, and time-frequency domain analysis to acquire significant information from EMG signals. The most common domain used in EMG signals is the time domain, which is represented by mean absolute value (MAV), zero crossings (ZC), nth-order autoregressive (AR), mean, log detector, variance (VAR), wavelength (WL), root-mean-square value (RMS), simple square integral (SSI), the sign of slope changes (SS), modified mean absolute value (MMAV), EMG integral (IEMG), average amplitude change (AAC), histogram, Willison amplitude and sample entropy [8].

The final stage in PR is the classification to specify to which class the extracted features fit. The most common machine learning classifiers are Support Vector Machine (SVM), decision tree, K-Nearest Neighbor, Hidden Markov (HM), and Linear discriminant analysis (LDA) [9]. Currently, the prosthetic technique is an artificial part of resolving the problem of amputees' parts. The appearance of the consumer-level 3D printers supports the prosthetics implementation. Thus, the developed prosthetics have gained popularity over traditional prosthetics, which are expensive and unaffordable to many users [10]. The main challenge from the mechanical standpoint is how to mix many degrees of freedom (DOF). As a result, the DOF increase is caused by decreasing in the gripping force, which leads to an unstable system. Also, increasing DOFs will lead to a rise in actuators number and high failure probability so that the prosthetic hand will need high maintenance and difficult manufacture [11].

2. METHODOLOGY

This work aims to design a system for pattern recognition (PR) of the EMG signal. The EMG signal data will collect from different subjects using the Myo armband that contains eight EMG sensors distributed in 360o. The data is collected by the Myo armband will be transferred to the PC using Bluetooth. The data is classified into six different motions using three feature techniques. This work used two classifiers; the support vector machine (SVM) and k- nearest neighbors (KNN). Figure (1) shows the block diagram of the proposed system.

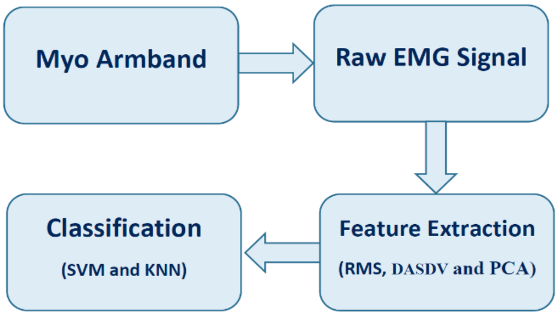

Figure 1: The block diagram of the proposed system

I. Data Collection

Myo armband is used to record six regular motion classes: bowing, clapping, handshaking, hugging, jumping, and running. The eight sensors of the Myo armband are corresponding to eight inputs of time series with a voltage amplitude of the required muscles, and each time series consist of 9000 samples. The eight sensors cover the following body part; right arm, left arm, right leg, and left leg, taking into consideration two muscles; the biceps and the triceps. Figure (2) shows the raw EMG signals for Bowing, Clapping, and Handshaking motions. Figure (3) shows the raw EMG signals for Hugging, Jumping, and Running motions.

II. Feature Extraction

The main idea behind the feature extraction is to avoid extensive data, for each motion will take a long time for signal processing. Therefore, many types of feature extraction are used to reduce the raw data dimensions and produce new vectors that will enter the classification stage instead of the raw data. So, the new vectors must contain all the required information to obtain fast training [12].

In this work, the EMG signals are analyzed in offline mode using Matlab 2019a. Then the recoded signals are segmented in a window size of 200 ms and an increment of 150 ms. The window is analyzed using the Root Mean Square (RMS), Difference Absolute Standard Deviation Value (DASDV), and principal components analysis (PCA) as features.

A. Root Mean Square (RMS)

RMS is a common feature that used to analyze the EMG signal, the RMS modeled as amplitude modulated Gaussian random process whose relates to constant force and non-fatiguing contraction equation is given as [13]:

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

Where; X: input RMS values, N: Number of raw input signals. This equation is used to select a group of raw EMG signals based on the root mean square calculations.

B. Difference Absolute Standard Deviation Value (DASDV)

DASDV is similar to the RMS feature, which is calculated the standard deviation of wavelength; the DASDV equation is given as [14]:

$$DASDV = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N-1}(X_{i+1} + X_i)^2} \tag{2}$$

Where; X: input RMS values, N: Number of raw input signals. This equation is used to select a group of raw EMG signals based on DASDV calculations.

C. Principle Component Analysis (PCA)

The PCA is an exploration method that presents transformation data using a statistical process; these transformations are groups of observations that convert to orthogonally related values. The PCA is started by taking the mean of matrix input data, taking the covariance, finding the eigenvalue and the corresponding eigenvector. Then, choosing the remarkable value eigenvector and multiplying the transposed matrix obtained by eigenvector to the mean-adjusted data transposed [15] [16]. The PCA is used for dimension reduction because of reduction in time and space complexities, the dimensions of the new components will be uncorrelated and orthogonal to each other. The reduction happens when PCA select the maximum variance and form new directions; figure (4) shows how the axis transform from the old axis to a new axis. From the figure, the old perpendicular black axis has a scattering plot that shows the set of data, while the green axes are the new PCA that are chosen depending on the maximum variance [17]. Also, the number of these components is less or equal to the original one and will never be more than the original data [18].

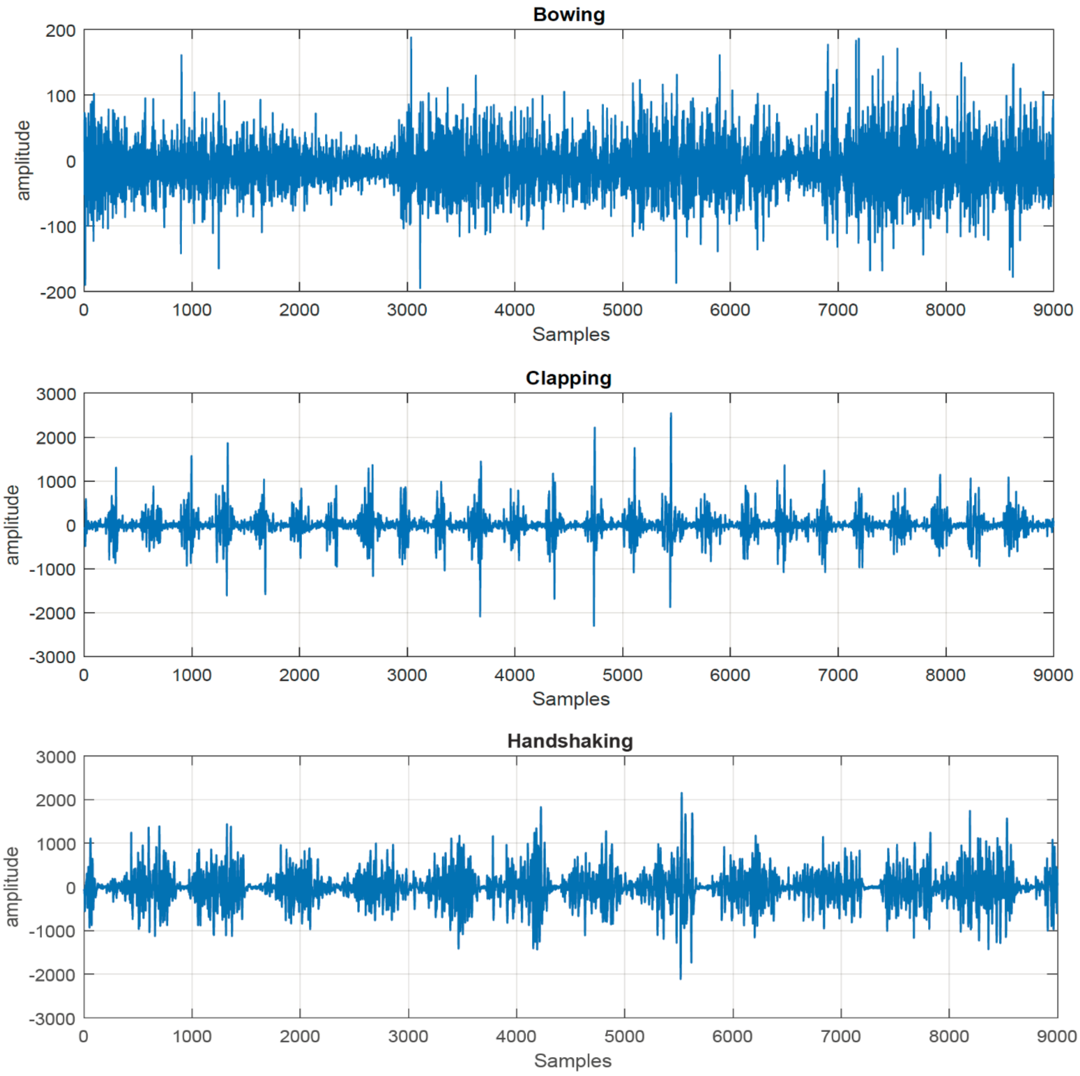

Figure 2: The raw EMG signals for Bowing, Clapping, and Handshaking motions.

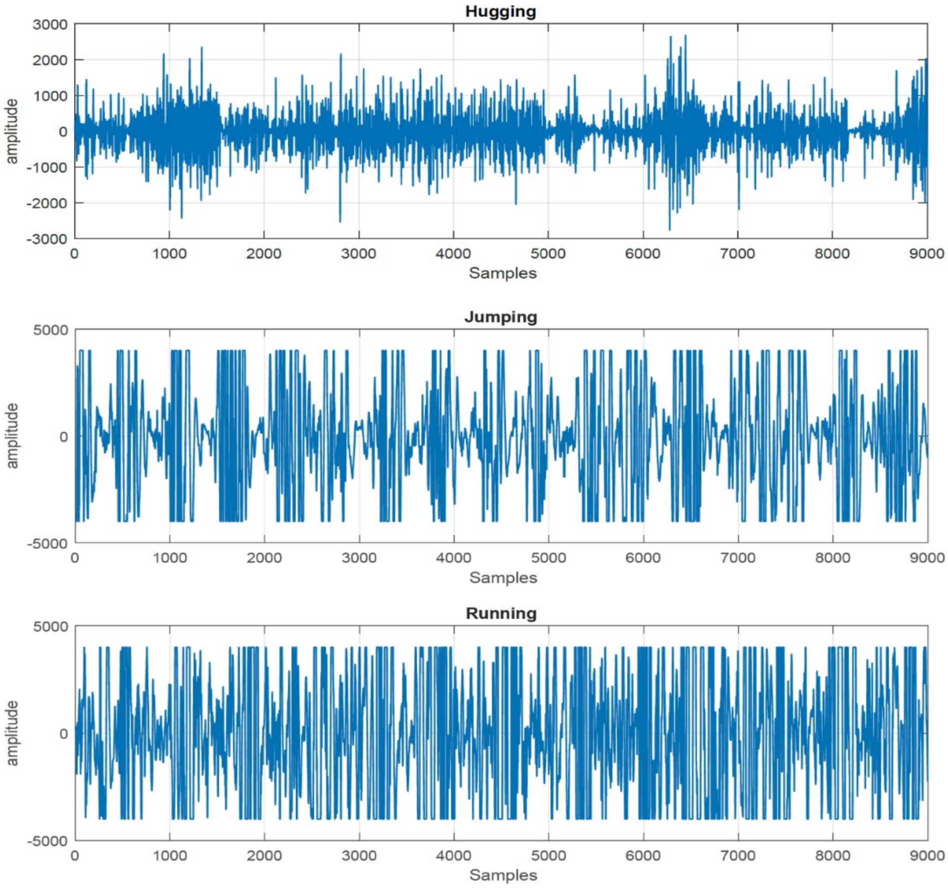

Figure 3: The six raw EMG signals Hugging, Jumping and Running motions.

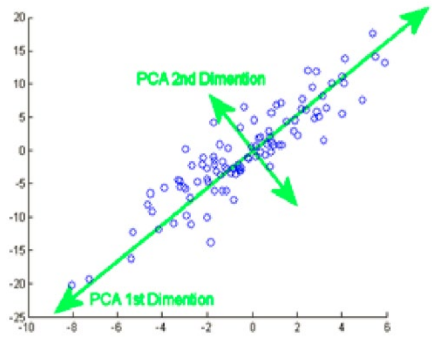

Figure 4: PCA axis transformation [18]

III. Classification Methods

The classifier's primary role is to take the features extracted from the raw data and generate a task to use an extracted feature and create the typical classes matching the required action [12].

In this work, the extracted features fed to the classifier using Matlab 2019a to detect the correct class motions. The data is divided for the training part and testing part, 80% of the data is for training, and 20% of data is for testing by using holdout validation. Then, the support vector machine and KNN have been used here as classifiers:

A. Support Vector Machine (SVM):

SVM is a supervised ML algorithm that is used for regression or classification issues. The SVM builds a hyperplane or groups of hyperplanes with infinite dimensionality space. Also, SVM Uses a subgroup of training (support vectors), and diverse Kernel functions can be particular for choosing the support vectors. The classifier used in medical diagnosis, also SVM, is useful for low memory. SVM deals with multidimensional space that takes the required class and divides it using the best-optimized method [19]. The cubic SVM is a type of SVM, and the equation of cubic SVM can be given as (3):

$$k(x_i, x_j) = (x_i^T, x_j)^3 \quad (3)$$

The cubic SVM is preferred due to the short time required for training in this work training time range from 45 sec to 50 sec.

B. K-Nearest-Neighbor's (KNN)

KNN is supervised learning a secure method but has a useful and accurate classification [20]. In the KNN method, the data represented in the vector space. The k represents a variable and selects as an essential factor, and h represents the least distance from the selected point if k is equal to h, which means unselected point. In this work, the Euclidean distance is used to measure the distance between points using the given equation (4):

$$D_i = \sum_{k=1}^{h} \sqrt{(x_i - x_k)^2 + (y_i - y_k)^2} \quad (4)$$

Weighted Distance Nearest Neighbor (WDNN) Classifier is a type of KNN [21] based on recalling the revealing cases and learning their weights for classification. Also, this algorithm is designed to classify the data without using regression [22]. Fine KNN Classifier is another type of KNN that makes it finely itemized to distinguish among different classes with the number of neighbors set to 1[23].

3. Results

In this work, three participants have been taken to collect the raw EMG signals by applying RMS, DASDV, and PCA to extract the features. The SVM and KNN methods are used to classify the motions. The performance of the classifier model is described using the confusion matrix. This matrix clarifies the relationship between the predictive and the actual events. Table I shows the confusion matrix model.

TABLE I: confusion matrix model

Confusion Matrix	Predicted No	Predicted Yes
Actual No	TN	FN
Actual Yes	FP	TP

Where: (TN) is True Negative, (FN) is False Negative, (FP) is False Positive, (TP) is True Positive. The accuracy of the confusion matrix can determine from the following equation:

$$Accuracy = \frac{TN+TP}{TN+TP+FN+FP} \quad (5)$$

The experimental results are divided into three parts. In the first part of experiments applied by using one subject for 6 different motions, the accuracy reaches 67.5%, 78.5%, and 87.9% for cubic SVM, weighted KNN, and Fine KNN, respectively. Figure (5) shows the confusion matrix for one subject.

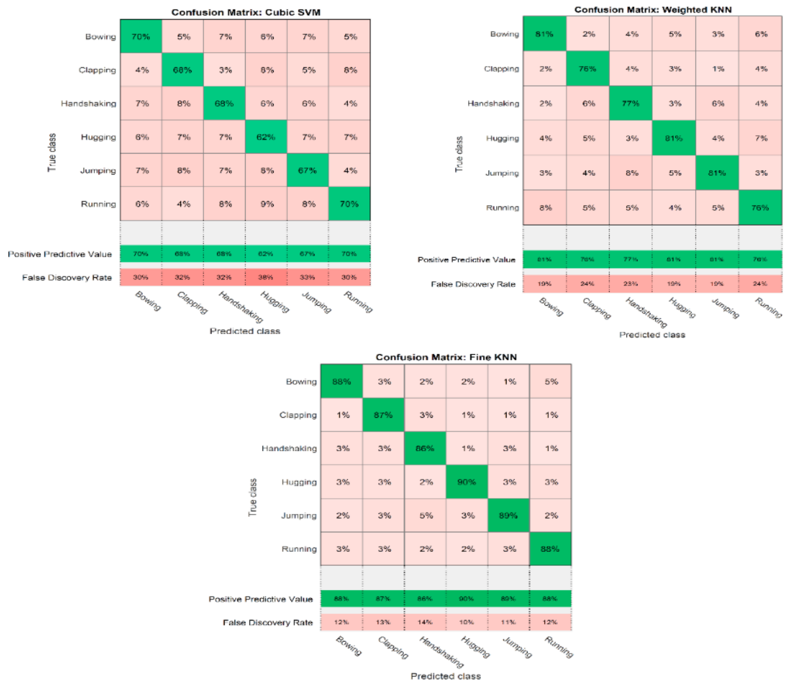

Figure 5: The confusion matrix for one subject.

The second part of experiments applied by using two subjects for 12 different motions, the accuracy reaches 92.1%, 93.7%, and 98.3% for cubic SVM, weighted KNN, and Fine KNN, respectively. Figure (6) shows the confusion matrix for the two subjects.

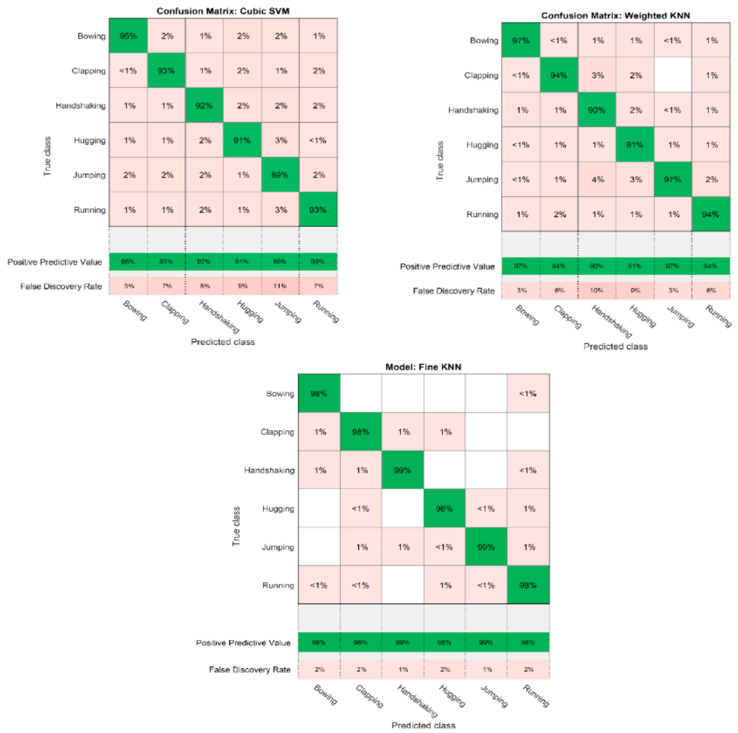

Figure 6: The confusion matrix for two subjects.

The third part of experiments applied by using three subjects for 18 different motions, the accuracy reaches 94.8%, 95.7%, and 98.9% for cubic SVM, weighted KNN, and Fine KNN, respectively. Figure (7) shows the confusion matrix for the three subjects.

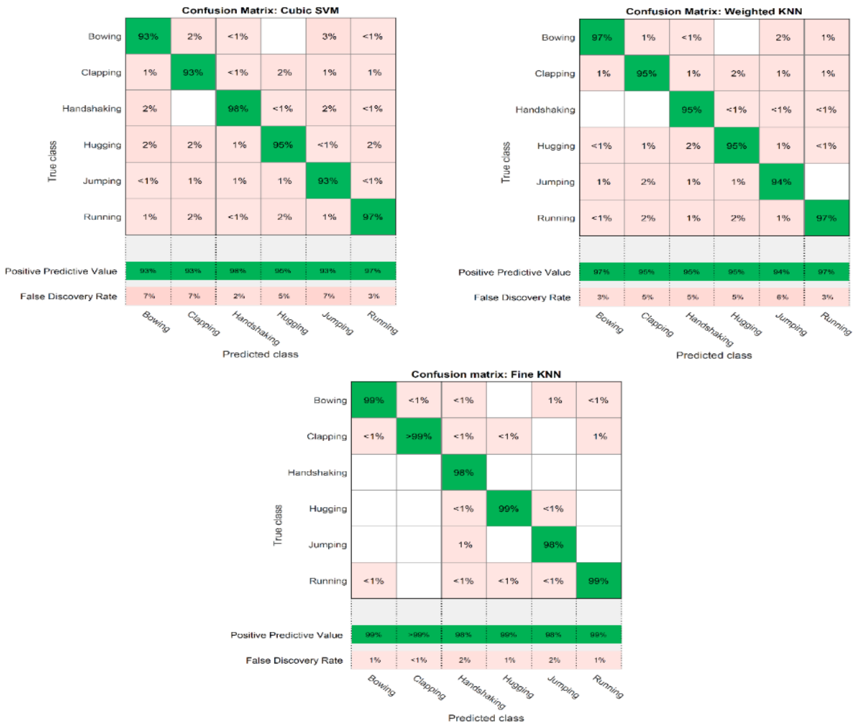

Figure 7: The confusion matrix for three subjects.

For the third part of the experiments, the scattering plot is applied to the classified data to observe the relationship between the variables. Where the dot represents the correct prediction, and x represents the incorrect prediction. Figure (8) shows the scattering plot for three subjects using cubic SVM, weighted KNN, and Fine KNN.

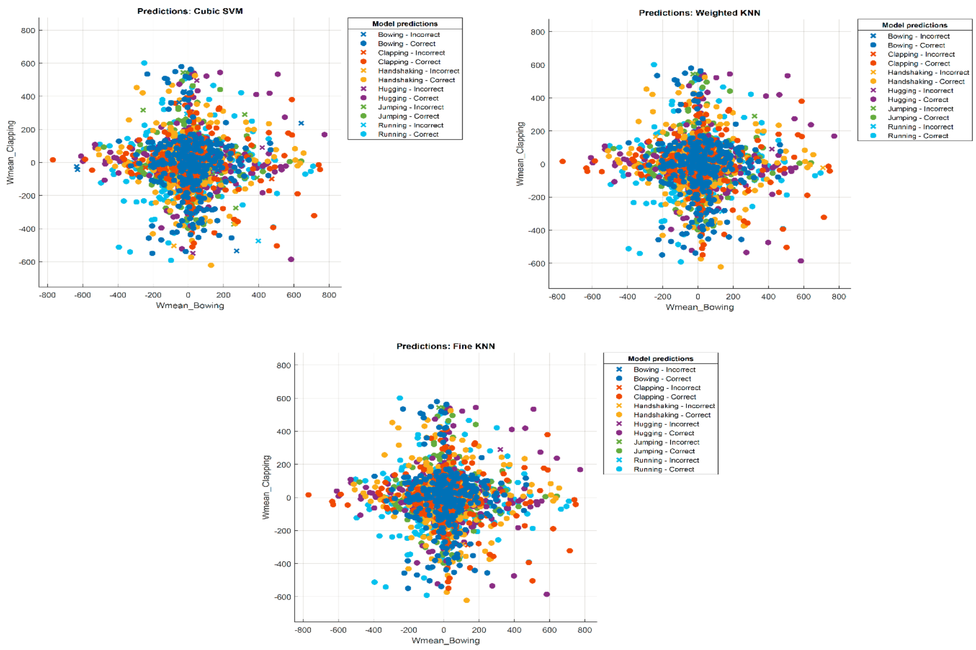

Figure 8: The scattering plot for three subjects using cubic SVM, weighted KNN, Fine KNN.

The third experiment applies the area under the curve (AUC) and Receiver Operating Characteristics (ROC) curve to know the ability of the model to distinguish among classes. Figure (9) shows the ROC for three subjects.

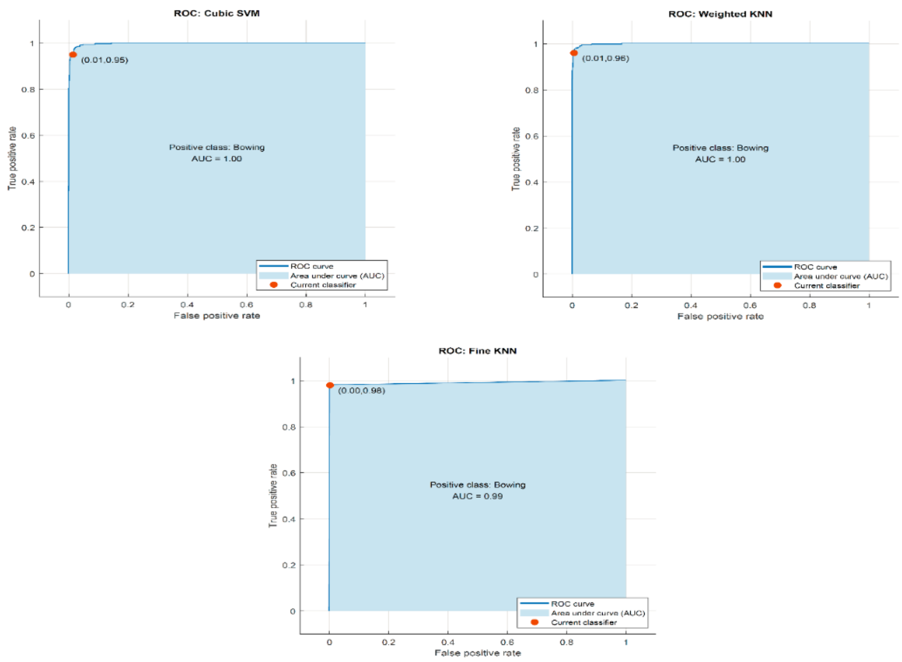

Figure 9: The ROC for three subjects.

4. DISCUSSION

Myo Armband is used in this work instead of the traditional sensors to solve noise issues and avoid the DC component in the raw signals. However, the data was collected for six movements, and there is no need to use any filter. The time-domain analysis is used instead of frequency domain analysis to extract the raw EMG signals' features because it shows better performance in classifying the EMG signals.

The Machine Learning (ML) algorithm is used in this work and shows fast training time; the average training time reaches 45.5 sec. The obtained classification result shows that both the fine KNN and the weighted KNN are better than the SVM.

From all the three experiments, as noticed from the results, the first experiment has the lowest accuracies in all classifiers due to the small raw data. When the data increase in the second experiment to use two subjects, the efficiencies also increase. The results achieve the highest accuracies when the data of the three subjects are used. In evidence, the increasing number of subjects means increasing in data that will be trained and will lead to increase inaccuracies. Figure (10) shows the classification performance using different numbers of subjects.

In the third experiment, due to the high performance, the scattering plot shows perfect prediction; as noticed in figure (8), the X point representing the incorrect prediction is very low.

Another proof that the third experiment is the best the figure (9) shows the AUC of 1.00 that clarifies the probability of the random positive is the highest among the random negative and this good indication of the high accuracy.

The results of this work have been compared with three previous works, as shown in Table (II). The first one takes only RMS and the SVM and achieved good accuracy. The second work chooses the RMS, and Integrated Absolute Value (IAV) as features also selects the KNN and Native Bayesian Pattern (NBP) as a classifier, and the authors achieved accuracy lower than the accuracy archived in this work. The last study used more than three features and used linear regression and SVM as a classifier and achieved lower accuracy than this work achieved.

TABLE II: compare this work with pervious works.

Author	Year	Features	Movements	Classifier	Accuracy
McCool, et al. [24]	2014	RMS	biceps brachii, quadriceps femoris, tibialis anterior.	SVM	95.75%
Praveen, et al.[12]	2018	RMS and IAV	Hand grasp (Opposition) and release (Re-position)	KNN and NBP	92%, 94%.
Khan, et al. [9]	2019	mean, Kurtosis, Peak to Peak, Shape Factor, energy, jitter, ZC, and Spectral Spread.	Healthy and patients	LR and SVM	SVM of 94.1%, LR of 95.1%
This work	2020	RMS, DASDV, and PCA	Left Hand Clip, Left Hand Down, Left Hand Up, Right Hand Clip, Right Hand Down, Right Hand Up.	KNN and SVM	SVM of 94.8%, weighted KNN of 95.7%, and Fine KNN of 98.9%.

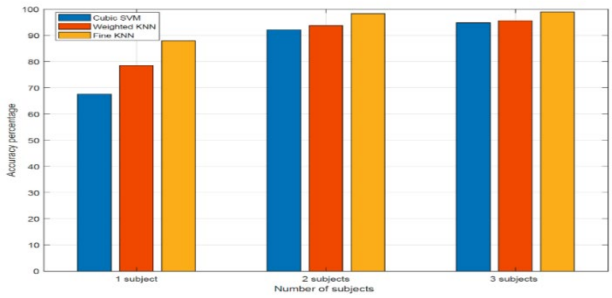

Figure 10: The classification performance using different numbers of subjects.

5. CONCLUSIONS

In this work, the proposed classification technique using machine learning (ML) shows a high speed of training with composite data. Also, the number of subjects (i.e., number of inputs to classifier) has a significant effect on increasing the accuracy. Therefore, the third experiment that uses three subjects shows the highest accuracy among other experiments. As noticed from the result, both Fine KNN and weighted achieve higher accuracies reach 98.9% and 95.7%. Meanwhile, the SVM achieves an accuracy of 94.8%. The high training accuracy of this work is useful to implement a genuine prosthetic human body part.

References

[1] M. Z. Jamal, Dong-Hyun Lee, Dong Jin Hyun, Real Time Adaptive Filter based EMG Signal Processing and Instrumentation Scheme for Robust Signal Acquisition Using Dry EMG Electrodes, 16th International Conference on Ubiquitous Robots (UR). IEEE, 2019, 18851324. 10.1109/URAI.2019.8768662

[2] V. P. Maurya, P. Kumar, S. Halder, Optimisation and Classification of EMG signal using PSO-ANN , 2019 Devices for Integrated Circuit (DevIC). IEEE, 2019, 18934718. 10.1109/DEVIC.2019.8783882

[3] O, Mohammadreza Asghari, H. Hu, Myoelectric control systems—A survey, Biomed. Signal Process., 2 (2007) 275-294. doi.org/10.1016/j.bspc.2007.07.009

[4] S. Rissanen, M. Kankaanpää, M. P. Tarvainen, Extraction of typical features from surface EMG signals in Parkinson's disease, 11th International Congress of Parkinson's Disease and Movement Disorders, Istanbul, Turkey. 2007.

[5] N. Sengar, M. K. Dutta, C. M. Travieso, Identification of amyotrophic lateral sclerosis using EMG signals, 4th IEEE Uttar Pradesh Section International Conference on Electrical, Computer and Electronics (UPCON). IEEE, 2017, 17471834 . 10.1109/UPCON.2017.8251093

[6] M. E. Benalcázar, A. G. Jaramillo, Jonathan, A. Zea, A. Páez, Hand gesture recognition using machine learning and the Myo armband, 25th European Signal Processing Conference (EUSIPCO). IEEE, 2017, 17290443. 10.23919/EUSIPCO.2017.8081366

[7] H. Huang, et al. , EMG pattern recognition using decomposition techniques for constructing multiclass classifiers, 6th IEEE International Conference on Biomedical Robotics and Biomechatronics (BioRob). IEEE, 2016 , 16192630. 10.1109/BIOROB.2016.7523810

[8] N. Nazmi, M. A. Abdul Rahman , S.Yamamoto , S. A. Ahmad , H. Zamzuri , S. A. Mazlan, A review of classification techniques of EMG signals during isotonic and isometric contractions, Sensors, 16 (2016). 10.3390/s16081304

[9] Khan, Muhammad Umar, et al. ,Classification of EMG Signals for Assessment of Neuromuscular Disorder using Empirical Mode Decomposition and Logistic Regression, International Conference on Applied and Engineering Mathematics (ICAEM). IEEE, 2019 , 19029859. 10.1109/ICAEM.2019.8853684

[10] N. Amancherla, A. Turlapaty, B. Gokaraju, SVM based Classification Of sEMG Signals using Time Domain Features for the Applications towards Arm Exoskeletons, Proceedings of the Advances in Robotics , 2019, 1-5. doi.org/10.1145/3352593.3352620

[11] K. Y. Choi, A. Akhtar, T. Bretl, A compliant four-bar linkage mechanism that makes the fingers of a prosthetic hand more impact resistant, 2017 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2017, 17057750. 10.1109/ICRA.2017.7989791

[12] L. S. Praveen, S. N. Nagananda, P. Shankapal, Design and Development of Real Time Bionic Hand Control Using EMG Signal, 2018 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), 2018, 18162760. 10.1109/CONECCT.2018.8482393

[13] S. R. Kurniawan, D. Pamungkas, MYO Armband sensors and Neural Network Algorithm for Controlling Hand Robot, 2018 International Conference on Applied Engineering (ICAE). IEEE, 2018, 18326507 . 10.1109/INCAE.2018.8579153

[14] Caesarendra, Wahyu, et al. , A classification method of hand EMG signals based on principal component analysis and artificial neural network, 2016 international conference on instrumentation, control and automation (ICA). IEEE, 2016 , 16585836. 10.1109/ICA.2016.7811469

[15] H. Abdi, L. J. Williams, Principal component analysis, Wiley Interdiscip. Rev.: Comput. Stat., 2 (2010) 433-459. https://doi.org/10.1002/wics.101

[16] J. Ryu, D. H. Kim, Real-time gait subphase detection using an EMG signal graph matching (ESGM) algorithm based on EMG signals, Expert Syst. Appl., 85 (2017) 357-365. doi.org/10.1016/j.eswa.2017.05.006

[17] Y. Narayan, R. M. Singh, L. Mathew, Surface EMG Signal Classification Using Ensemble Algorithm, Adv. Eng. Inform.,(2018) 218-230. https://doi.org/10.1007/978-981-13-3140-4_20

[18] M. F. I. Ibrahim, A. A. Al-Jumaily, PCA indexing based feature learning and feature selection," 2016 8th Cairo International Biomedical Engineering Conference (CIBEC), 2016, 16623015. 10.1109/CIBEC.2016.7836122

[19] Lin, Yi-Lin, G. Wei, Speech emotion recognition based on HMM and SVM, 2005 international conference on machine learning and cybernetics, 8, 2005 , 8705055 . 10.1109/ICMLC.2005.1527805

[20] G. Guo, H. Wang, D. Bell, Y. Bi, K. Greer, KNN model-based approach in classification, OTM Confederated International Conferences" On the Move to Meaningful Internet Systems, 2003, 986–996 . https://doi.org/10.1007/978-3-540-39964-3_62

[21] M. Z. Jahromi , E. Parvinnia, R. John , A method of learning weighted similarity function to improve the performance of nearest neighbor, Inf. Sci., 179 (2009) 2964-2973. https://doi:10.1016/j.ins.2009.04.012

[22] H. Dubey , Efficient and accurate kNN based classification and regression." Ms.c. Thesis, International Institute of Information Technology , India , 2013.

[23] S. S. Alfere, A. Y. Maghari, Prediction of Student's Performance Using Modified KNN Classifiers, 1st International Conference on Engineering & Future Technology, 2018.

[24] P. McCool, G. D. Fraser, A. Chan, L. Petropoulakis, J. J. Soraghan, Identification of contaminant type in surface electromyography (EMG) signals, IEEE Trans Neural Syst Rehabil Eng., 22 (2014) 774-783. 10.1109/TNSRE.2014.2299573