A User-Specific Hand Gesture Recognition Model Based on Feed-Forward Neural Networks, EMGs, and Correction of Sensor Orientation

Marco E. Benalcázar *,†,‡, Ángel Leonardo Valdivieso Caraguay †,‡ and Lorena Isabel Barona López †,‡

Artificial Intelligence and Computer Vision Research Lab, Departamento de Informática y Ciencias de la Computación (DICC), Escuela Politécnica Nacional, Quito 170517, Ecuador; angel.valdivieso@epn.edu.ec (Á.L.V.C.); lorena.barona@epn.edu.ec (L.I.B.L.)
* Correspondence: marco.benalcazar@epn.edu.ec; Tel.: +593-2-2297-6300 (ext. 4706)
† Current address: Ladrón de Guevara E11-253, Quito170525, Ecuador.
‡ These authors contributed equally to this work.

Received: 17 October 2020; Accepted: 22 November 2020; Published: 1 December 2020

Abstract: Hand gesture recognition systems have several applications including medicine and engineering. A gesture recognition system should identify the class, time, and duration of a gesture executed by a user. Gesture recognition systems based on electromyographies (EMGs) produce good results when the EMG sensor is placed on the same orientation for training and testing. However, when the orientation of the sensor changes between training and testing, which is very common in practice, the classification and recognition accuracies degrade significantly. In this work, we propose a system for recognizing, in real time, five gestures of the right hand. These gestures are the same ones recognized by the proprietary system of the Myo armband. The proposed system is based on the use of a shallow artificial feed-forward neural network. This network takes as input the covariances between the channels of an EMG and the result of a bag of five functions applied to each channel of an EMG. To correct the rotation of the EMG sensor, we also present an algorithm based on finding the channel of maximum energy given a set of synchronization EMGs, which for this work correspond to the gesture wave out. The classification and recognition accuracies obtained here show that the recognition system, together with the algorithm for correcting the orientation, allows a user to wear the EMG sensor in different orientations for training and testing, without a significant performance reduction. Finally, to reproduce the results obtained in this paper, we have made the code and the dataset used here publicly available.

Keywords: classification; correction of rotation of EMG sensors; electromyography; EMG; feed-forward neural network; hand gesture recognition

1. Introduction

A gesture can be considered as a symbol that communicates expressions or physical behavior. Many investigations use gestures as a communication tool between a computer and a human to interact or control devices in different fields, including medical applications, human–machine interfaces, automotive field, robot control, sign language recognition, sign language translation, movement modeling, and virtual reality [1–6]. Gesture recognition allows us to determine the user’s intention based on the identification of facial expressions, eye gaze, hand motion, and the movement of the body segment or posture. The development of a hand gesture recognition model is a challenging problem because it requires high accuracy to identify the class of a given movement, the instant of time when it
happens, and its duration. Moreover, to develop a system with acceptable portability and usability characteristics, it needs to be able to work with limited computational resources [2,7,8].

Currently, several works are focused on hand gesture classification instead of recognition [8–11]. Classification models only identify the class of a given hand movement from a predefined set of gesture classes. In other words, the classification only returns a label that identifies the gesture (which class), whereas the recognition returns the label as well as its timestamp (when it is produced and how long it lasts). As a result, the problem of gesture recognition is harder than classification. Moreover, a recognition model as part of its architecture, usually includes a classification model [7].

There are different types of sensors to develop hand gesture recognition models such as gloves equipped with force and flexion sensors [12,13], optical and infrared cameras [14], electromyographic (EMG) sensors [15], and inertial measurement units (IMUs) [16–19]. Each of these sensors has problems to solve to achieve promising results in hand gesture recognition applications. For example, in vision-based gesture recognition models, the detection and tracking of the hand movement may be affected by changes of illumination and occlusions [20]. Using a glove can be uncomfortable if the glove does not fit the user’s hand properly or if it is used for a long period of time. Signals measured with IMUs (e.g., angular velocity and acceleration) are usually noisy, thus making it difficult to extract the part of the signal that conveys information of the gesture executed.

For this work, EMG sensors are used because they do not suffer from some of the problems mentioned above. For example, EMG sensors do not have problems because of occlusion and changes in illumination; they are usually easy and comfortable to use, and EMG sensors return less noisy signals than IMUs. EMG sensors measure electric signals on skeletal muscles. These electric signals, called EMGs, contain information of a voluntary movement commanded by the brain. EMG sensors can be of two types: intramuscular and surface sensors. On one hand, intramuscular sensors use a set of invasive electrodes (i.e., needles), which make good and localized contact with the part of a skeletal muscle under examination, thus making it possible to record EMGs from deep skeletal muscles, with little interference. However, intramuscular sensors are not practical for implementing gesture recognition models because one or more needles must be inserted in the muscle causing pain to the user. On the other hand, a surface sensor consists of a set of non-invasive electrodes, such as wearable bracelets or adhesive pods, so wearing this type of sensor is easy and comfortable [21,22]. For this reason, in this work, we use a surface EMG sensor called a Myo armband produced by Thalmic Labs (Kitchener, Ont., Canada).

As we mentioned before, an EMG sensor measures the electrical signals when a voluntary movement is produced due to the contraction of skeletal muscles. This contraction is produced by the activation of a set of muscle fibers, called a motor unit (MU), where each fiber of the MU is connected to the axon of a motor neuron. The force produced by a muscle depends on the number of MUs activated by the brain. MUs contract when a train of electrical pulses, generated by the brain, reaches each of its fibers. This train of pulses that reach a MU is called Motor Unit Action Potential Train (MUAPT). The superposition (i.e., linear summation) of the MUAPTs of several motor units is what we call an EMG. Additionally, an EMG contains not only MUAPTs but also Gaussian noise [23] due to signals generated by nearby organs, environment noise, or noise produced by the electrical equipment used to measure the EMG. It is important to highlight that an EMG signal behaves as a non-stationary process because the distribution of these signals change with time. The statistics of an EMG depend on several factors such as the location of the sensor, corporal temperature, and body fatness [24]. Consequently, defining a mathematical model to characterize an EMG signal is a difficult task. In Figure 1, we show an example of EMGs measured by the Myo armband for each of the five gestures recognized in this work: wave in, wave out, fist, fingers spread, and double tap. These five gestures are the ones recognized by the proprietary system of the Myo armband, a sensor that returns EMGs composed of eight channels measured at 200 Hz. It is worth mentioning that, in other works, some authors used sensors with EMGs sampling rates from 200 to 1000 Hz obtaining sundry results in their recognition systems [7]. However, the Myo armband sensor sampling rate—200 Hz—has been
demonstrated to provide high-performance results in hand gesture recognition applications, which make it one of the most popular EMGs sensor in the market [7,25].

There are two types of hand gesture recognition models: user-specific and general. On one hand, a general model is designed to recognize gestures from any user. For the training phase, a general model uses the data gathered from a set of users, and then the model can be tested with other users. Therefore, a general model must be trained only once. On the other hand, a user-specific model must be trained and tested with the data of each user. Although user-specific models must be retrained for a new user, it is well-known that these types of models provide higher accuracy than general models. Consequently, for this research, we have chosen to design a user-specific model.

![Figure 1. Examples of EMGs of the five hand gestures used in this work.](image)

It is important to note that even a slight variation of the orientation of the EMG sensor between training and testing can cause the reduction of the performance in terms of accuracy for user-specific and general models [26]. This variation happens when the sensor’s pods are placed in a different position every time a user wears the armband. Therefore, a correction of the sensor orientation is needed to compensate for any change in the orientation between training and testing. There are few works in the literature that address the orientation correction problem of the EMG sensor. For example, a set of experiments was carried out in [26] that demonstrate that shifting the sensor 2 cm can cause the classification accuracy to drop considerably with accuracy between 50% and 60%. In another work, it is demonstrated that, if the sensor is rotated, a remapping can be performed to reorganize the distribution of the sensors and thus try to avoid drops in the classification accuracy due to the incorrect position of the sensor [27]. Finally, another method can be applied by analyzing the angle shift of the EMG sensor [28]. However, such approaches were computationally expensive due to the high complexity of the calculations required to perform the orientation correction procedure.

Some researchers [11,27,28] use the Myo armband or the Delsys sensor [29] to measure EMGs. Even though they claim to propose hand gesture recognition systems; in practice, they only perform classification. As it was mentioned before, classification is different from recognition, so most of the literature provides classification systems instead of recognition systems. In [26], a high-density electrode array is used. This sensor array is composed of 96 channels that measure EMGs used to distinguish 11 hand and wrist movements. However, this proposal is focused on the classification task only. Regarding the correction of sensor orientation, there are a few works that tackle this problem using the Myo armband [27,28]. Despite these two works solving the problem of the correction of sensor rotation, they do not perform recognition. Therefore, it is important to design systems that perform recognition with the correction of sensor orientation.
In this work, a user-specific model with correction of sensor orientation is proposed. The proposed algorithm for the correction of orientation identifies the EMG channel with the maximum energy, and then the EMG channels are rearranged based on this channel. To the best of our knowledge, this is the first proposal that applies the correction of orientation for a user-specific hand recognition model. Therefore, the main contribution of this work is to propose a user-specific hand gesture recognition model that allows a user to wear the bracelet in different orientations between training and testing, without experiencing a significant reduction in the recognition accuracy. In this way, we avoid the need of training a user-specific model each time it needs to be used, which is the usual approach for user-specific models.

To implement our hand recognition model, five different modules are proposed: data acquisition, pre-processing, feature extraction, classification, and post-processing. The data acquisition module is in charge of gathering the signals returned by the EMG sensor. The collected EMGs are related to five hand gestures executed for 5 s each. The gestures that we use for this work are the same ones recognized by the proprietary recognition system of the Myo armband. The pre-processing module is responsible for preparing and cleaning the collected EMG data. Typical approaches used at this module include filtering [30], offset compensation [7], rectification [31], normalization [32], and segmentation [22]. The feature extraction and feature selection module allowed us to extract independent features that represent accurately each EMG to be processed. The main challenge of this module is to determine the most significant and meaningful features to be fed to the classification module. Performing feature selection is a difficult task because it is a combinatorial problem [33]. The feature extraction methods can provide time, frequency, and time-frequency features from the EMGs. Some of the feature extraction methods that can be used are the mean absolute value (MAV), root mean square (RMS), variance (VAR), standard deviation (SD), among others [13,34,35]. Once the features are selected, this information is classified in different classes using the classification module. Some of the classifiers that can be used are K-nearest neighbor [36], decision trees [37], and Artificial Neural Networks (ANN) [38]. Finally, the post-processing module consists of refining the classifier responses by smoothing or filtering spurious labels [39].

The collected EMGs are related to five hand gestures executed for 5 s each. The gestures that we use for this work are the same ones recognized by the proprietary recognition system of the Myo armband. The pre-processing module is responsible for preparing and cleaning the collected data in order to reduce noise. The feature extraction and feature selection module allowed us to extract independent features that represent accurately each EMG to be processed. The main challenge of this module is to determine the most significant and meaningful features to be fed to the classification module. Performing feature selection is a difficult task because it is a combinatorial problem [33]. Once the features are selected, this information is classified in different classes using the classification module. Finally, the post-processing module consists of refining the classifier responses by smoothing or filtering spurious labels [39].

The main contributions of this paper include: (i) a user-specific hand gesture recognition model together with (ii) an algorithm for the correction of sensor orientation. The proposed recognition model is based on the use of a shallow feed-forward artificial neural network that takes as input a feature vector composed of the covariances between the channels of an EMG and a bag of five functions applied to the channels of an EMG. To reproduce the results obtained in this paper and analyzing the fine details of the proposed model, we have made the code and dataset publicly available.

Following this introduction, the remaining of this document is organized into four sections. In Section 2, we describe the dataset and the methodology used in this work. In Section 3, we present the results obtained with the proposed model. In Section 4, we present an analysis and a discussion of the results obtained. Finally, in Section 5, we present the conclusions and outline future work.
2. Dataset and Proposed Model

2.1. Dataset

The dataset used for this work is composed of EMG signals. These signals were measured using the Myo armband, which is a commercial and low cost EMG bracelet manufactured by Thalmic Labs (Figure 2). The bracelet has eight pods distributed uniformly along its circumference. The digital EMG signals returned by the sensor are organized into eight channels, in accordance with the order of the Myo pods, and are measured using a sampling frequency of 200 Hz.

Once the Myo armband is worn on the forearm, the user is requested to make five hand gestures: wave in, wave out, fist, fingers spread, and double tap. The rest position, as well as other gestures, are referred to as no gesture. Figure 2 highlights the gestures used in this work in their final positions.

The dataset used here is composed of EMGs of 120 users which were randomly divided into two groups: training and testing. Subsequently, the EMGs of each user in each of these two groups were randomly divided into two subsets: one for training and the other for testing (see Figure 3). The training subset for each user $D_{train} = \{(F_1, Y_1), \cdots, (F_N, Y_N)\}$ is composed of a total of $N = 130$ EMGs, where $F_i \in [-1, 1]^{1000 \times 8}$ is a matrix that contains the i-th EMG signal recorded for 5 seconds. The label for $F_i$ is denoted by the categorical variable $Y_i \in \{wave\ in, wave\ out, fist, fingers\ spread, double\ tap\}$, with $i = 1, 2, \cdots, N$. In the subset $D_{train}$, we have 25 EMGs for each of the five gestures to be recognized and five EMGs for the hand relaxed. Each of the five gestures for this work was performed following a sequence of three positions: rest, final position, and rest.

The testing subset $D_{test} = \{(G_1, Y_1), \cdots, (G_M, Y_M)\}$ for each user is also composed of a total of $M = 130$ EMGs, with 25 EMGs for each gesture to be recognized and 5 EMGs for the hand relaxed. Each matrix $G_i$ belongs to the matrix space $[-1, 1]^{1000 \times 8}$, with $i = 1, 2, \cdots, M$.

The dataset used for this paper contains a total of 15,600 EMGs. In this dataset, 75% of users are men and the remaining 25% are women. The age of the users varies from 17 up to 29 years. Finally, 91% of users are right-handed and the remaining 9% are left-handed, and 16% of the users suffered some injury to the arm. For the left-handed users, the EMGs were recorded on the right forearm.
2.2. Proposed Model

In this subsection, we describe the process applied to each EMG for the training and testing subsets of each user, which is represented in Figure 4.

Data Acquisition: To simulate the recognition model functioning in online mode, the concept of a sliding window was used for data acquisition as can be observed in Figure 5. In this work, we simulate the operation of the system because the data were acquired offline. We used a window of $n = 66$ points, with a stride of also 66 points. This window was subsequently divided into six non-overlapping sub-windows of length $m = 11$ points each. The segment of EMG observed through a sub-window was stored in a matrix $A^{(i)} \in [-1, +1]^{m \times 8}$, where $i$ denotes the number of the sub-window from which the data came from, with $i = 1, 2, .., 6$. It should be taken into account that $F_i$ denotes the EMG measured during the 5 s of data acquisition for each sample, while $A^{(i)}$ denotes exclusively the portion of an EMG viewed through an 11-point sub-window. Figure 5 illustrates how a 66-point window through sub-windows is divided in an EMG recording. The EMG signal is non-stationary, but if small windows are taken, the statistical variation of the probability distribution of the features is lower than large windows. We use sub-windows because empirically we observe that predicting a label for each EMG sub-window and then making a majority vote among these labels to predict the window label gave better results than predicting a single label considering only the 66-point window.

Pre-Processing: Each obtained sub-window $A^{(i)}$ goes through a process of rectification and filtering [30,31]. This process is used to obtain the envelope of each channel. The EMG envelope is assumed to contain the actual information of the gesture performed. For rectification, we used the absolute value function, thus obtaining the matrix $B^{(i)} = \text{abs}(A^{(i)})$, which has its values in the range $[-1, +1]$. The rectified envelopes $B^{(i)}$ are then further processed to obtain the final classification. The process is shown in Figure 4.

Results: The final result is obtained by comparing the obtained classification with the ground truth, which is the actual gesture performed. The recognition rate is calculated by counting the number of correctly classified gestures and dividing it by the total number of gestures. The performance of the model is evaluated using various metrics such as accuracy, precision, recall, and F1-score.

Figure 3. Scheme for dataset EMG-EPN-120.

Figure 4. Hand gesture recognition process for EMG signals by using the Myo Armband sensor.
interval $[0, 1]$. For filtering, a second-order Butterworth low-pass filter, with a cutoff frequency of 1 Hz, was applied to each column of the matrix $B(t)$, obtaining the matrix $C(t)$ with the signal filtered, with $i = 1, 2, \ldots, 6$.

**Feature Extraction:** In this step, we want to represent each of the pre-processed sub-window observations with a feature vector that belongs to a feature vector space, where the clusters of feature vectors of each class are compact and well separated from each other. For this purpose, we computed the covariance between the pre-processed channels of each subwindow observation. Together with the covariance, we also used the following functions for feature extraction: band power, mean frequency, occupied bandwidth, mean absolute value, and waveform length. The input of these five functions was the raw EMG signal observed through each sub-window. Below, we describe briefly each function used for feature extraction:

- **Covariance:** This feature is used to measure how the values of an EMG channel vary with respect to the values of other seven channels. The covariance is calculated between each pair of channels, and it returns a vector $\sigma$ composed of $\binom{8}{2} = 28$ elements.
- **Band power:** This function represents the mean power of the EMG signal. It calculates the average power for each channel and returns a vector $bp$ composed of eight elements.
- **Mean frequency:** This feature is defined as the normalized mean frequency of the signal power spectrum in the time domain. The result of applying this function is a vector $mf$ composed of eight elements.
- **Occupied bandwidth:** This function calculates the difference between two frequencies which enclose 99% of the power of the input signal. The result is a vector $obw$ composed of eight elements.
- **Mean Absolute Value:** This function is the average of the absolute value of the EMG signal amplitude. It is a popular feature used in EMG hand gesture recognition applications [40]. The vector $mav$ has eight elements.
- **Waveform length:** This feature represents the cumulative length of waveform over time [41] and returns a vector $wl$ of eight elements.

![Figure 5. EMG sliding window acquisition with 66 EMG points and six non-overlapping sub-windows.](image)
All vectors described above were concatenated as follows: \( \mathbf{x} = [x_{fr}, x_{bp}, x_{mf}, x_{obw}, x_{mav}, x_{wl}] \).
This new vector contains 68 elements and will be used as input of the classification module.

**Classification:** To design the classification module, we used a feed-forward artificial neural network (ANN) because it includes universal approximators (i.e., ANN with a continuous squashing function in the output is capable of approximating any Borel function, provided sufficiently many hidden units) [42]. The ANN used for this work contains three layers: input, hidden, and output. The input layer contains 68 nodes, each of which takes one component of the feature vector \( \mathbf{x} \). The hidden layer is composed of 500 neurons with the ReLu activation function. The output layer has six neurons, with the softmax activation function, where each neuron estimates the conditional probability \( P(Y | \mathbf{x}) \) that the feature vector \( \mathbf{x} \) belongs to the class \( Y \in \{ \text{wave in}, \text{wave out}, \text{fist}, \text{fingers spread}, \text{double Tap}, \text{relax} \} \). For training the ANN, we minimized a loss function that is the sum of the weights (regularization) plus the cross entropy between the six probabilities predicted by the network and the one-hot encoding of the actual label of the feature vector \( \mathbf{x} \). For minimizing the cross entropy, we used 500 epochs of the Conjugate Gradient Backpropagation with a Polak–Ribiére Updates algorithm that adjusts the step size at each iteration [43].

The number of training samples for every individual model is the total number of sub-window observations from the 130 EMGs and would be around 11,830 (i.e., \( 130 \times 91 \)). Each user has 130 EMGs, and each EMG includes 1000 points. These points can be divided into around 91 sub-window observations of 11 points. It is worth mentioning that, for each user (in both training and testing user groups), the same network architecture was trained using only his/her EMGs. This implies that, during the training process, an individual adaptation is carried out, and it is encoded in the weights of his/her corresponding network. To reduce the chance of model overfitting, we use regularization by weight decay (with \( \lambda = 10^{-5} \)) similarly as proposed by Chen et.al. [44].

**Post-Processing:** In this step, in order to refine the decision returned by the classification module, two steps were applied: reduction and filtering atypical values.

- **Reduction:** The label for a window observation, composed of \( n = 66 \) points, was defined based on the labels of its six sub-window observations by applying a majority voting scheme. Ties were broken randomly. The label of each sub-window observation, composed of \( m = 11 \) points, was obtained using the steps described above. In this way, for each EMG, we obtained a vector of labels \( \hat{l} = (\hat{l}_1, \ldots, \hat{l}_p) \), where \( p \) is approximately equal to 15. In addition to the vector of labels, we also obtained a vector of times \( \hat{t} = (\hat{t}_1, \ldots, \hat{t}_p) \), where the time \( \hat{t}_i \) indicates the instant where the label \( \hat{l}_i \) was predicted, with \( i = 1, 2, \ldots, p \).

- **Filtering atypical values:** This step is based on future and past labels. This step verifies two possible scenarios:
  - If the value of \( \hat{l}_{i-1} \) is equal to the value of \( \hat{l}_{i+1} \), then the value of \( \hat{l}_i \) should be replaced with the value of \( \hat{l}_{i-1} \).
  - If the values of \( \hat{l}_{i-1}, \hat{l}_i \) and \( \hat{l}_{i+1} \) are different to each other, then the value of \( \hat{l}_i \) should be replaced with the value of \( \hat{l}_{i-1} \).

The selection of a baseline model was achieved after the evaluation of different hyper-parameters combinations as can be observed in the Appendix A, and more specifically in Appendix A.1. The chosen baseline model was selected as the hyper-parameter combination with the best recognition accuracy results over the validation tests. Using this baseline model, an analysis of the effect and importance of the pre-processing and post-processing modules is carried out in Appendix A.2. From both of these analyses, we present in Table 1 a summary of the hyper-parameters on each module that best fit our data-set for the proposed hand gesture recognition model.
Table 1. Summary of the hyper-parameters in the proposed hand gesture recognition model.

| Module              | Hyper-Parameter       | Value          |
|---------------------|-----------------------|----------------|
| Data acquisition    | window size           | 66 points      |
| Data acquisition    | stride                | 66 points      |
| Data acquisition    | sub-windows length    | 11 points      |
| Data acquisition    | num sub-windows       | 6              |
| Pre-processing      | low-pass filter order | 2              |
| Pre-processing      | low-pass filter cutoff frequency | 1 Hz |
| Feature extraction  | number of features    | 68             |
| Classification      | ANN input             | 68 units       |
| Classification      | ANN hidden layer      | 500 units, RELU|
| Classification      | ANN output layer      | 6 classes, softmax |
| Classification      | Training algorithm    | Conjugate Gradient descent, \(\lambda = 10^{-5}\), 500 epochs |

2.3. Method for Correction of Sensor Orientation

Correction of orientation is required in EMG-based hand gesture recognition systems because a rotation of the sensor causes a shift in the sequence of features fed to a classifier. This shift can cause a significant reduction of the recognition accuracy. To illustrate this problem, let us assume that we have a very simple recognition system composed of only two modules: feature extraction and classification (Figure 6). This system takes as input EMGs of \(n \in \mathbb{Z}^+\) channels, where each channel is denoted by \(CH_i\), with \(i = 1, 2, \ldots, n\), and returns a label \(y\) corresponding to the gesture performed. The feature extractor maps an EMG to a feature vector \(x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n\), where \(x_i = f(CH_i)\), with \(i = 1, 2, \ldots, n\). The classifier \(g\) returns a label \(y = g(x_1, x_2, \ldots, x_n)\). Note that the result of the classifier depends on the sequence with which the features are arranged in the vector \(x\). If the EMG sensor is rotated, then \(x_i = f(CH_i)\), with \(i, j = 1, 2, \ldots, n\) and \(i \neq j\), except for rotations of \(\pm 360\) degrees, thus causing a shift in the sequence of features fed to the classifier, a shift which in turn reduces the accuracy of the recognition system.

![Figure 6. Illustration of the importance of correction of sensor orientation](image)

Now, we describe the proposed algorithm for the correction of orientation. The aim of this algorithm is to correct, by software, the rotation of the EMG sensor by arranging its channels with respect to the channel with the highest energy, using a set of synchronization gestures. By default, we assume that the sequence in which the pods of the EMG sensor are arranged is \(S = 0, 1, \ldots, 7\). Given a set of synchronization gestures \(G_{\text{sync}} = \{E_1, E_2, \ldots, E_m\}\), where \(E_i \in [-1, +1]^{k \times 8}\), with \(k \in \mathbb{Z}^+\) and \(i = 1, 2, \ldots, m\), and the sequence \(S\), the algorithm for the correction of orientation consists of the following steps:
For each EMG $E_i \in \mathcal{G}_{\text{sync}}$:

(a) Compute the envelope $\hat{E}_i$ of the EMG $E_i$. For this task, we first rectify the EMG by computing its absolute value and then filter the rectified EMG using a second-order Butterworth low-pass filter $\Psi$, with a cutoff frequency of 10 Hz, thus obtaining the signal $\hat{E}_i = \Psi(\text{abs}(E_i))$, where $\hat{E}_i \in [-1, +1]^{1 \times 8}$

(b) Compute the vector of energies $e_i = (e_{i,1}, e_{i,2}, ..., e_{i,8})$ for $\hat{E}_i$, where $e_{ij}$ denotes the energy of the $j$-th channel (i.e., column) of the signal $\hat{E}_i$, with $j = 1, 2, ..., 8$. The value of $e_{ij}$ is computed by trapezoidal numerical integration of the squared values of the $j$-th column of $\hat{E}_i$.

(c) Normalize the vector $e_i = (e_{i,1}, e_{i,2}, ..., e_{i,8})$ by dividing each of its values by the length $k$ of each channel of $E_i$, obtaining in this way the new vector of normalized energies $e_i^* = (e_{i,1}/k, e_{i,2}/k, ..., e_{i,8}/k)$.

2. Compute the component-wise sum of the normalized vectors of energies to obtain the vector $e^* = \sum_{i=1}^{m} e_i^*$.

3. Find the index $i^*$ of the component of the vector $e^*$ with the largest value.

4. Compute the new sequence $\tilde{S} = \tilde{s}_1, \tilde{s}_2, ..., \tilde{s}_8$ in which the pods of the EMG sensor should be arranged, where $\tilde{s}_j = \text{mod}(s_j + i^*, 8)$, where $s_j \in S$, with $j = 1, 2, ..., 8$.

Selecting the gesture for synchronization is a key aspect for the correction of orientation to be successful. In this work, we selected the gesture for synchronization by testing each of the five gestures considered here using 20 users from the training set. For each of these 20 training users and for each candidate gesture for synchronization, we selected randomly four EMGs out of the 25 training EMGs of that candidate. With these four EMGs for synchronization, we followed the procedure stated for experiment 4, which is described in the next section. In this way, we estimated the recognition accuracy for each of the five candidate gestures for synchronization over 20 training users: wave in (83.68%), wave out (96.76%), fist (91.80%), fingers spread (83.96%), and double tap (84.12%). Therefore, based on these recognition accuracies, we selected wave out as the synchronization gesture for this work.

The way how, in practice, the recognition model combined with the algorithm for correcting the orientation is as follows. Here, we assume that the recognition model is already trained. After wearing the bracelet, a user has to perform the synchronization gesture a given number of times. Then, the algorithm for correcting the orientation finds an arrangement of the pods based on the synchronization EMGs. After this, the recognition systems use this new arrangement of the pods to rearrange the channels of an EMG before it passes to the recognition system. Note that the process of acquiring the synchronization EMGs must be performed each time a user wants to use the recognition system.

3. Experiments and Results

In this section, we describe the experiments that we executed in this work and also analyze the results obtained. For reproducing the results of the proposed model, we make publicly available the Matlab code [45] and the dataset used for this paper [46].

3.1. Experiments

For this work, a total of four experiments were executed combining rotation and no rotation of the bracelet with and without the correction of orientation proposed here. The details of each experiment are presented below.

Experiment 1: In this experiment, we evaluated the performance of the proposed algorithm when: (a) there is no rotation of the bracelet for training and testing, and (b) there is no correction of the orientation of the training and testing EMGs. This experiment simulates the ideal scenario where each user trains and tests the recognition model with EMGs acquired placing the bracelet always in exactly
the same orientation. For this experiment, all users wore the bracelet, for training and testing, in the orientation recommended by Thalmic Labs, the manufacturer of the Myo armband. This recommended position implies that a user should wear the bracelet in such a way that pod number 4 is always parallel to the palm of the hand.

**Experiment 2** : In this experiment, we evaluated the performance of the proposed algorithm when: (a) there is no rotation of the bracelet for training, but there is rotation for testing, and (b) there is no correction of the orientation of the training and testing EMGs. In this experiment, the training EMGs were acquired with the bracelet placed in the orientation recommend by the manufacturer of the Myo armband.

**Experiment 3** : In this experiment, we evaluated the performance of the proposed algorithm when: (a) there is no rotation of the bracelet for training, but there is rotation for testing, and (b) there is correction of the orientation of the training and testing EMGs. In this experiment, the training EMGs were also acquired with the bracelet placed in the orientation recommend by the manufacturer of the Myo armband.

**Experiment 4** : In this experiment, we evaluated the performance of the proposed algorithm when: (a) there is rotation of the bracelet for training and testing, and (b) there is correction of the orientation of the training and testing EMGs. This experiment simulates a general scenario where the user trains the recognition model using an orientation different from the orientation of testing. The experiments described above are summarized in Table 2.

### Table 2. Settings of each experiment executed in this work.

| Experiment Number | Rotation of the Bracelet | Correction of Orientation |
|-------------------|--------------------------|---------------------------|
|                   | Training | Testing | Training | Testing |
| 1                  | no       | no      | no       | no      |
| 2                  | no       | yes     | no       | no      |
| 3                  | no       | yes     | yes      | yes     |
| 4                  | yes      | yes     | yes      | yes     |

**Simulation of rotation of the bracelet** : All the EMGs of the dataset used for this work were acquired with the users wearing the bracelet in the orientation recommended by Thalmic Labs. To simulate rotations of the bracelet, we assume that, by default, the pods of the Myo armband are ordered according to the sequence $S = 0, 1, ..., 7$. Then, with uniform probability, we selected randomly a number $r$ from the set $\{-3, -2, -1, 0, +1, +2, +3, +4\}$. Then, we simulate the rotation of bracelet by computing the new sequence $\tilde{S} = \tilde{s}_1, \tilde{s}_2, ..., \tilde{s}_8$ of the pods, where $\tilde{s}_i = \text{mod}(s_i + r, 8)$, with $s_i \in S$ and $i = 1, 2, ..., 8$. Note that, in this way, we simulated rotations of the bracelet clockwise and counterclockwise in steps of 45 degrees.

**Protocol for model evaluation** : The recognition model was designed and trained using the EMGs of the 60 users from the training subset processed according to the settings established for each experiment. Likewise, for testing, we used the EMGs of the 60 users from the testing subset. For each user from the testing subset, the model was trained using the 25 training EMGs for each gesture to be recognized and the five EMGs of the hand relaxed. Once the model was trained, it was applied to the 25 testing EMGs for each gesture to be recognized, including the five testing EMGs of the hand relaxed, using a 66-point window, with a stride of 66 points. In this way, for each testing EMG, we obtained two vectors: a vector $\hat{l} = (\hat{l}_1, \ldots, \hat{l}_p)$, containing the labels predicted by the model, and a vector $\hat{t} = (\hat{t}_1, \ldots, \hat{t}_p)$ containing the instants of time to which these labels correspond.

For evaluating the recognition accuracy, we used two sets: $A$ and $B$. The set $A = \{(\hat{y}_i, t_i)\}_{i=1}^k$ is formed by the ordered pairs $(\hat{y}_i, t_i)$, which are obtained from the vectors $\hat{l}$ and $\hat{t}$, where $k$ denotes the number of time points of the EMG, with $k = 1000$ for this work. For example, given the predicted vectors of labels and times $\hat{l} = (1, 3, 3, 3, 2, 1, 1)$ and $\hat{t} = (10, 20, 30, 40, 50, 60, 70)$,
respectively, the first 10 pairs of the set $A$ will be $(1,1), (1,2), (1,3), ..., (1,10)$, the next 10 pairs will be $(3,11), (3,12), (3,13), ..., (3,20)$, and so on. The set $B = \{(y_i, t_i)\}_{i=1}^{k}$ is the ground truth for the recognition, where the pair $(y_i, t_i)$ contains the actual label $y_i$ for the time $t_i$ of the EMG, with $i = 1, 2, ..., k$. The values of $y_i$ and $t_i$ were obtained from the manual segmentation of the EMGs.

A recognition was considered successful if each of the following three conditions are met: (a) all the labels of the ordered pairs in the set $A^*$ must be equal to each other, where the set $A^*$ is obtained from the set $A$ by excluding from $A$ all the ordered pairs that contain the label "no gesture." (b) All the ordered pairs in the set $A^*$ must be connected, which means that for at least one permutation of the times of the pairs in $A^*$, the absolute value of the difference between adjacent times in that permutation must be exactly equal to 1; and (c) the overlapping factor computed using the following equation must be equal to or greater than a given threshold $\rho$:

$$\rho = 2 \times \frac{|A^* \cap B^*|}{|A^*| + |B^*|}$$  \hspace{1cm} (1)

where the set $B^*$ is obtained from the set $B$ by excluding from $B$ all the pairs that contain the label "no gesture." For this work, we used a threshold $\rho = 0.25$. It is important to note that, for evaluating the recognition, we only processed the 25 testing EMGs of each of the five gestures of interest, thus excluding the class "no gesture." An illustration of the calculation of value $\rho$ through overlapping among ground-truth and the vector of predictions can be observed in Figure 7. At this point, it is worth emphasizing that classification and recognition are different concepts. Classification compares a predicted label with a ground true label of a given EMG sample, which allows for identifying the corresponding hand gesture class for such sample. On the other hand, recognition involves not only inferring the label of a given sample, but also requires calculating the instants of time where the gesture was performed. For this purpose, it is required to compare the vector of points related to the ground truth with the corresponding vector of predictions for a given EMG sample by using the overlapping factor $\rho$.

FIGURE 7. Representation of calculation process by using the overlapping factor among the ground truth and the vector of predictions. If the threshold $\rho > 0.25$, then the recognition process is considered successful.
The time response of the proposed model was evaluated using a desktop computer with a 4-core processor of 4 GHz and 16 GB of RAM. The time response obtained for the proposed model is the average of the times that takes the model to process each window observation of 66 points.

3.2. Results

The averages of the classification and recognition accuracies, and the time of processing using the EMGs from the testing subset are presented in Table 3. For experiments 3 and 4, we tested the algorithm for correction of the orientation using 1, 2, and 3 synchronization EMGs, which for this work correspond to EMGs of the gesture wave out.

Finally, in Table 4, we present the classification and recognition accuracies classified according to sex (female and male) and handedness (left and right-handed) for experiment 4, with four EGMs for synchronization. It is important to mention that the classes’ sex and handedness are not mutually exclusive. The time included in Table 4 corresponds to the time that the model takes to classify a 66-point window.

| Experiment Number | Classification Accuracy (%) | Recognition Accuracy (%) | Time (ms) |
|-------------------|-----------------------------|--------------------------|-----------|
| 1                 | 96.92 ± 2.14                | 95.32 ± 3.92             | 59.07 ± 13.64 |
| 2                 | 17.38 ± 8.06                | 8.03 ± 6.96              | 55.96 ± 11.44 |
| 3 (1 sync.)       | 81.54 ± 31.50               | 78.73 ± 34.06            | 64.61 ± 11.01 |
| 3 (2 sync.)       | 88.24 ± 24.42               | 86.11 ± 26.42            | 57.78 ± 8.78 |
| 3 (3 sync.)       | 88.62 ± 23.74               | 86.12 ± 26.13            | 51.77 ± 10.23 |
| 3 (4 sync.)       | 93.78 ± 14.90               | 91.93 ± 16.44            | 53.82 ± 10.76 |
| 4 (1 sync.)       | 83.01 ± 28.37               | 80.01 ± 31.87            | 67.46 ± 11.03 |
| 4 (2 sync.)       | 87.60 ± 24.40               | 85.52 ± 26.51            | 62.42 ± 7.53 |
| 4 (3 sync.)       | 94.65 ± 12.99               | 93.05 ± 14.59            | 52.43 ± 10.62 |
| 4 (4 sync.)       | 95.88 ± 9.07                | 94.40 ± 10.93            | 50.04 ± 5.58 |

4. Discussion

For all the experiments, the difference between the classification and the recognition accuracies is due to the way that these performances are computed. Let us remember that, for classification, the time in which a gesture is executed is not taken into account—unlike for recognition, where the time matters in the computation of the result.

The results presented in Table 3 show that the highest classification and recognition results are obtained in experiment 1, where we simulate an ideal case in which a user always wears the bracelet in exactly the same orientation both for training and testing. It is worth pointing out that the
results of experiment 1 are the targets for the results of the remaining three experiments. When, in experiment 2, we simulate the rotation of the sensor for testing, and when the recognition model does not include a correction of orientation, the classification and recognition accuracies hugely deteriorate, making the recognition model almost useless since a random model will achieve a classification accuracy of 17% for six classes. For the case of experiment 3, where there is rotation of the sensor, for testing and training, the EMGs are acquired in the recommended orientation for the Myo armband, and the correction of orientation is included for the training and testing EMGs. For experiment 3, the classification and recognition accuracies improve significantly with respect to the accuracies of experiment 2. The highest accuracies of experiment 3 are obtained with four synchronization EMGs. For the case of experiment 4, the training and testing EMGs are acquired in different orientations from the recommended orientation for the Myo armband, and the orientation of the sensor for training is different from the orientation of testing, and includes the correction of orientation for training and testing; we also observe an improvement of the accuracies with respect to the ones of experiment 2. Once again, among the accuracies of experiment 4, the highest results are obtained with four synchronization EMGs. The high standard deviations of the best accuracies of experiments 3 and 4 (both achieved with four synchronization EMGs) indicate that the recognition model with the correction of orientation proposed in this work has a performance that varies significantly from one user to another. This result could be caused due to a miss performance of the algorithm for the correction of orientation for some users.

The results presented in Table 4 allow us to understand better the recognition and classification accuracy of the proposed model under the configuration that was set for experiment 4, with four synchronization EMGs. For the case of the class sex that is composed of the categories male (with 45 testing users) and female (with 25 testing users), the model achieved recognition accuracies of \((95.89 \pm 2.80)\%\) and \((93.90 \pm 12.52)\%\), respectively. It is interesting to see that the proposed model produced, on average, better recognition accuracies for females than for males. However, the overlapping of the confidence intervals \(\mu \pm \sigma\) of the recognition accuracy for females and for males suggests that this difference might not be statistically significant. The standard deviation for the classification and recognition accuracies is significantly lower for females than for males. Regarding the class handedness, on average, better recognition accuracy was achieved for the two left-handed testing users who, according to our protocol for EMG acquisition, performed the gestures using their right hand. As in the case of the class sex, for the class handedness, we cannot conclude that there is statistically significant difference between the performance of our model with left-handed and right-handed users, since the confidence intervals \(\mu \pm \sigma\) of the recognition accuracies of these two categories overlap. In-depth research needs to be performed to analyze if there is a statistical difference between the classification and recognition accuracies for the classes sex and handedness.

Finally, regarding the processing time, the average times of all the experiments do not have significant variation among each other. This suggests that the algorithm for the correction of orientation is cheap in the processing time, making its use in real-time recognition systems possible, which should respond in less than 300 ms.

5. Conclusions and Future Work

In this work, we have presented a model for hand gesture recognition. Together with this model, we have also presented an algorithm for correcting, by software, the orientation of the EMG bracelet that is used to measure EMGs on the forearm for recognizing, in real time, five gestures of the hand: wave in, wave out, fist, fingers spread, and double tap. The proposed gesture recognition model is user-specific, which means that it is trained and tested with EMGs of the same user, but its architecture is the same for all users. For the architecture of the recognition model, we use a shallow feed-forward artificial neural network, which takes as input a feature vector composed of the covariances between the channels of an EMG and a bag of five functions applied to each channel of an EMG. The algorithm
for the correction of orientation is based on finding the index of the channel of maximum energy for a set of synchronization EMGs, which for this work corresponds to the gesture wave out.

The average classification and recognition accuracies of the proposed model are relatively high and have low standard deviations when the bracelet is placed in exactly the same orientation for training and testing. However, by simulating rotations of the bracelet for testing (which is very likely to occur in reality), we have demonstrated that the accuracies of the proposed model deteriorate so much that they behave almost as bad as a random recognition algorithm. The inclusion, as part of the recognition model, of the algorithm for correcting the orientation of the bracelet significantly improves the classification and recognition accuracies. This improvement of the accuracy increases on average by including more synchronization EMGs. In this work, using four synchronization EMGs allowed us to obtain average accuracies that are almost as good as when the bracelet is placed on exactly the same orientation for training and testing. However, the main drawback of the proposed model for gesture recognition is its high standard deviation when we include the algorithm for correcting the orientation of the EMG bracelet.

Even though the experiments and the results presented in this paper have been executed and obtained using EMGs measured with the Myo armband, the proposed model has the potential to be used with EMGs acquired using other sensors (i.e., bracelets) whose designs are similar to the Myo armband—for example, with the new EMG sensor gForce-pro.

Future work includes testing the proposed algorithm for correcting the orientation of the EMG bracelet for general gesture recognition models. Another future work includes studying statistically the difference in the recognition accuracy for the classes sex and handedness.

**Author Contributions:** Conceptualization, L.I.B.L., Á.L.V.C. and M.E.B.; investigation, L.I.B.L., Á.L.V.C. and M.E.B.; project administration, L.I.B.L., Á.L.V.C. and M.E.B.; resources, L.I.B.L., Á.L.V.C. and M.E.B.; supervision, L.I.B.L., Á.L.V.C. and M.E.B.; validation, L.I.B.L., Á.L.V.C. and M.E.B.; data-curation, L.I.B.L., Á.L.V.C.; software, M.E.B.; visualization, L.I.B.L., Á.L.V.C. and M.E.B.; methodology, M.E.B.; formal analysis, M.E.B.; funding acquisition, M.E.B.; writing original draft, L.I.B.L., Á.L.V.C. and M.E.B.; writing-review & editing, L.I.B.L., Á.L.V.C. and M.E.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** The authors gratefully acknowledge the financial support provided by the Escuela Politécnica Nacional (EPN) for the development of the research project “PIGR-19-07 Reconocimiento de gestos de la mano usando señales electromiográficas e inteligencia artificial y su aplicación para la implementación de interfaces humano—máquina y humano—humano”.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Appendix A. Model Parameters**

**Appendix A.1. Hyper-Parameters Tuning**

This Appendix A.1 provides details of the hyper-parameter tuning process that we performed to obtain the best possible model for our hand gesture recognition model. For this purpose, we consider the number of hidden units, the regularization parameter \( \lambda \), and the number of epochs (i.e., to analyze early stopping) as the tuning hyper-parameters on our model. We summarize in Table A1 the combinations tested over the validation user set (i.e., users 31–60), and the configuration used for testing is highlighted, which was the best configuration obtained based on the recognition metric. It is noticed that the best result is sorted by recognition accuracy.

In Table A2, we present the confusion matrix of the best obtained model which was selected as a baseline for this work. As mentioned before, the confusion matrix is a tool for analyzing classification results. It can be observed that the lowest sensitivity is on relax 46.67%; this means that several relax EMGs were classified incorrectly. On the other hand, the highest precision is obtained by the gesture fingers spread 95.45%. 

---

**Author Contributions:** Conceptualization, L.I.B.L., Á.L.V.C. and M.E.B.; investigation, L.I.B.L., Á.L.V.C. and M.E.B.; project administration, L.I.B.L., Á.L.V.C. and M.E.B.; resources, L.I.B.L., Á.L.V.C. and M.E.B.; supervision, L.I.B.L., Á.L.V.C. and M.E.B.; validation, L.I.B.L., Á.L.V.C. and M.E.B.; data-curation, L.I.B.L., Á.L.V.C.; software, M.E.B.; visualization, L.I.B.L., Á.L.V.C. and M.E.B.; methodology, M.E.B.; formal analysis, M.E.B.; funding acquisition, M.E.B.; writing original draft, L.I.B.L., Á.L.V.C. and M.E.B.; writing-review & editing, L.I.B.L., Á.L.V.C. and M.E.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** The authors gratefully acknowledge the financial support provided by the Escuela Politécnica Nacional (EPN) for the development of the research project “PIGR-19-07 Reconocimiento de gestos de la mano usando señales electromiográficas e inteligencia artificial y su aplicación para la implementación de interfaces humano—máquina y humano—humano”.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Appendix A. Model Parameters**

**Appendix A.1. Hyper-Parameters Tuning**

This Appendix A.1 provides details of the hyper-parameter tuning process that we performed to obtain the best possible model for our hand gesture recognition model. For this purpose, we consider the number of hidden units, the regularization parameter \( \lambda \), and the number of epochs (i.e., to analyze early stopping) as the tuning hyper-parameters on our model. We summarize in Table A1 the combinations tested over the validation user set (i.e., users 31–60), and the configuration used for testing is highlighted, which was the best configuration obtained based on the recognition metric. It is noticed that the best result is sorted by recognition accuracy.

In Table A2, we present the confusion matrix of the best obtained model which was selected as a baseline for this work. As mentioned before, the confusion matrix is a tool for analyzing classification results. It can be observed that the lowest sensitivity is on relax 46.67%; this means that several relax EMGs were classified incorrectly. On the other hand, the highest precision is obtained by the gesture fingers spread 95.45%.
Table A1. Validation tests for hyper-parameters tuning.

| Hidden Units | Lambda   | Epochs | Classification Accuracy (%) | Recognition Accuracy (%) |
|--------------|----------|--------|-----------------------------|--------------------------|
| 100          | $1 \times 10^{-5}$ | 300    | 96.95 ± 2.05                | 92.13 ± 5.32             |
| 250          | $1 \times 10^{-5}$ | 300    | 97.41 ± 1.9                 | 92.32 ± 5.48             |
| 500          | $1 \times 10^{-5}$ | 300    | 97.36 ± 2.03                | 92.37 ± 6.17             |
| 100          | $1 \times 10^{-4}$ | 300    | 97.13 ± 2.18                | 92.48 ± 5.06             |
| 100          | $1 \times 10^{-5}$ | 500    | 96.64 ± 1.94                | 92.48 ± 4.88             |
| 500          | $1 \times 10^{-4}$ | 300    | 97.31 ± 2.17                | 92.51 ± 6                |
| 250          | $1 \times 10^{-4}$ | 300    | 97.36 ± 1.98                | 92.53 ± 5.33             |
| 100          | $1 \times 10^{-4}$ | 500    | 96.87 ± 2.22                | 92.61 ± 4.35             |
| 250          | $1 \times 10^{-4}$ | 500    | 97.13 ± 1.75                | 92.72 ± 5.22             |
| 250          | $1 \times 10^{-5}$ | 500    | 96.85 ± 1.58                | 92.91 ± 5.36             |
| 500          | $1 \times 10^{-4}$ | 500    | 96.64 ± 2.16                | 93.15 ± 5.28             |
| 500          | $1 \times 10^{-5}$ | 500    | 96.87 ± 2.3                 | 93.25 ± 5.05             |

Table A2. Confusion matrix of the proposed model.

| Targets         | waveIn | waveOut | Fist  | fingersSpread | doubleTap | relax | Predictions Count (Precision %) |
|-----------------|--------|---------|-------|---------------|-----------|-------|---------------------------------|
| waveIn          | 747    | 1       | 2     | 0             | 1         | 8     | 759 98.42%                     |
| waveOut         | 2      | 746     | 0     | 8             | 4         | 11    | 771 96.76%                     |
| Fist            | 1      | 0       | 741   | 6             | 3         | 12    | 763 97.12%                     |
| fingersSpread   | 0      | 3       | 1     | 735           | 3         | 28    | 770 95.45%                     |
| doubleTap       | 0      | 0       | 6     | 1             | 739       | 21    | 767 96.35%                     |
| relax           | 0      | 0       | 0     | 0             | 0         | 70    | 70 100%                        |
| Targets Count   | 750    | 750     | 750   | 750           | 750       | 150   | 3900                           |
| (Sensitivity %) | 99.6%  | 99.47%  | 98.8% | 98%           | 98.53%    | 46.67% | 96.87%                       |

Appendix A.2. Analysis of Pre-Processing and Post-Processing Modules

In Table A3, we present the results with and without the pre-processing and/or the post-processing module. This evaluation was carried over the validation user set following the baseline obtained previously.

Table A3. Results with and without pre-processing and post-processing.

| Pre-Processing | Post-Processing | Classification Accuracy (%) | Recognition Accuracy (%) |
|----------------|-----------------|-----------------------------|--------------------------|
| No             | No              | 96 ± 1.82                   | 59.57 ± 10.33            |
| Yes            | No              | 96.87 ± 2.3                 | 64.35 ± 10.25            |
| No             | Yes             | 96 ± 1.82                   | 93.12 ± 5.08             |
| Yes            | Yes             | 96.87 ± 2.3                 | 93.25 ± 5.05             |
In Table A3, it can be observed that classification results are similar in these tests (around 96%, see Tables A4–A6). On the other hand, recognition accuracy results show the effects and importance of the post-processing module increasing performance from 64.35% to 93.25%. With these results, we justify the pre and post-processing modules used in this hand gesture recognition model.

**Table A4.** Confusion matrix without pre-processing or post-processing.

| Targets     | Predictions Count | Precision % |
|-------------|-------------------|-------------|
| waveIn      | 745               | 95.03%      |
| waveOut     | 743               | 96.12%      |
| fist        | 745               | 97.39%      |
| fingersSpread | 734            | 95.45%      |
| doubleTap   | 739               | 95.85%      |
| relax       | 38                | 100%        |

**Table A5.** Confusion matrix including pre-processing but excluding post-processing.

| Targets     | Predictions Count | Precision % |
|-------------|-------------------|-------------|
| waveIn      | 747               | 98.42%      |
| waveOut     | 746               | 96.76%      |
| fist        | 741               | 97.12%      |
| fingersSpread | 735            | 95.45%      |
| doubleTap   | 739               | 96.35%      |
| relax       | 70                | 100%        |

**Table A6.** Confusion matrix including both pre-processing and post-processing.

| Targets     | Predictions Count | Precision % |
|-------------|-------------------|-------------|
| waveIn      | 747               | 98.42%      |
| waveOut     | 746               | 96.76%      |
| fist        | 741               | 97.12%      |
| fingersSpread | 735            | 95.45%      |
| doubleTap   | 739               | 96.35%      |
| relax       | 70                | 100%        |
Table A6. Confusion matrix excluding pre-processing including post-processing.

| Targets            | Predictions Count (Precision %) |
|--------------------|----------------------------------|
|                    | waveIn  | waveOut | fist  | fingersSpread | doubleTap | relax  |
| waveIn             | 745     | 4       | 2     | 2             | 4         | 27     |
|                    |         |         |       |               |           | 784    |
|                    |         |         |       |               |           | 95.03% |
| waveOut            | 3       | 743     | 0     | 7             | 2         | 18     |
|                    |         |         |       |               |           | 773    |
|                    |         |         |       |               |           | 96.12% |
| fist               | 1       | 0       | 745   | 7             | 0         | 12     |
|                    |         |         |       |               |           | 765    |
|                    |         |         |       |               |           | 97.39% |
| fingersSpread      | 1       | 3       | 2     | 734           | 5         | 24     |
|                    |         |         |       |               |           | 769    |
|                    |         |         |       |               |           | 95.45% |
| doubleTap          | 0       | 0       | 1     | 0             | 739       | 31     |
|                    |         |         |       |               |           | 771    |
|                    |         |         |       |               |           | 95.85% |
| relax              | 0       | 0       | 0     | 0             | 0         | 38     |
|                    |         |         |       |               |           | 38     |
|                    |         |         |       |               |           | 100%   |
| Targets Count      | 750     | 750     | 750   | 750           | 750       | 150    |
| (Sensitivity %)    | 99.33%  | 99.07%  | 99.33%| 97.87%        | 98.53%    | 25.33% |
|                    | 3900    |         |       |               |           |        |
|                    | 96%     |         |       |               |           |        |

References

1. Young, S.; Stephens-Fripp, B.; Gillett, A.; Zhou, H.; Alici, G. Pattern Recognition for Prosthetic Hand User’s Intentions using EMG Data and Machine Learning Techniques. In Proceedings of the 2019 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM), Hong Kong, China, 8–12 July 2019; pp. 544–550.
2. Jiménez, L.A.E.; Benalcázar, M.E.; Sotomayor, N. Gesture recognition and machine learning applied to sign language translation. In Proceedings of the VII Latin American Congress on Biomedical Engineering CLAIB 2016, Bucaramanga, Santander, Colombia, 26–28 October 2016; Springer: Berlin/Heidelberg, Germany, 2017; pp. 233–236.
3. Saggio, G.; Sbernini, L. New scenarios in human trunk posture measurements for clinical applications. In Proceedings of the 2011 IEEE International Symposium on Medical Measurements and Applications, Bari, Italy, 30–31 May 2011; pp. 13–17.
4. Cavrini, F.; Bianchi, L.; Quitadamo, L.R.; Saggio, G. A fuzzy integral ensemble method in visual P300 brain-computer interface. Comput. Intell. Neurosci. 2016, 2016, 9845980. [CrossRef] [PubMed]
5. León, M.A.; Romero, P.A.; Quevedo, W.X.; Arteaga, O.B.; Terán, C.; Benalcázar, M.E.; Andaluz, V.H. Virtual Rehabilitation System for Fine Motor Skills Using a Functional Hand Orthosis. In Proceedings of the International Conference on Augmented Reality, Virtual Reality and Computer Graphics, Otranto, Italy, 24–27 June 2018; Springer: Berlin/Heidelberg, Germany, 2018; pp. 78–94.
6. Saggio, G.; Bisegna, P.; Latessa, G.; Bocchetti, S. Mechanical modeling of bend sensors exploited to measure human joint movements. In Proceedings of the 2009 IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks & Workshops, Kos, Greece, 15–19 June 2009; pp. 1–4.
7. Jaramillo-Yáñez, A.; Benalcázar, M.E.; Mena-Maldonado, E. Real-Time Hand Gesture Recognition Using Surface Electromyography and Machine Learning: A Systematic Literature Review. Sensors 2020, 20, 2467. [CrossRef] [PubMed]
8. Chung, E.A.; Benalcázar, M.E. Real-Time Hand Gesture Recognition Model Using Deep Learning Techniques and EMG Signals. In Proceedings of the 2019 27th European Signal Processing Conference (EUSIPCO), A Coruna, Spain, 2–6 September 2019; pp. 1–5.
9. Qi, J.; Jiang, G.; Li, G.; Sun, Y.; Tao, B. Surface 539 EMG hand gesture recognition system based on PCA and 540 GRNN. Neural Comput. Appl. 2019. [CrossRef]
10. Tavakoli, M.; Benussi, C.; Lopes, P.A.; Osorio, L.B.; de Almeida, A.T. Robust hand gesture recognition with a double channel surface EMG wearable armband and SVM classifier. *Biomed. Signal Process. Control* **2018**, *46*, 121–130. [CrossRef]

11. Zhang, Z.; Yang, K.; Qian, J.; Zhang, L. Real-Time Surface EMG Pattern Recognition for Hand Gestures Based on an Artificial Neural Network. *Sensors* **2019**, *19*, 3170. [CrossRef] [PubMed]

12. Abhishek, K.S.; Qubeley, L.C.K.; Ho, D. Glove-Based Hand Gesture Recognition Sign Language Translator Using Capacitive Touch Sensor. In Proceedings of the 2016 IEEE International Conference on Electron Devices and Solid-State Circuits (EDSSC), Hong Kong, China, 3–5 August 2016; pp. 334–337.

13. Wang, N.; Lao, K.; Zhang, X. Design and Myoelectric Control of an Anthropomorphic Prosthetic Hand. *J. Bionic Eng.* **2017**, *14*, 47–59. [CrossRef]

14. Palmeri, M.; Vella, F.; Infantino, I.; Gaglio, S. Sign Languages Recognition Based on Neural Network Architecture. In *Smart Innovation, Systems and Technologies*; Springer: Cham, Switzerland, 2018; Volume 76, pp. 109–118. [CrossRef]

15. Benatti, S.; Rovere, G.; Bossé, J.; Montagna, F.; Farella, E.; Glaser, H.; Schonle, P.; Burger, T.; Fateh, S.; Huang, Q.; et al. A Sub-10mW Real-Time Implementation for EMG Hand Gesture Recognition Based on a Multi-Core Biomedical SoC. In Proceedings of the 2017 7th International Workshop on Advances in Sensors and Interfaces, IWASI 2017, Vieste, Italy, 15–16 June 2017, pp. 139–144.

16. Kim, S.Y.; Han, H.G.; Kim, J.W.; Lee, S.; Kim, T.W. A hand gesture recognition sensor using reflected impulses. *IEEE Sensors J.* **2017**, *17*, 2975–2976. [CrossRef]

17. Zhang, X.; Chen, X.; Li, Y.; Lantz, V.; Wang, K.; Yang, J. A framework for hand gesture recognition based on accelerometer and EMG sensors. *IEEE Trans. Syst. Man Cybern. Part Syst. Hum.* **2011**, *41*, 1064–1076. [CrossRef]

18. Iyer, D.; Mohammad, F.; Guo, Y.; Al Safadi, E.; Smiley, B.J.; Liang, Z.; Jain, N.K. Generalized Hand Gesture Recognition for Wearable Devices in IoT: Application and Implementation Challenges. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Cham, Switzerland, 2016; Volume 9729, pp. 346–355. [CrossRef]

19. Moschetti, A.; Fiorini, L.; Esposito, D.; Dario, P.; Cavallo, F. Recognition of Daily Gestures with Wearable Inertial Rings and Bracelets. *Sensors* **2016**, *16*, 1341. [CrossRef]

20. Singh, J.; Roy, A.; Laskar, R.H. Dynamic hand gesture recognition using vision-based approach for human–computer interaction. *Neural Comput. Appl.* **2018**, *29*, 1129–1141. [CrossRef]

21. Luh, G.C.; Lin, H.A.; Ma, Y.H.; Yen, C.J. Intuitive muscle-gesture based robot navigation control using wearable gesture armband. In Proceedings of the 2015 International Conference on Machine Learning and Cybernetics (ICMLC), Guangzhou, China, 12–15 July 2015; Volume 1, pp. 389–395.

22. Benalcazar, M.E.; Mutoche, C.; Zea, J.A.; Jaramillo, A.G.; Anchundia, C.E.; Zambrano, P.; Segura, M.; Benalcazar Palacios, F.; Perez, M. Real-Time Hand Gesture Recognition Using the Myo Armband and Muscle Activity Detection. In Proceedings of the 2017 IEEE 2nd Ecuador Technical Chapters Meeting (ETCM 2017), Salinas, Ecuador, 16–20 October 2017; Volume 2017-Janua, pp. 1–6. [CrossRef]

23. Fortunato, A.; King, J.; Bonneau, P.J. Data Display, Acquisition and Feedback System for Biomedical Experiments. 2006. Available online: https://digitalcommons.wpi.edu/mqp-all/340 (accessed on 26 November 2020)

24. Weiss, L.D.; Weiss, J.M.; Silver, J.K. *Easy EMG: A Guide to Performing Nerve Conduction Studies and Electromyography*; Elsevier: Amsterdam, The Netherlands, 2015; p. 304.

25. Vimos, V.H.; Benalcazar, M.; Oha, A.F.; Cruz, P.J. A Novel Technique for Improving the Robustness to Sensor Rotation in Hand Gesture Recognition Using sEMG. In Proceedings of the International Conference on Computer Science, Electronics and Industrial Engineering (CSEI), Ambato, Ecuador, 28–31 October 2019; Springer: Berlin/Heidelberg, Germany, 2019; pp. 226–243.

26. Boschmann, A.; Plattner, M. Reducing classification accuracy degradation of pattern recognition based myoelectric control caused by electrode shift using a high density electrode array. In Proceedings of the 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Diego, CA, USA, 28 August–1 September 2012; pp. 4324–4327.
27. Zhang, Y.; Chen, Y.; Yu, H.; Yang, X.; Lu, W.; Liu, H. Wearing-independent Hand Gesture Recognition Method Based on EMG Armband. Pers. Ubiquitous Comput. 2018, 22, 511–524. [CrossRef]
28. Xu, Z.; Shen, L.; Qian, J.; Zhang, Z. Advanced Hand Gesture Prediction Robust to Electrode Shift with an Arbitrary Angle. Sensors 2020, 20, 1113. [CrossRef][PubMed]
29. Wahid, M.F.; Tafreshi, R.; Langari, R. A Multi-Window Majority Voting Strategy to Improve Hand Gesture Recognition Accuracies Using Electromyography Signal. IEEE Trans. Neural Syst. Rehabil. Eng. 2019, 28, 427–436. [CrossRef][PubMed]
30. Tavakoli, M.; Benussi, C.; Lourenco, J.L. Single Channel Surface EMG Control of Advanced Prosthetic Hands: A Simple, Low Cost and Efficient Approach. Expert Syst. Appl. 2017, 79, 322–332. [CrossRef]
31. Liang, S.; Wu, Y.; Chen, J.; Zhang, L.; Chen, P.; Chai, Z.; Cao, C. Identification of Gesture Based on Combination of Raw sEMG and sEMG Envelope Using Supervised Learning and Univariate Feature Selection. J. Bionic Eng. 2019, 16, 647–662. [CrossRef]
32. Barros, P.; Maciel-Junior, N.T.; Fernandes, B.J.; Bezerra, B.L.; Fernandes, S.M. A Dynamic Gesture Recognition and Prediction System Using the Convexity Approach. Comput. Vis. Image Underst. 2017, 155, 139–149. [CrossRef]
33. Duda, R.O.; Hart, P.E.; Stork, D.G. Pattern Classification; John Wiley & Sons: Hoboken, NJ, USA, 2012.
34. Ullah, A.; Ali, S.; Khan, I.; Khan, M.A.; Faizullah, S. Effect of Analysis Window and Feature Selection on Classification of Hand Movements Using EMG Signal. In Intelligent Systems and Applications; Arai, K., Kapoor, S., Bhavia, R., Eds.; Springer International Publishing: Cham, Switzerland, 2020; pp. 400–415.
35. Berrneo-Calderon, J.; Velasco, M.A.; Rojas, J.L.; Villarreal-Lopez, J.; Galvis Resrepo, E. Movement Control System for a Transradial Prosthesis Using Myoelectric Signals. In AETA 2019—Recent Advances in Electrical Engineering and Related Sciences: Theory and Application; Cortes Tobar, D.F., Hoang Duy, V., Trong Dao, T., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 273–282.
36. Liu, H.; Wang, L. Gesture Recognition for Human-Robot Collaboration: A Review. Int. J. Ind. Ergon. 2018, 68, 355–367. [CrossRef]
37. Lu, L.; Mao, J.; Wang, W.; Ding, G.; Zhang, Z. A Study of Personal Recognition Method Based on EMG Signal. IEEE Trans. Biomed. Circuits Syst. 2020, 14, 681–691. [CrossRef]
38. Jabbari, M.; Khushaba, R.N.; Nazarpour, K. EMG-Based Hand Gesture Classification with Long Short-Term Memory Deep Recurrent Neural Networks. In Proceedings of the 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Montreal, QC, Canada, 20–24 July 2020; pp. 3302–3305.
39. Reaz, M.B.I.; Hussain, M.S.; Mohd-Yasin, F. Techniques of EMG signal analysis: detection, processing, classification and applications. Biomed. Online Process. 2006, 8, 11–35. [CrossRef]
40. Too, J.; Abdullah, A.R.; Saad, N.M. Classification of Hand Movements based on Discrete Wavelet Transform and Enhanced Feature Extraction. Int. J. Adv. Comput. Sci. Appl. 2019, 10. [CrossRef]
41. Too, J.; Abdullah, A.R.; Zawawi, T.T.; Saad, N.M.; Musa, H. Classification of EMG signal based on time domain and frequency domain features. Int. J. Hum. Technol. Interact. (IJHaTI) 2017, 1, 25–30.
42. Castro, J.; Mantas, C.; Benítez, J. Neural networks with a continuous squashing function in the output are universal approximators. Neural Netw. 2000, 13, 561–563. [CrossRef]
43. Al-Bayati, A.; Abbo, K.K.; Saleh, I.A. Conjugate Gradient Back-propagation with Modified Polack-Rebiere updates for training feed forward neural network. Iraqi J. Stat. Sci. 2011, 11, 164–173. Available online: https://stats.mosuljournals.com/article_27897_02399b3781a5ae6ed5cb753db5641bb3.pdf (accessed on 26 November 2020).
44. Chen, L.; Fu, J.; Wu, Y.; Li, H.; Zheng, B. Hand Gesture Recognition Using Compact CNN via Surface Electromyography Signals. Sensors 2020, 20, 672. [CrossRef][PubMed]
45. Dataset EMG-EPN-120. Available online: https://laboratorio-ia.epn.edu.ec/en/resources/dataset/2020_emg_dataset_120 (accessed on 26 November 2020).
46. laboratorioAI/2020_ROT_SPECIFIC_ANN_EMG_120: Code Example for the Paper “A User-Specific Hand Gesture Recognition Model Based on Feed-Forward Neural Networks, EMGs and Correction of Sensor Orientation”. Available online: https://github.com/laboratorioAI/2020_ROT_SPECIFIC_ANN_EMG_120 (accessed on 26 November 2020).

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations. © 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).