DESIGN AN SYSTEM FOR HAND GESTURE RECOGNITION WITH EMG SIGNAL BY NEURAL NETWORK

Vaishali M. Gulhane¹, Dr. Amol Kumbhare²

¹Electronics & Communication Dept., Dr. A.P.J.Abdul Kalam University, Indore /MP/India
²Asso. prof. in Electronics Department, Dr. A.P.J.Abdul Kalam University, Indore /MP/India

ABSTRACT

The intellectual computing of an effective human-computer interaction (HCI) or human alternative and augmentative communication (HAAC) is vital in our lives in today's technological environment. One of the most essential approaches for developing a gesture-based interface system for HCI or HAAC applications is hand gesture recognition. As a result, in order to create an advanced hand gesture recognition system with successful applications, it is required to establish an appropriate gesture recognition technique. Human activity and gesture detection are crucial components of the rapidly expanding area of ambient intelligence, which includes applications such as robots, smart homes, assistive systems, virtual reality, and so on. We proposed a method for recognizing hand movements using surface electromyography based on an ANN. The CapMyo dataset based on the Myo wristband (an eight-channel sEMG device) is utilized to assess participants' forearm sEMG signals in our technique. The original sEMG signal is preprocessed to remove noise and detect muscle activity areas, then signals are subjected to time and frequency based feature extraction. We used an ANN classification model to predict various gesture output classes for categorization. Finally, we put the suggested model to the test to see if it could recognize these movements, and it did so with an accuracy of 87.32 percent.

Keywords—Hand Gesture Recognition, Human Computer Interaction, Electromyography, Artificial Neural Networks.

1. INTRODUCTION

Human computer interaction (HCI) or human alternative and augmentative communication (HAAC) is becoming increasingly essential in our daily lives in the current intelligent computing environment. One of the most important research techniques in HCI or HAAC applications is gesture recognition. Gestures are communicative, meaningful body motions or body language expressions that involve physical actions of the fingers, hands, arms, face, head, or torso with the aim of transmitting meaningful information or engaging with the environment in general. Hands are utilized to perform the majority of essential body language expressions since they are more flexible and controlled elements of the human body. As a result, hand gestures are appropriate for communicating information such as expressing a sentiment, indicating a number, pointing out an object, and so on. Sign language and gesture-based computer control both employ hand gestures as a main interface method [1], [7]. Simple mechanical devices such as a keyboard and mouse are utilized for man-machine interaction in a conventional HCI system. These gadgets, on the other hand, limit the pace and spontaneity with which man and machine communicate. On the other hand, due to its natural interaction ability, interaction methods based on hand gestures and computer vision have become a popular alternative communication modality for man-machine contact in recent years. For successful HCI or HAAC applications such as robots, sign language communication, virtual reality, and so on, a proper design of hand gesture recognition framework may be utilized to build an advanced hand gesture based interface system. The method of recognizing the movements made by a human hand or hands is known as hand gesture recognition [13].

Because hand gestures vary in time and location, as well as across ethnic backgrounds, it's nearly difficult to identify any single one. Any gesture, broadly defined, is a deliberate hand movement that involves movements of the fingers, palm, and wrist to communicate information. When employed to generate a range of postures or motions, gestures are regarded to extend to the arms as well as the hands in the field of gesture recognition. A computer or other machine that detects human gestures is one that can be operated efficiently using the user's hands and arms. Because gestures come easily to humans and are an essential element of how we communicate, using gestures for human-computer interactions is a simple process. Users should be able to instruct computers to perform complicated activities using only a single posture or a few basic, continuous, dynamic hand gestures [23]. Hand gesture recognition technologies are divided into two groups depending on sensing techniques: (i) gloves, sensor, or wearable hand based approaches, and (ii) vision-based techniques.
EMG signals are used in a variety of applications, including neuromuscular illness diagnosis, prosthetic/orthotic device control, human-machine interfaces, virtual reality gaming, and the creation of muscle-oriented exercise equipment. EMG signals have been used in a variety of research projects to interface with machines and computers. The goal of gesture recognition research is to develop a system that can recognize distinct human gestures and utilize them to communicate information or control devices [4] [6] [8] [14]. Gestures can come from any body action or state, although they are most often made with the hands or the face. EMG signals, which are used to assess muscle activity, are used to record hand movements. The development of EMG-based control has gotten a lot of attention in the last three decades since it will improve the social acceptability of handicapped and elderly people by enhancing their quality of life. However, pattern identification of EMG data is the most difficult element of designing myoelectric control-based interfaces [21] [30]. This is due to substantial differences in EMG signals, which have various characteristics based on age, muscle activity, motor unit pathways, skin-fat layer, and gesture style. EMG signals include more complex kinds of noise than other bio signals, which are generated by intrinsic equipment and ambient noise, electromagnetic radiation, motion artefacts, and the interaction of various tissues. It might be difficult to extract valuable characteristics from an amputee's or handicapped person's remaining muscles. When dealing with a multiclass classification problem, this problem becomes even more complex to solve. Many studies have explored various types of EMG signal categorization techniques with satisfactory recognition results. The goal of this study is to create a system architecture that can detect hand movements that communicate specific information utilizing an EMG signal.

The following is a breakdown of the paper's structure. The work of several researchers on the categorization of the EMG signal is discussed in section II. The suggested technique is presented in Section III, which includes signal processing, namely segmentation, feature extraction, and classification. In part IV, the findings are examined, and in section V, the conclusion is presented.

II. RELATED WORK

The majority of the studies utilized hand gesture recognition methods that may be divided into two categories: i) glove, sensor, or wearable hand based techniques, and ii) vision-based techniques, all of which are detailed below.

P N Huu et al. [1] researched, surveyed, and performed research in order to give an overview of human hand gesture recognition, covering the main processes of hand gesture recognition, as well as the most often used approach and methodology. A convolutional neural network-based gesture recognition technique is presented, in which training and testing are carried out with various convolutional neural networks, as opposed to other known methods and designs [2]. For the problem of dynamic hand gesture detection, an efficient strategy for utilizing knowledge from various modalities in training unimodal 3D convolutional neural networks (3D-CNNs) is described [3]. A. Chahid and colleagues [4] devised a technique. The Quantization-based position Weight Matrix (QuPWM) feature extraction approach for multiclass classification to improve the interpretation of biomedical signals has shown promise in extracting important features from a variety of biomedical signals, including EEG and EMG signals. A technique [5] was suggested to categorize hand motions by one smartphone utilizing inaudible high-frequency sound, and the model classified 8 hand gestures with 94.25 percent accuracy. A. Devaraj et al. [6] utilized SVM and KNN machine learning classifiers to recognize a specific hand motion from an EMG signal recorded by a sensor-based band. B. Besma et al. [7] presented a method for identifying two movements (wrist flexion and wrist extension) by decoding surface electromyography (EMG) signals of amputees, which has proved to have an overall accuracy of about 80% and may be utilized for prosthesis. A comparison is presented between the created band and the Myo Armband, which recognizes gestures using surface-Electromyography (s-EMG) [8]. Based on spatio-temporal characteristics such as Histogram of Oriented Gradients 3D Hog and Histogram of Oriented Optical Flow 3D Hoof, a gesture recognition system was described [9]. [10] suggested an effective deep convolutional neural networks technique for hand gesture detection that used transfer learning to overcome the need of a big labelled dataset.

The area under the ROC curve metric achieved 98 percent overall performance using a machine learning technique developed [11] for real-time identification of 16 movements of user hands using the Kinect sensor that respects such constraints. Hand movements and finger detection in still pictures and video sequences are the subject of T. Bravenec et al. [12]. Using a wrist-mounted tri-axial accelerometer, a computational solution for human-robot interaction was given [13]. Surface electromyography was used to identify hand movements using a technique based on support vector machines (SVM) [14]. (sEMG). A model [15] for real-time hand gesture identification that takes as input electromyography (EMG) data recorded on the forearm and uses an auto-encoder for automated feature extraction and an artificial feed-forward neural network for classification, utilizing the commercial sensor Myo Armband. A novel 3D hand gesture identification technique [16] is based on a deep learning model in which sequences of hand-skeletal joint locations are analyzed by parallel convolutions in a new Convolutional Neural Network (CNN). Experimental results are consistent with theoretical estimates and illustrate the benefits of the suggested gesture recognition system design [17], which utilizes a hand detector to identify hands in the frame and then switches to a gesture classifier if a hand is found. A hand gesture recognition solution based on LSTM-RNNs and 3D Skeleton Features [18] presents, with experimental results showing that the suggested technique has a resilience of 92.196 percent on a self-defined dataset. The use of the temporal inter-frame pattern on the identification of both static and dynamic hand gestures is explored in a three-level system [19]. A sensor-based system that deciphers this sign language of hand gestures for English alphabets has been created[20].
A hand gesture identification model [21] developed that used surface electromyography in the transient state, support vector machines (SVM), and discrete wavelet transforms to recognize hand motions that lasted a brief period (i.e., short-term gestures) (DWT). Author [22] has provided a successful dynamic hand gesture and movement trajectory recognition system that may be utilized in real-time for effective HRI interaction. E. Kaya et al. [23] developed a hand gesture recognition method based on surface electromyography (EMG) signals collected from a wearable device, the Myo armband, to classify and recognize numbers from 0 to 9 in Turkish Sign Language. To recognize the hand gestures, they used machine learning techniques such as kNN, SVM, and ANN. The suggested approach by J. Kim et al. [24] transforms reflected and recorded sound data into an image in a short time using a short time Fourier transform, and then applies the acquired data to a convolutional neural network (CNN) model to identify hand motions. On three datasets, a neural network design consisting of two types of recurrent neural network (RNN) cells was built [25], revealing that this very modest network beats state-of-the-art hand gesture identification techniques that depend on multi-modal data by a wide margin. K. N. Krisandria et al. [26] use hand motions to create interactions between humans and computers, which are recognized by the palm of the hand, which is derived from the findings of human skeletal segmentation using the Kinect camera. The framework combines incoming signals [27] at the semantic level, a method similar to that used in multi-agent systems, where modals give local semantics before entering the fusion module. A novel human hand gesture dataset is given, which was collected using a low-cost, wearable IoT-based device with accelerometer and gyroscope sensors. A real-time hand gesture recognition accelerator based on hand skeleton extraction has been suggested [28] [29]. Surface electromyography (sEMG) was collected from six hand and forearm muscles and categorized using three distinct techniques [30] [31] [32].

III. PROPOSED WORK

Process, analyze, and recognize the hand gesture signal is the goal of the EMG-based hand gesture recognition system. The whole process of a hand gesture recognition system may be broken down into two phases: training and testing. Figure 1 shows a schematic representation of a hand gesture recognition system. The preprocessing, feature extraction, and feature selection processes are the same in both the training and testing phases. Preprocessing, feature extraction, feature reduction, and classification are the phases of a hand gesture recognition system in general.

![Fig 1: The schematic diagram of a hand gesture recognition system.](image-url)
1. Hand Gesture Dataset

The CapgMyo is a benchmark database of high-density sEMG (HD-sEMG) recordings of diverse participants’ hand gestures, based on an 8x16 electrode array and an acquisition equipment as illustrated in fig 2.

![Original surface electromyography (sEMG) signals recorded by the MYO armband.](image)

2. Preprocessing

Because of its sensitivity, EMG signals are often polluted by external noise sources and artefacts. Using these tainted signals will also result in a poor classification result, which is undesirable. Electrode noise, motion artefacts, power line noise, ambient noise, and intrinsic noise in electrical and electronic equipment are the most common sources of noise, artefacts, and interference that can contaminate EMG signals. The first three forms of noise can be removed by employing standard filtering techniques such as a band pass filter or a band-stop filter, or by utilizing high-quality equipment with correct electrode placement. Other noises/artifacts and interferences of random noise that are in between the main frequency range of EMG are difficult to eliminate.

Wavelet-based approaches are useful for studying many forms of non-stationary data, such as EMG. For example, the Discrete Wavelet Transform (DWT) scales and shifts the mother wavelet and decomposes a discrete-time signal $x[k]$ into a collection of signals. Finding the appropriate number of wavelet decomposition levels is the first step in the DWT decomposition. At the same time, the signal $x[k]$ passes through the high-pass and low-pass filters. In wavelet decomposition, detail (D) represents the signal at high frequency, while approximation (A) represents the low-frequency component (A). The similarity between the signal and the wavelet functions is measured by these coefficients. After down sampling the resultant filtered signal by two, this procedure is repeated on the low pass approximation coefficients obtained at each level. This research focuses on frequency decomposition levels. At each decomposition level of the DWT, the resulting detailed coefficients reflect distinct frequency bands of the EMG signals. In this experiment, we discovered that DWT with DB wavelet provided the best results. Wavelet-based feature extraction methods create a vector that is far too large to be utilized as a classifier input. This approach reduces the amount of characteristics that may be extracted from wavelet coefficients. The chosen characteristics of EMG signals are extracted using DWT. After obtaining DWT coefficients, statistical characteristics for each of the five DWT sub-bands are retrieved.

3. Features Extraction and Selection

Because of the complexity of EMG signals, effective feature selection is critical for successful classification. The characteristics utilized to represent the raw EMG signals have a huge impact on the pattern classification system's performance. Because it is difficult to extract a feature parameter that completely reflects the unique characteristic of the recorded EMG signals to a motion instruction, many feature parameters are required for EMG signal categorization. For the categorization of EMG signals, traditional characteristics derived from the time domain, frequency domain, time-frequency domain, and time-scale domain are used. After wavelet transformation or 3 level DB decomposition of signal, we employed hand created, that is, manually derived features in the time and frequency domain. Mean absolute value (MAV), waveform length (WL), zero crossings (ZC), slope sign changes (SSC), RMS, and standard deviation are all assessed in the time domain. Skewness, mean frequency, and kurtosis are all terms used in the frequency domain.
4. Classification (train, validate and test)

After collecting feature dataset for train and test with its corresponding output labels, hand gesture classification was done using Artificial Neural Network via train, validate, and test stages to produce anticipated output as hand gesture. The ANN employed in this study is a dynamic and strong back-propagation (BP) type network. Its state evolves over time until it reaches the final equilibrium point, which is attained by successful training. The Widrow-Hoff learning rule is applied to a multiple-layer network with a nonlinear differentiable transfer function to generate BP. The learning rule for neural network propagation determines how the weights between the layers change. ANN in which hidden layer, activation function, epoch, error rate, and learning rate are utilized as hyper parameters to adjust or train the classifier for the optimal validation of the training process.

Where the training data has already been labelled, the classifiers from the supervised learning model are utilized. In machine learning, there are many different types of categorization algorithms. The Multi-Layer Perceptron (MLP) Artificial Neural Network (ANN) based on the Levenberg–Marquardt algorithm is used for classification in this study since it is a robust approach in this particular instance. According to the literature, the accuracy of an artificial neural network’s classification depends on the feature set, network topology, and training technique chosen. A series of input and output units linked together to form a network is what an ANN is. There are three layers in the network: an input layer, a tan-sigmoid hidden layer, and a linear output layer. The advantages of neural networks are primarily their high tolerance for noisy input and their ability to classify untrained patterns. They might also be beneficial if there isn’t enough information about the relationships between characteristics and classes. Hidden layers and output layer nodes were activated using a hyperbolic tangent sigmoid function and a linear function, respectively. The acquired characteristics of sEMG signals are fed into the ANN, and the network output is the categorization, or the estimated movement caused. The network's overall diagram is depicted in Fig.5.

![General Architecture of Artificial Neural Network](image)

Fig 3: General Architecture of Artificial Neural Network

The hidden layer size is an essential parameter for ANN, since it adds to network accuracy. The weights linking the hidden layer to the output get smaller as the number of neurons in the hidden layer grows. Increasing this value typically increases the network's training performance, but it doesn't always assist with generalization. Under-fitting occurs when a network's weight and biases are improperly adjusted during training due to the use of a small number of neurons in the hidden layers. If the number of hidden neurons is increased beyond a certain point, the network's accuracy may suffer. To save a high number of network variables, it is preferable to have a lot of memory, thus the training process becomes complicated. The right number of neurons in the hidden layer is chosen to balance the network's efficiency and complexity. The number of neurons in a neural network is not determined by a rule of thumb. For the selection of neurons in the neural network in this study, a trial and error technique is employed. To minimise overfitting, the training input data were randomly separated into three sets, with 70% of the samples assigned to the training set, 15% to the validation set, and 15% to the test set. To explore the influence of hidden layer size on class estimation accuracy, we raised the number of neurons from 1 to 15. Under-fitting is shown in neural networks with 1 hidden neuron, while over-fitting is seen in neural networks with 15 hidden neurons, with 5 being the ideal amount of hidden neurons. In the input layer, four neurons correspond to the input feature, while two neurons in the output layer represent the two classes. During training, the back-propagation method is used to modify the weights and biases while reducing the difference between the goal and neural network output.

More information isn't necessarily better in machine learning applications, as feature selection approaches demonstrate. Following the feature extraction procedure, it was discovered that a specific collection of feature(s) may impair or provide no value to the classifier's performance. Counting the number of times a feature divides a tree can be a useful metric for feature selection.

IV. RESULTS AND DISCUSSION

In this experimentation, we used MATLAB R2018b to construct a recommended architecture to assess the proposed model. On a desktop computer with an Intel® Core™ i5 CPU and 8GB of RAM, the suggested model was tested. The CapgMyo dataset [33] is a benchmark collection of high-density sEMG (HD-sEMG) recordings of hand gestures made by diverse individuals (able-bodied people ranging in age from 23 to 26 years) utilizing an 8x16 electrode array and a newly built acquisition equipment. We employed eight distinct hand motions in this project, as shown in Fig. 4. Each move was performed 10 times and held for 3 to 10 seconds each time (10 trials).
Fig 4: The used eight hand gestures from CapgMyo Dataset [22].

The performance of the classifier model is described using a confusion matrix also called error matrix. It’s a matrix in which each row represents examples from an actual class and each column represents instances from a predicted class, or vice versa. The confusion matrix is used to evaluate the performance using the accuracy, sensitivity, and specificity criteria.

\[
\text{Accuracy} = \frac{(TP + TN)}{(TP + TN + FP + FN)}
\]

\[
\text{Sensitivity} = \frac{TP}{(TP + FN)}
\]

\[
\text{Specificity} = \frac{TN}{(TN + FP)}
\]

Where, TP – true positive, TN – true negative, FP – false positive, FN – false negative

Fig 5: Decomposition of test signal using DWT
Figure 5 displays a sample of the test signal waveform after discrete wavelet transformation preprocessing (DWT). Training, validation, and testing are the three steps of our system's evaluation. 70% of the data samples are utilized in the training and validation phases, whereas 30% are used in the testing phase. After the train data has been validated, test samples are analyzed to determine the proper hand gesture as a projected output. Table 1 lists the parameters that were assessed for testing and compares them to existing techniques used by researchers, also graphically compared in fig. 6.

**Table 1: Comparative Results**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>81.25</td>
<td>83.1</td>
<td>86.0</td>
<td>87.32</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>70.48</td>
<td>-</td>
<td>-</td>
<td>75.44</td>
</tr>
<tr>
<td>Specificity</td>
<td>59.72</td>
<td>-</td>
<td>-</td>
<td>70.35</td>
</tr>
</tbody>
</table>

This paper provided a hand gesture detection algorithm based on CapgMyo datasets from the Myo armband device. The approach first preprocesses the data before extracting features from it using temporal and frequency domain statistics. Finally, utilizing 70% of the dataset feature vectors, an artificial neural network with feed forward backpropagation network is used to create a classifier. In this experiment, we test the remaining 30% of feature vectors. Instead of merely obtaining the results recognized by any gesture, every test feature vectors must be categorized by a classifier so that each feature vector may be recognized correctly. A feature vector is classed as no gesture if it is not detected. Our suggested model has an accuracy of 87.32 percent, which is greater than existing approaches. In the future, we'll try to implement the approach in a real-time application. Based on the outcomes of this study, it is clear that PPG can provide similar HGR results as s-EMG.

**VI. REFERENCES**


