Prediction of Flexion and Extension Movements of 4 Fingers of the Hand Using a New Labeled Method

J A García Torres, A Ibarra Fuentes, E Morales Sánchez and A Hernández Zavala

Instituto Politécnico Nacional. CICATA-IPN Unidad Querétaro. Cerro Blanco. No. 141. Querétaro, Qro. CP. 76090, México.

an hernandezz@ipn.mx

Abstract. This work presents a neural network classifier for identifying the flexion and extension movements for four fingers from the hand, out of the surface electromyography signals in the forearm muscles. A new labeled data method was proposed based on time segmentation to relate the sEMG signal with the corresponding finger movement. This is a different way of labeling the data for training the neural network, allowing to reduce the amount of training gesture hand. The experiment consists of 10 sessions in which the fingers are flexed randomly, one at a time for 2 minutes with a 16ms sample time. The absolute mean value (MAV) is used as a feature extractor in the time domain to average 5 samples and the normalized data is used for the neural network. Results show that this system with the labeled data method, provides a 98.83% precision value for the index finger, 93.46% for the ring finger, 80.34% for the middle finger, and 68.46% for the little finger. The results are similar to those found in the literature where they classify different gestures using the conventional labeling method.

1. Introduction

When performing a body movement, the involved muscles produce electrical signals which are known as electromyographic signal EMG. These signals are obtained with the use of intramuscular needles or by superficial electrodes known as surface electromyographic signals (sEMG) [1]. Due to their relative facility to obtain these electrical signals, sEMG has been widely used to analyze muscle activation patterns and movement prediction. The work from [2], uses 4 bipolar electrodes placed on the forearm to predict 3 movements: closed hand, open hand, and change of position. In the work presented in [3] the Myo Armband bracelet commercial sensor was used to classify 8 hand movements [4].

To interpret sEMG signals and predict gestures, different machine learning algorithms are used, such as neural networks or support vector machines. In [5], a deep neural network is used to classify 8 hand gestures. The training of this network was carried out with the database from [6] which contains five different positions of the upper limb of different subjects.

A large amount of labeled data is necessary to obtain good performance in neural network models. In the case of prediction of movements using sEMG [7], it is necessary to associate the sEMG data with a movement from the finger and the label to which it belongs. Getting these data is hard work, since it requires a variety of movements that belongs to the same class. Otherwise, the network will only be able...
to distinguish certain portions of the class. For example, if a network is trained with movements of a
duration of one second, the same performance may not be obtained for movements of different durations.
Or if you make the training with movements in which a lot of force is applied, the network may not
distinguish movements with less force. For this reason, data labeling alternatives are being sought to
solve these problems.

In this work, it is developed a neural network classifier with unique weights for each user, capable of
identifying the flexion and extension gestures of 4 fingers from the hand. A different method of easy
acquisition of labeled data is proposed and the neural network is trained using these data. The method
consists of splitting the whole movement into labeled instants of time, measuring the sEMG and the
position of the fingers on every sample. In this way, it is expected not to process exorbitant amounts of
data.

2. sEMG in the Forearm
The sEMG has been studied for the development of systems capable of predicting human movement
[8]. Its purpose varies like in the remote control of devices [9] and the control of robotic prostheses [10].
This work focuses on the prediction of flexion and extension movements of 4 fingers of the hand: index
finger, middle finger, ring finger, and little finger. Figure 1 shows the muscles found in the back and
front of the forearm, which are activated by the movement of the hand fingers, wrist, and elbow. The
responsible for flexing and extending the fingers of the hand are the common extensor muscle of the
fingers and the extensor muscle proper to the little finger.

The sEMG-360 technique allows obtaining the electromyographical activity around the entire forearm,
being adequate given that although certain muscles are not attributed to these gestures, they still register
electrical activity which contributes to the discrimination exerted by the neural network at classifying.

![Figure 1. Forearm muscles, a) Back of the forearm and b) Anterior part of the forearm.](image)

3. Labeled data acquisition
The amount of data with which a neural network is trained is very relevant to obtain a good performance.
For a neural network to be able to identify the sEMG hand gestures, the sEMG signals are recorded
during the performance of the gesture. Then, these signals are labeled within a membership class during
all the movement, and with this information, the neural network learns. The greater the variety of data
belonging to the same class is used for training, the better performance of the classification is obtained.
Due to the labeling task, data generation is often slow and complicated. For that reason, an easy and
automatic method for labeling is proposed, in which moments of the gesture are labeled instead of the
entire movement, reducing the number of gestures necessary for the neural network.
3.1. Instrumentation
The Myo Armband is a commercial device capable of classifying 5 hand gestures through 8 sensors integrated of 3 dry electrodes positioned using the sEMG-360 technique on the forearm. For this work, it was used as a data logger for sEMG of the muscles of the right forearm, obtaining values from the 8 electrodes at a sample frequency of 60 Hz. Figure 2 shows the system setup consisting of the Myo-Armband and the Finger Flexion Detector (FFD), the data used for training is generated by the two devices. The Myo Armband serves as an sEMG logger for obtaining the data from the 8 sensors. The FFD circuit is used for data labeling, this device was designed to detect the current position of each finger (flexion = 1, extension = 0), for each of the 4 fingers of the hand: index, middle, ring, and little. It works with 4 mechanical switches which are activated by flexing each finger sending a 5V signal to a microcontroller which registers a 1x4 array with the current state of each finger at a frequency of 60 Hz.

3.2. Data generation
The experiment to obtain the data was carried out with the same test subject (male, 26 years old). Since the applications of systems like these are normally used by specific users, so it is preferable to train the network for one user and not for a variety of users, since the sEMG of each user are different due to physical characteristics as type skin, muscle mass, and interface electrode-skin. The user performed 10 sessions each session consisted of randomly flexing and extending each finger at a time for 2 minutes and 8 seconds and a sample time of 16 milliseconds, therefore each session provided 8000 samples. Two vectors per sample were obtained: \(\mathbf{x}_t\) that contains the data of the 8 sEMG sensors and is 1x8, and \(\mathbf{y}_t\) that corresponds to the current state of each of the fingers and is 1x4. By concatenating the 10 sessions, the results are the bidimensional arrays \(\mathbf{X}\) from size 80000x8 and \(\mathbf{Y}\) from size 80000x4.

3.3. Data processing
Once the data was acquired, it was smoothed by applying the absolute mean value (MAV) filter \([11]\). Five samples were averaged applying the equation (1) to matrix \(\mathbf{X}\) and equation (2) to matrix \(\mathbf{Y}\), obtaining as results \(\mathbf{X}_{MAV}\) matrix of size 16000x8 and \(\mathbf{Y}_{MAV}\) matrix of size 16000x4.

\[
X_{MAV} = \frac{1}{5} \sum_{i=1}^{5} |X_i|
\]

\[
Y_{MAV} = \text{Round}\left(\frac{1}{5} \sum_{i=1}^{5} Y_i\right)
\]

where \(X_i\) is the sample value; \(X_{MAV}\) is the smoothed sample; \(Y_i\) is the current state of the finger of the hand and \(Y_{MAV}\) is the category of the smoothed sample. Figure 3 shows 4 of the 8 sensors corresponding to \(X_{MAV}\) and the labeling values of each of the fingers \(Y_{MAV}\).
The matrix $X_{\text{MAV}}$ is normalized to handle numbers between 0 and 1 by applying the equation (3) to each column of the matrix.

$$X_{\text{Normalized}} = \frac{X_{\text{MAV}} \cdot X_{\text{Min}}}{X_{\text{Max}} - X_{\text{Min}}}$$  \hspace{1cm} (3)

where $X_{\text{MAV}_i}$ is the value of a column of the sample to be normalized; $X_{\text{Min}}$ is the minimum value of the column of $X_{\text{MAV}_i}$; $X_{\text{Max}}$ is its maximum value and $X_{\text{Normalized}}$ is the normalized sample.

![Figure 3. sEMG signal after applying MAV and their corresponding classification, flexion=100, extension=0.](image)

To generate a sequence of data that serves as input features for the neural network, 5 samples were integrated, 4 previous samples and the current sample. The final package contains 40 values associated with the 8 sensors of the Myo-Armand device and 4 values that categorize the state of the fingers of the hand. In this way, the number of samples was not reduced only more features were added to each one.

### 4. Neural network

The data set for the neural network was divided into 80% for training and 20% for testing. A 15% fraction of the training set was used for validating its behavior in a three-layer full-connected neural network. The number of nodes in the input layer and the hidden layer was chosen based on their performance in various training. The input layer has 700 nodes, the hidden layer 200, and the output layer has 4 nodes as there are four classes: index, ring, middle and little finger. All the layers use non-linear activation sigmoid function as in equation (4). Normally, the activation function used for multiclass classification is softmax, but the sigmoid function allows to choose a better decision threshold to increases or decreases the number of false positives, which is useful for systems like this. Each layer was applied a 20% dropout regularization.

The Adam optimizer was used with a learning rate of $1 \times 10^{-4}$ in a batch size of 64, these parameters were also chosen based on their performance in various training times.

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$  \hspace{1cm} (4)

where $Z$ corresponds to the multiplication of the weights by the input of the layer plus the bias ($W^T x + b$).

The binary cross entropy is used as loss function like in equation (5). It provides the error between the predictions and the real values. For a better understanding of this concept, if the loss function is equal to 0 it means that the predictions are totally equal to the real ones [12].

$$E = -\frac{1}{N} \sum_{i=1}^{N} y_i \cdot \log(p(y_i)) + (1 - y_i) \cdot \log(1 - (p(y_i)))$$  \hspace{1cm} (5)
where $y_i$ is the real label and $p(y_i)$ is the predicted probability of belonging to a class from the input data.

5. Results

Once the data is separated with the specifications mentioned above, the program was executed during 150 epochs calculating the loss value for the validation set in each epoch. The model with the lowest value was selected since it presents the best performance. The loss value for the training set and the validation set for each epoch is presented in Figure 4a. Finally, a value of 0.1572 was obtained for the validation set and 0.1515 for the test set in the loss function.

Due to the nature of the project and the method for obtaining data, each class had more zeros than ones given that there were more instants with extended fingers than flexed, for that reason the accuracy is not informative. Therefore, it was decided to observe the precision which indicates for this case, the percentage of correct flexing instants predictions (true positives) over the total number of flexing instants predictions (true positives + false positives). The recall provides for this case, the percentage of correct flexing instants predictions (true positives) over the actual total number of flexing instants (true positives + false negatives).

The precision and recall curve provide a graph on which the decision threshold can be selected for defining from which probability the class membership is determined. By observing the graph, we selected the threshold for having the highest percentage of precision without affecting the recall value. This indicates having the least number of false positives. A false positive in the network is predicting the flexing of a non-moving finger. Having a high precision value also increases the number of false negatives, indicating neural network predicts not flexing when a finger is moving, which is preferable to predicting an unintended movement (false positive). Figure 4b shows the precision and recall curve for the index finger class with a threshold of 0.86. This threshold allows having a precision value of 98.83% and a recall value of 72.28%. Table 1 shows the precision and recall values obtained for each class.

![Figure 4](image)

**Figure 4.** Performance curve. a) Loss value for the validation and training set through the epochs. b) Precision vs Recall curve.

**Table 1.** Performance obtained for each of the classes.

|                | Index finger | Middle finger | Ring finger | Little finger |
|----------------|--------------|---------------|-------------|--------------|
| Precision      | 98.83%       | 80.34%        | 93.46%      | 68.46%       |
| Recall         | 72.28%       | 66.79%        | 69.14%      | 62.23%       |
The network was tested with the data set from session 10. Figure 5 shows the prediction performance of the network and the label for the index finger.

Figure 5. Blue: real label for each sample; Red: prediction from the network for each sample.

6. Discussion
The system was able to present similar results when comparing it with other gesture identifiers for the index finger and ring finger classes. In [13], the gesture of cylindrical grip, hook grip and ball grip presented a precision of 93%, 85% and 90% respectively, in which they use a database with 450 gestures. The proposed labeling method has an average of 8 gestures per session, a total of 80 gestures performed to obtain 16,000 tagged data. In Figure 5, it is observed that the prediction of the neural network is similar to the real label, identifying all the flexion movements made by the index finger. There are small lags between the real label and the predicted one in 1 or 2 instants, remember that the labeled instants are every 16 ms, which for example, for the control application of a robotic hand in real time for prostheses are not detrimental. It can be also noticed that sometimes there are unexpected peaks. It seems to be a detrimental effect to the objective of this project, but it is proposed to combine the results of this network with conditionals, in this way these peaks could be easily suppressed.

The middle finger and little finger classes presented a lower performance, it is considered that this is due to the difficulty of isolating the flexion movement of these fingers, that is, when the index finger is moved intentionally the ring finger also moves, which makes the sEMG measures similar between the gestures of these fingers, hindering the discrimination exerted by the neural network.

7. Conclusions
This work proposes an easy way to obtain labeled data from finger movement. This method consists of labeling each instant of time during the movement (time segmentation). Labeled data were used to train a fully connected neural network to identify flexion and extension gestures of 4 fingers from the hand.

The results showed a prediction gesture for the index finger and the ring finger with a precision value of 98.83% and 93.46% respectively, whereas for the middle finger and the little finger it was obtained a lower value, 80.34% and 68.46% of precision, which may be due to the difficulty of isolating the movement of these fingers. The proposed method of labeled data based on time segmentation reduces the required training gesture hand when is compared with the conventional method of labeling.

References
[1] Correa Figueroa J L, Morales Sánchez E, Huerta Ruelas J A, González-Barbosa J J and Cárdenas Pérez C R 2016 Sistema de adquisición de señales SEMG para la detección de fatiga muscular Rev. Mex. Ing. Biomed. 37 27–17
[2] Tavakoli M, Benussi C and Lourenco J L 2017 Single channel surface EMG control of advanced prosthetic hands: A simple, low cost and efficient approach Expert Syst. Appl. 79 332–322
[3] Sánchez Velasco L E, Arias Montiel M, Guzmán Ramírez E and Lugo González E 2020 A low-cost EMG-controlled anthropomorphic robotic hand for power and precision grasp Biocybern. Biomed. Eng. 40 221–37
[4] Visconti P, Gaetani F, Zappatore G A and Primiceri P 2018 Technical features and functionalities of Myo armband: An overview on related literature and advanced applications of myoelectric armbands mainly focused on arm prostheses *Int. J. Smart Sens. Intel. Syst.* **11** 25–1

[5] Mukhopadhyay A K and Samui S 2020 An experimental study on upper limb position invariant EMG signal classification based on deep neural network *Biomed Signal Process Control* **55** 8–1

[6] Khushaba R N, Takruri M, Miro J V and Kodagoda S 2014 Towards limb position invariant myoelectric pattern recognition using time-dependent spectral features *Neural Networks* **55** 58–42

[7] Chollet F 2020 *The Sequential model* [internet] available in https://keras.io/guides/sequential_model/

[8] Jia G, Lam H K, Liao J and Wang R 2020 Classification of electromyographic hand gesture signals using machine learning techniques *Neurocomputing* **401** 248-236

[9] Ali M, Riaz A, Usmani W U and Naseer N 2020 EMG based control of a quadcopter *Int. Conf. Mech. Electron. Comput. Ind. Technol.* **3** 250–4

[10] Shi W T, Lyu Z J, Tang S T, Chia T L and Yang C Y 2018 A bionic hand controlled by hand gesture recognition based on surface EMG signals: A preliminary study *Biocybern. Biomed. Eng.* **38** 126–35

[11] Alkan A and Günay M 2012 Identification of EMG signals using discriminant analysis and SVM classifier *Expert Syst. Appl.* **39** 44–7

[12] Gómez R 2018 Understanding categorical cross-entropy loss, binary cross-entropy loss, softmax loss, logistic loss, focal loss and all those confusing names [internet] available in http://gombru.github.io/2018/05/23/cross_entropy

[13] Rascón Madrigal L H, Sinecio Sidrian M A, Mejía Muñoz J M, Díaz Román J D, Canales Valdiviezo I and Botello Arredondo A I 2020 Estimación en la intención de agarres cilíndrico, esférico y gancho utilizando redes neuronales profundas *Rev. Mex. Ing. Biomed.* **41** 117–27