

# Classification of Different Shoulder Girdle Motions for Prosthesis Control Using a Time-Domain Feature Extraction Technique

Huda M. Radha<sup>1</sup>, Alia K. Abdul Hassan<sup>2</sup> and Ali H. Al-Timemy<sup>3</sup>

<sup>1</sup>Department of Computer Science, College of Science, University of Baghdad, Baghdad, Iraq

<sup>2</sup>Department of Computer Science, College of Science, University of Technology, Baghdad, Iraq

<sup>3</sup>Department of Biomedical Engineering, Al-Khwarizmi College of Engineering, University of Baghdad, Baghdad, Iraq

**Abstract**—The upper limb amputation exerts a significant burden on the amputee, limiting their ability to perform everyday activities, and degrading their quality of life. Amputee patients' quality of life can be improved if they have natural control over their prosthetic hands. Among the biological signals, most commonly used to predict upper limb motor intentions, surface electromyography (sEMG), and axial acceleration sensor signals are essential components of shoulder-level upper limb prosthetic hand control systems. In this work, a pattern recognition system is proposed to create a plan for categorizing high-level upper limb prostheses in seven various types of shoulder girdle motions. Thus, combining seven feature groups, which are root mean square, four-order autoregressive, wavelength, slope sign change, zero crossing (ZC), mean absolute value, and cardinality. In this article, the time-domain features were first extracted from the EMG and acceleration signals. Then, the spectral regression (SR) and principal component analysis dimensionality reduction methods are employed to identify the most salient features, which are then passed to the linear discriminant analysis (LDA) classifier. EMG and axial acceleration signal datasets from six intact-limbed and four amputee participants exhibited an average classification error of 15.68 % based on SR dimensionality reduction using the LDA classifier.

**Index Terms**—Biosignal analysis, Dimensionality reduction, LDA classifier, Time domain

## I. INTRODUCTION

The amputation of an upper limb is a common problem that affects people worldwide. The root causes are many and include vascular disorders such as trauma from

accidents, diabetes, and regional conflict-associated traumatic amputations (Choo, Kim and Chang, 2022). The amputee's arm is incomplete if the shoulder has been disarticulated or the upper limb has been amputated below the elbow (Jang, et al., 2011). This means that the amputee cannot use the prosthetic arm because of a lack of muscle power. High-level amputations can be replaced with a body-powered prosthetic hand powered by mechanical cables or a very modern (i.e., TMR) surgical procedure requiring advanced intervention (Craelius, 2021).

Only a few studies have determined that a shoulder girdle electric prosthesis is the newest biological alternative to a limb. For this process to be successful, signals must be transmitted based on the absence of remanent and the shape of the scapula. The focus should therefore be on finding a way to solve the motion detection problem so that adequate prostheses function well as part of the control framework, as the task of the electromagnetic prosthesis is to send the necessary operational signals to detect shoulder motions (Sharba, 2020, Nsugbe et al., 2022).

Rivela, et al., 2015, employed a pattern recognition (PR) system to identify eight healthy people's shoulder motions into nine different groups. The data were segmented using a window length of  $L = 500$  ms and an increment of  $I = 62$  ms as a starting point. Their analysis used waveform length (WL) and root mean square (RMS) as features. The linear discriminant analysis (LDA) was used as a classifier in this research, and the percentage of correctly classified patterns was 92.1 %. Jiang, et al., 2020, examined the ability to build EMG signals from 12 muscles for the upper arm motion pattern identification. In this study, different muscles near the shoulder were shown to produce varied grips. The cross-subjects convolution neural network (CNN) model was employed in this research, and the obtained accuracy was 79.64% in recognizing motion patterns.

Analysis of the data collecting sequence used by Sharba, Wali and Timemy, 2020, has been conducted in another study carried out by Sharba, Wali and Timemy, 2020. This utilized



dataset can be employed for various purposes, including enhancing the operation of prosthetic hands. In recent years, there has been much interest in detecting the intention to move the upper limbs. Forearm muscle activity recordings were used in pattern recognition techniques to identify hand and wrist movements. However, it is impossible to infer the coordinated motion of the body from those signals alone. As a result, the actions of the prosthesis may appear unnatural when viewed as part of the total body, and a dynamic connection between the user and the prosthesis is inconceivable.

A higher-level amputee cannot use these systems because they rely only on contractions of the forearm muscle to generate energy (Li, et al., 2021). However, the reason behind using the time domain analysis for feature extraction is that the time series is a collection of data points that typically include a series of measurements taken at distinct time intervals. A time series analysis statistical technique is designed to examine such measures for feature extraction, model identification, parameter estimation, model validation, and forecasting. A time series model or some statistical aspects of time series data can be recovered as the key damage-sensitive features using the effective feature extraction technique known as time series analysis. In time series analysis, choosing the right model class is crucial. The kind and nature of the time series data and the availability or lack thereof of input or excitation data all affect how this process works (Pulliam et al., 2011). Time-invariant linear models are the best options for feature extraction when vibration measurements are linear and stationary. The most important and fundamental problem in time series modeling is choosing an appropriate and precise order that enables the target model to produce uncorrelated residuals (Entezami, 2021).

The main contributions of the present research are summarized below:

1. For each subject, we combined the seven feature parameters, which are the RMS, four-order autoregressive (AR), WL, slope sign change (SSC), zero crossing (ZC), mean absolute value (MAV), and cardinality to extract 60-dimensional feature vectors. This was accomplished by combining the seven feature groups mentioned previously.
2. The effects of SR and PCA dimensionality reduction algorithms in error analysis testing are then investigated.

The remainder of the work is structured as follows. The following section describes the proposed work plan in detail, and the theoretical background of time-domain feature extraction principles and dimensionality reduction is examined. In Section III, the experimental data and analysis are presented. Finally, Section IV concludes the work and gives suggestions for future investigations.

## II. THE PROPOSED WORK SCHEME

The proposed work scheme for classifying different shoulder girdle motions based on time-domain feature extraction contains signals pre-processing (cross-validation

and segmentation) stage, time-domain features extraction and dimensionality reduction, and then classification stages. First, as indicated in Fig. 1, the cross-validation of training and testing data is applied to pre-process the provided signals, followed by segmentation with overlap window size. After that, time-domain features are conducted. Finally, the LDA classifier is employed to detect seven more shoulder girdle motions for prosthesis control after the SR and PCA dimension reduction presentations. The classes of motion are shown in Fig. 2. The steps are described in more detail in the following subsections.

### A. Signal pre-processing

This section outlines the fundamental ideas that underlie the specific data collection techniques employed by Sharba, Wali and Timemy, 2020. The information was received from six participants with intact limbs (Fig. 3a) and four participants with amputees (Fig. 3b). Seven movements were selected for the classification of shoulder girdle motions: Elevation, depression, protraction, retraction, upward rotation, downward rotation, and rest. The extracted data were recorded from five EMG channels. A 3-axis accelerometer sensor (ADXL335) was also mounted on top of the shoulder (Fig. 4) to provide three acceleration signals (Fig. 5).

### Cross-validation

Cross-validation was utilized to exclude one trial from this study. To train the classifier in each fold throughout the optimization procedure, eight trials were used. The remaining trials were utilized to evaluate the classifier and calculate the classification error rate. This procedure was done 8 times (eight runs) to calculate the average error rates for the eight runs.

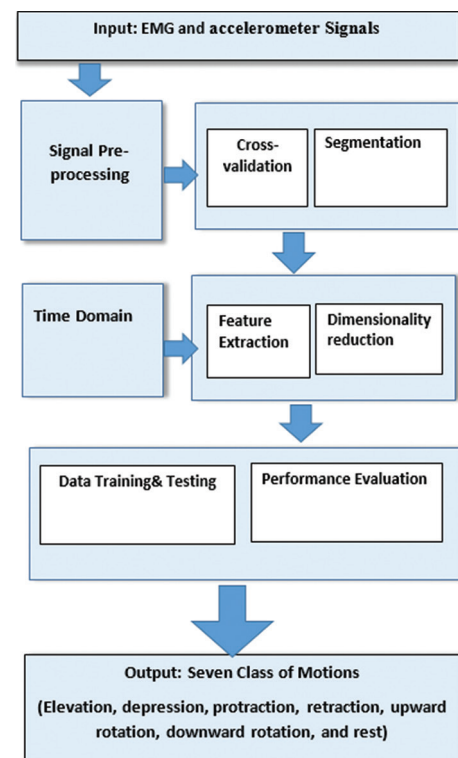


Fig. 1. The general design of the proposed work scheme.

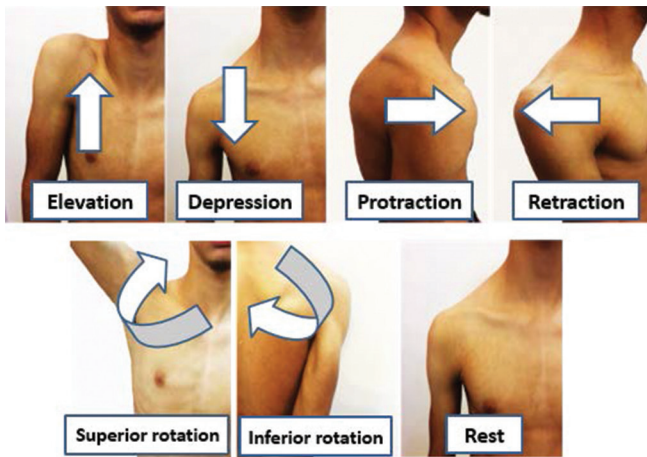


Fig. 2. Classes of shoulder girdle motions examples (Sharba, Wali and Timemy, 2020).

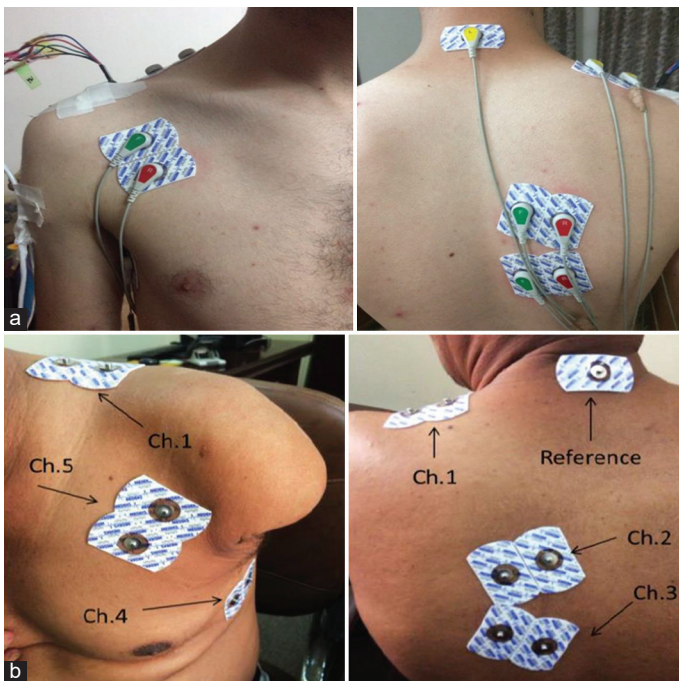


Fig. 3. (a) An example of an intact-limbed subject's specific electrode positions. (b) An example of an amputee subject's specific electrode positions.

*Segmentation*

The raw signals are sampled at a rate of 1 kHz. The data were also segmented using an overlapped segmentation approach with a window size of 150 ms and a window increment of 50 ms; Fig. 3 displays the EMG signals from five channels. Fig. 4 depicts the three acceleration signals channels.

*B. Time-domain-based feature extraction methods*

The time-domain (TD) features are commonly utilized in classification studies and are the most advantageous. The main advantage of TD features extraction is that they are easily extracted and produce excellent results compared to other approaches such as frequency- domain (FD) and time-

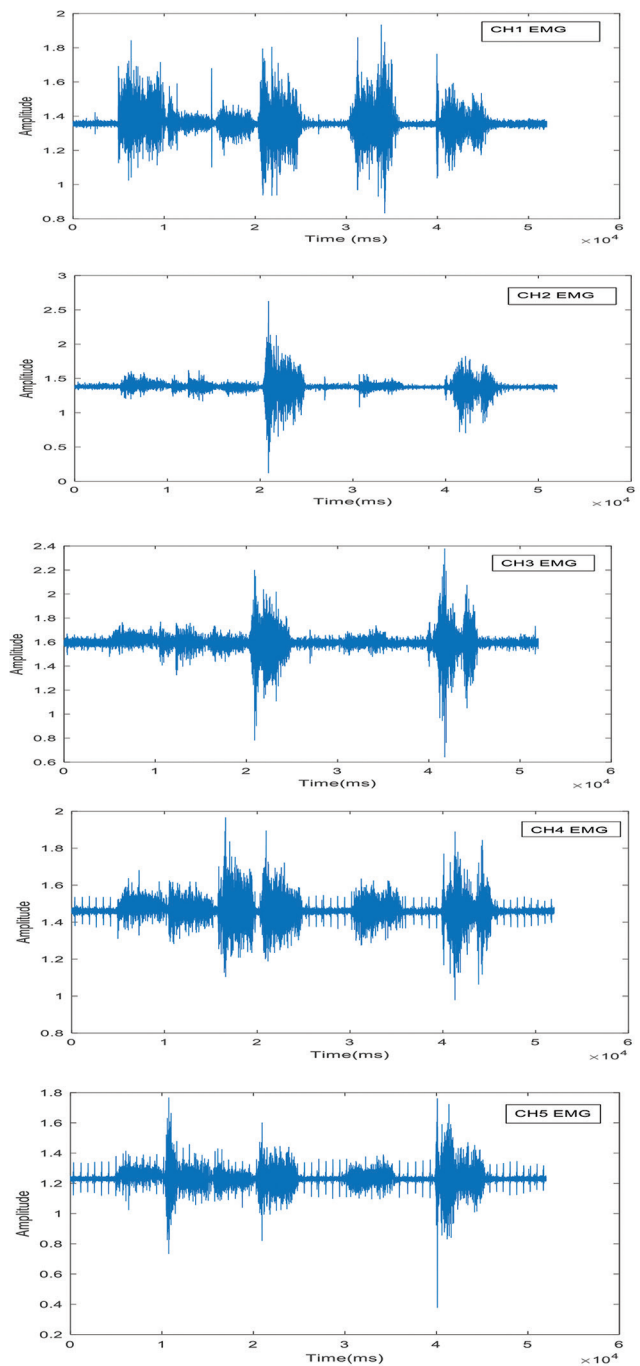


Fig. 4. Examples of EMG data (5 channels) in the pre-processing condition.

FD (TFD) features. The previous research demonstrated the use of TD, particularly in terms of its speed, ease of implementation, and absence of any required transformation (Padfield, 2022). However, the nature of the available signals determines whether or not the time or frequency features should be used. Examining the temporal features of the signal is ineffective in some situations. Thus, researchers must look at it from a new perspective (Boashash, Khan and Ben-Jabeur, 2015). The TDs fundamental problem is that features are formed from the signal's stationary properties. There may be large differences in the components when dealing with



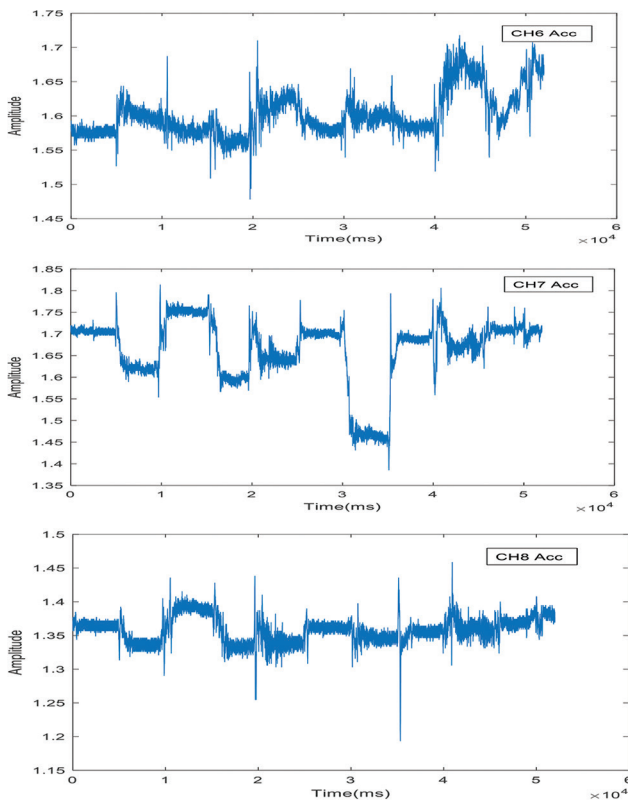


Fig. 5. Examples of 3-axis accelerometer sensor data (three channels) in the pre-processing condition

non-stationary signals like EMG that is typically collected during dynamic movements (Samuel, et al., 2018).

Since the characteristics of the TD have evolved entirely based on the amplitude of the EMG, it is particularly susceptible to noise captured during data collection. Differentiating between stratigraphic movements requires temporal and spectral information. Furthermore, this will be the main criterion for distinguishing between the FD and the TFD in the classification. However, not all of these qualities were used at the same time by a large number of researchers. To achieve efficient classification, important features were selected (Rampichini, et al., 2020). The time-domain features used in this study are as follows:

**RMS**

The RMS has been described as a Gaussian process analogous to the usual technique of muscular contraction. Furthermore, it mimics the standard deviation SD method (Phinyomark, Khushaba and Scheme, 2018). The RMS mathematical model has been represented as:

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

Where,  $xi$  represents the EMG signal, whereas  $N$  represents the sample number of a signal and  $RMS$  is a ubiquitous feature that utilizes tested seven TD features with the EMG analysis.

**MAV**

The MAV feature is ubiquitous and widely exploited by EMG signal researchers. The integrated EMG (IEMG) value

is derived from the corrected EMG moving average. This characteristic has been known by numerous different names, including ARV (i.e., average rectified value) (Phinyomark, Khushaba and Scheme, 2018). The formula for calculating MAV is as follows:

$$MAV = \frac{1}{N} \sum_{i=1}^N |xi| \tag{2}$$

Where,  $xi$  represents the signal of the EMG, whereas  $N$  represents the signal's sample number.

**ZC**

In temporal domain analysis, the word “ZC” refers to something linked to frequency. It is a measure of the spectral components that occur when the number of EMG magnitudes is more than the level of zero amplitude (Phinyomark, Khushaba and Scheme, 2018). To avoid low-voltage fluctuations or background noise, the threshold condition must be achieved, and the mathematical definition of this condition is as follows:

$$ZC = \sum_{i=1}^{n-1} [sgn(xi \times xi + 1) \cap |xi - xi + 1| \geq threshold] \tag{3}$$

$$sgn(x) = \begin{cases} 1, & \text{if } x \geq threshold \\ 0, & \text{otherwise} \end{cases}$$

Meanwhile, one characteristic measure the upward ZC split by the peaks (NP) (Al-Timemy, et al., 2015). Only their spectral moments can be used to examine this feature. The following equation can be used to express the corresponding quality as follows:

$$IF = \frac{ZC}{NP} = \frac{SM2}{\sqrt{SM0 \times SM4}} \tag{4}$$

**WL**

WL can be defined as an EMG complexity measure that describes the total summation of fluctuations throughout every signal segment. This attribute is also referred to as the wavelength, and it is referred to as the total value of absolute derivative signals (WAVE) (Rampichini, et al., 2020). The equation for calculating WL is expressed as:

$$WL = \log\left(\frac{\sum_{i=0}^{n-1} |\Delta x|}{\sum_{i=0}^{n-1} |\Delta 2x|}\right) \tag{5}$$

**SSC**

The SSC is a ZC trait with a recognizable character, which is how one may characterize it. Calculations are made in which differences in the slope sign are used to indicate information regarding signal frequency. Within their threshold function, positive and negative slope changes have been counted three times in sequence. As a direct consequence of this, background noise in the EMG will not be present (Toledo-Pérez, et al., 2019). The corresponding mathematical expression for this property is as follows:

$$SSC = \sum_{i=2}^{n-1} [f[(xi - xi_{i-1}) \times (xi - xi_{i+1})]] \tag{6}$$

$$f(x) = \begin{cases} 1, & \text{if } x \geq threshold \\ 0, & \text{otherwise} \end{cases}$$

The threshold parameter of this feature should have a value between 50 and 100 mV, as recommended (Tkach, Huang and Kuiken, 2010). Nevertheless, it may be different if the instrument's gain value and the background noise amount are not adjusted to the same level (Badr and Abdul-Hassan, 2021).

#### *Auto-regressive coefficient*

When the peak location of the EMG signal is known, the AR feature can be developed on a statistical method using spectral information. It is a prediction model in which the EMG signal is described as a linear combination of the previous samples  $x_{i-p}$  and white noise  $w_i$  (Phinyomark, Khushaba and Scheme, 2018). Several classification techniques make use of the AR coefficient as a feature vector. The following is the basic AR model:

$$x_i = \sum_{p=1}^P a_p x_{i-p} + w_i \quad (7)$$

Where,  $P$  represents the AR order at a certain autoregressive coefficient  $a_p$  between the fourth-order (*AR4*) and sixth-order (*AR6*) (Chen, Luo and Hua, 2021), research works have been proposed for ideal AR to utilize in the analysis of the EMG.

#### *Cardinality feature*

Cardinality is a recently proposed and potentially useful feature. It is denoted by counting the number of components in a group of items while excluding all comparable objects from the elements in that collection (Barton, 2020).

Compared to other commonly used individual features in the literature, cardinality was shown to be one feature that can achieve high levels of accuracy despite the variations in the sampling frequency, window segments, and number and type of movement classes (Barton, 2020). As a result, cardinality has become a fundamental focus of future research. Cardinality was shown to be one feature that can achieve high levels of accuracy despite the variations in the sampling frequency (Praciano, et al., 2021).

As was indicated earlier, researchers have made extensive use of all six of the TD features that were chosen for this study. The previous research, such as that conducted by Samuel, et al., 2018, and Phinyomark, Khushaba and Scheme, 2018, has proposed and used the coupling of TD features with AR. They showed that the feature could be utilized to improve EMG signal classification, which is a significant achievement.

Compared with any FD and TFD hand movement detection system, these features provide extremely high classification accuracy (Karheily, et al., 2022). This was the primary reason behind this research's decision to use the TD features described previously in the EMG data acquired for this investigation.

### *C. Dimensionality reduction methods*

In several numbers of data processing disciplines, such as data mining, machine learning, information retrieval, and pattern recognition, dimensionality reduction has been a significant challenge. The performance of supervised machine

learning algorithms degrades when they are presented with multiple features that are not essential for making accurate predictions of the desired output (i.e., prediction accuracy) (Zeng, et al., 2009). One of the most critical aspects of information discovery, machine learning, pattern recognition, and computer vision is isolating a few distinctive and useful characteristics. It is common practice to approach the resolution of this problem by employing methods that involve dimensionality reduction (Sarker, 2021). In this paper, we propose the two most essential methods for dimensionality reduction: Spectral regression (SR) and principal component analysis (PCA). Both of these algorithms are known for their ability to reduce the number of dimensions.

#### *SR*

Training sample embedding results are provided by conventional manifold learning techniques such as the locally linear embedding, Laplacian Eigenmap, and isomap. The out-of-sample extension is a challenge; therefore, SR develops a regression model that can avoid the Eigen decomposition of dense matrices, which resolves the challenge of learning and embedding functions (Dong, 2021).

In real-world applications, the resulting data representations are frequently highly dimensional. Practical algorithms generally behave poorly when faced with many unnecessary features. Finding a way to merge them in a lower-dimensional unified space may, therefore, be beneficial for some tasks, such as those involving pattern recognition and regression problems (Jia, et al., 2022). However, DR techniques, including unsupervised, supervised, and semi-supervised methods, were frequently used in many information processing sectors despite the varying assumptions regarding the distribution of the data or the availability of data labeling (Adadi, 2021). Regression- and spectral graph-based SR avoids the Eigen decomposition of dense matrices and operates more effectively at a faster learning rate. The conditions can also be supervised, unsupervised, or semi-supervised (Thudumu, et al., 2020).

Traditional spectral dimensionality reduction strategies need Eigen decomposition of the dense matrices, which has a high computational cost in terms of time and memory, to find an embedding function that minimizes the objective function. The SR algorithm uses the least squares method to determine the ideal projection direction rather than computing the features' density matrix, enabling it to learn substantially more quickly. A G affinity graph with labeled and unlabeled points was developed to examine the complexity of the underlying geometry and learn responses from given data. The embedding function is realized utilizing these responses and standard regression (Nanga, et al., 2021).

#### *PCA*

PCA is a mathematical method that can be defined as a technique that simultaneously reduces the dimensionality of data while keeping the majority of its variance 1. It accomplishes this reduction by identifying the major components or directions contributing to maximum data fluctuation. Using a limited number of components, each sample can be characterized by a small number of integers

TABLE I  
RESULTS OF TESTING ERROR FOR PRINCIPAL COMPONENT ANALYSIS WITH  $No_{FEATURES}=35$  AND LINEAR DISCRIMINANT ANALYSIS CLASSIFIER FOR SIX INTACT LIMBED

Fold No	Intact-limb 1 (%)	Intact-limb 2 (%)	Intact-limb 3 (%)	Intact-limb 4 (%)	Intact-limb 5 (%)	Intact-limb 6 (%)
Fold 1	13.92	21.24	22.34	14.02	17.83	9.51
Fold 2	18.13	13.92	19.73	19.93	14.12	10.52
Fold 3	12.02	11.42	15.73	12.32	28.95	12.72
Fold 4	12.42	16.63	15.83	15.03	15.23	11.02
Fold 5	12.82	11.82	20.24	12.32	18.43	11.22
Fold 6	15.23	7.71	19.93	14.02	22.74	9.81
Fold 7	22.34	10.32	14.32	14.12	22.44	16.83
Fold 8	18.83	15.83	14.72	14.82	27.95	12.92
Mean	15.71	13.61	17.86	14.57	20.96	11.82

TABLE II  
RESULTS OF AVERAGE TESTING ERROR FOR SPECTRAL REGRESSION AND LINEAR DISCRIMINANT ANALYSIS CLASSIFIER FOR SIX INTACT LIMBED

Fold No	Intact-limb 1 (%)	Intact-limb 2 (%)	Intact-limb 3 (%)	Intact-limb 4 (%)	Intact-limb 5 (%)	Intact-limb 6 (%)
Fold 1	12.22	16.53	17.83	12.42	14.72	8.31
Fold 2	16.73	11.42	19.23	13.72	12.02	8.61
Fold 3	12.12	8.91	10.82	9.21	24.94	10.42
Fold 4	8.31	15.93	11.32	12.02	13.12	10.02
Fold 5	12.82	9.61	14.02	9.11	16.13	9.31
Fold 6	13.92	7.11	15.23	12.52	19.23	8.11
Fold 7	18.63	8.81	14.02	9.41	17.93	15.60
Fold 8	18.73	13.02	12.82	10.92	21.44	9.91
Mean	14.19	11.42	14.41	11.17	17.44	10.04

instead of the thousands of values associated with the tens of thousands of variables (Jolliffe and Cadima, 2016). On the other hand, PCA provides an orthogonal transformation that converts samples with linearly associated features into data with correlated variables. The main components are new features with fewer or equivalent variables to the starting ones. Since the PCA is an unsupervised approach, the data label information is not included. Ordinarily dispersed data have self-contained principal components (Groth, et al., 2013). The PCA is an easy nonparametric method for extracting the most important information from a collection of redundant or noisy data, and this is why it should be used. The benefits of using them extend far beyond image analysis and data compression to include pattern identification and visualization, as well as the prediction and regression of time series (Nanga, et al., 2021). According to Khalid, Khalil, and Nasreen, 2014, PCA has a few drawbacks, which are as follows:

1. It suggests that the relationships between the variables are linear.
2. It can only be interpreted if all variables are scaled numerically.
3. It lacks a probabilistic model framework that has been regarded as crucial in certain contexts, such as Bayesian decision-making and mixed modeling.

### III. EXPERIMENTAL RESULTS AND ANALYSIS

#### A. Experimental setup

Data were initially segmented using an overlapping segmentation approach with a 150 ms window and a 50 ms

TABLE III  
RESULTS OF TESTING ERROR FOR PRINCIPAL COMPONENT ANALYSIS WITH  $No_{FEATURES}=35$  AND LINEAR DISCRIMINANT ANALYSIS CLASSIFIER FOR FOUR AMPUTEES

Fold No	Amputee 1	Amputee 2	Amputee 3	Amputee 4
Fold 1	29.65	22.64	24.45	9.91
Fold 2	20.54	16.23	24.44	15.3
Fold 3	22.24	21.14	26.75	18.23
Fold 4	22.34	17.43	21.04	13.12
Fold 5	18.93	22.34	25.25	18.73
Fold 6	15.83	16.43	28.25	13.72
Fold 7	20.74	20.64	19.13	11.02
Fold 8	28.15	17.13	22.84	12.52
Mean	22.30	19.24	24.01	14.06

TABLE IV  
RESULTS OF AVERAGE TESTING ERROR FOR SPECTRAL REGRESSION AND LINEAR DISCRIMINANT ANALYSIS CLASSIFIER FOR FOUR AMPUTEES

Fold No	Amputee 1	Amputee 2	Amputee 3	Amputee 4
Fold 1	27.35	17.43	20.44	9.11
Fold 2	18.33	14.92	25.15	17.7
Fold 3	20.74	19.63	28.45	15.03
Fold 4	20.84	14.62	23.04	11.02
Fold 5	17.63	18.53	23.54	15.83
Fold 6	15.73	14.22	25.55	12.32
Fold 7	18.63	19.93	17.43	10.42
Fold 8	25.65	19.13	21.94	12.52
Mean	20.61	17.30	23.19	12.99

window increment. In each of the folds, classifiers were trained on eight trials. Classes and error rates for classifiers were calculated using individual trials. This method has been repeated eight times. After completing the eight runs, the average error rate for those eight trials was computed.

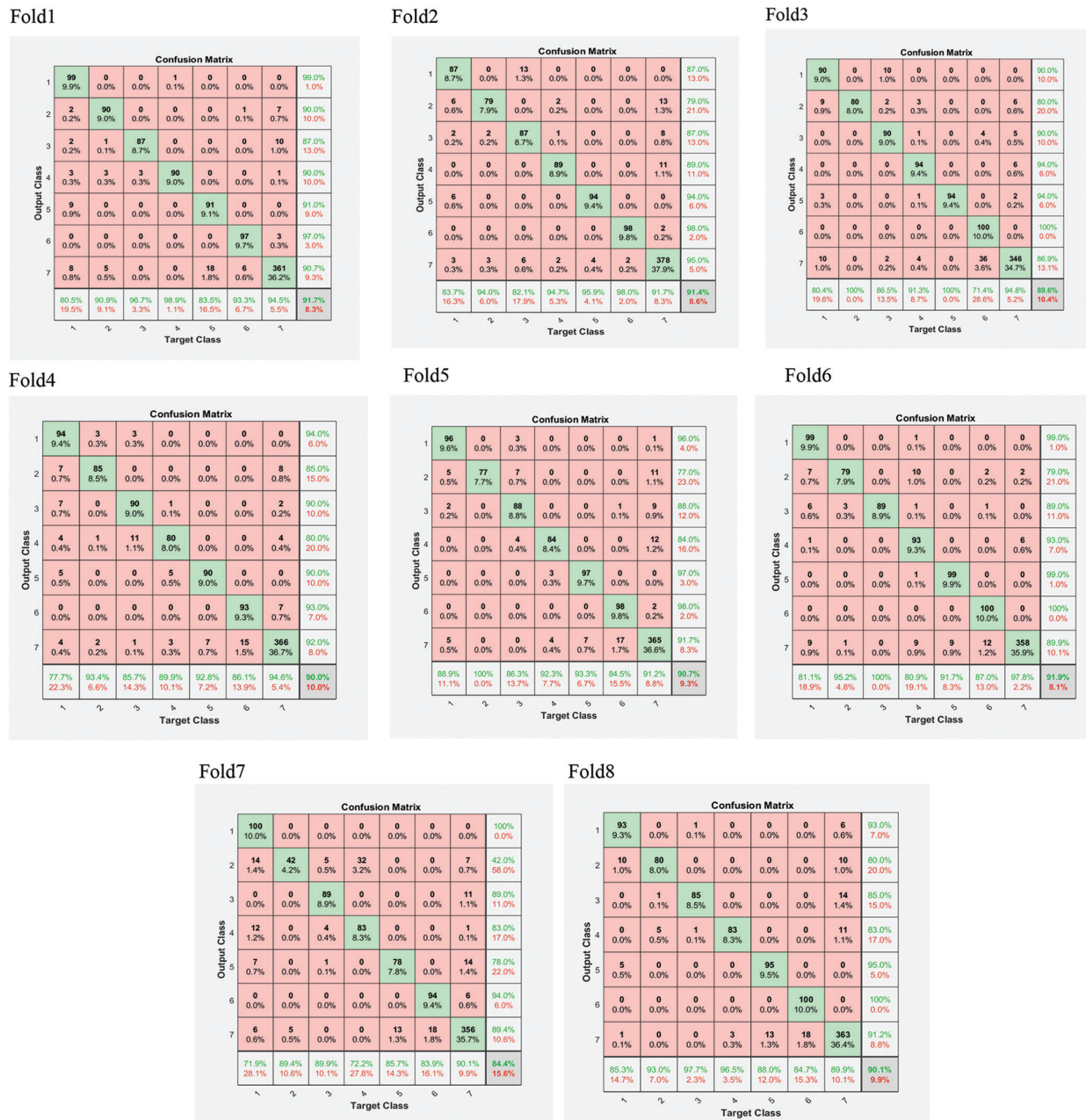


Fig. 6. Confusion matrices of accuracy for each fold of intact-limb 6.

TABLE V

THE OBTAINED AVERAGE TESTING ERROR USING SPECTRAL REGRESSION AND PRINCIPAL COMPONENT ANALYSIS WITH THE LINEAR DISCRIMINANT ANALYSIS CLASSIFIER FOR SIX INTACT LIMBED

Dimensionality reduction	Intact-limb 1 (%)	Intact-limb 2 (%)	Intact-limb 3 (%)	Intact-limb 4 (%)	Intact-limb 5 (%)	Intact-limb 6 (%)	Mean (%)
SR	14.19	11.42	14.41	11.17	17.44	10.04	13.11
PCA	15.71	13.61	17.86	14.57	20.96	11.82	15.75

SR: Spectral regression, PCA: Principal component analysis

TABLE VI

THE OBTAINED AVERAGE TESTING ERROR USING SPECTRAL REGRESSION AND PRINCIPAL COMPONENT ANALYSIS WITH THE LINEAR DISCRIMINANT ANALYSIS CLASSIFIER FOR FOUR AMPUTEES

Dimensionality reduction	Amputee 1	Amputee 2	Amputee 3	Amputee 4	Mean (%)
SR	20.61	17.30	23.19	12.99	18.52
PCA	22.30	19.24	24.01	14.06	19.90

SR: Spectral regression, PCA: Principal component analysis

*B. Testing error based on PCA features reduction*

The first experiment is conducted using the group of TD features, which are RMS, 4<sup>th</sup> Autoregressive, WL, SSC, ZC, MAV, and cardinality. In addition, eight different values for each fold (cross-validation) have been tested. The obtained average testing error is calculated for every intact-limb



TABLE VII  
COMPARATIVE ANALYSIS OF THE BIOSIGNALS TO CLASSIFY SHOULDER GIRDLE MOTIONS TO OTHER STATE-OF-THE-ART

Authors	Method	Channels	Classifier	Classification accuracy (%)
X. Li et al., Li et al., 2017	Time-domain for feature extraction	10 channels EMG and 10 channels EEG	LDA	84.2
Nsugbe and Al-Timemy, 2022	Time-frequency domain	5 channels EMG and 3 channels Acc	LDA	70
The proposed scheme	Time-domain for feature extraction	5 channels EMG and 3 channels Acc	LDA	84.9

EMG: Electromyography, LDA: Linear discriminant analysis, EEG: Electroencephalography

subject value-based PCA (dimensionality reduction) with the LDA classifier, as shown in Table I.

Again, the second experiment is conducted using the group of TD features such as RMS, 4<sup>th</sup> Autoregressive, WL, SSC, ZC, MAV, and cardinality. Every subject value is used to compute the attained average testing error. In addition, the cross-validation method has examined eight distinct values for each fold. As shown in Table II, the LDA classifier and the confusion matrices, as shown in (Fig. 6) for each fold of subject 6, as examples, the obtained average testing error is determined for each subject value based on SR.

The third experiment is done with a set of TD features, such as RMS, 4<sup>th</sup> Autoregressive, WL, SSC, ZC, MAV, and cardinality. Cross-validation has also been used to test eight different values for each fold. For every amputee's subject value-based PCA (dimensionality reduction) with the LDA classifier, the average testing error is calculated, as shown in Table III.

On the other hand, in the fourth trial, we use a variety of TD features, including RMS, 4<sup>th</sup> Autoregressive, WL, SSC, ZC, MAV, and cardinality. There have been eight alternative values for each fold tested using cross-validation. Table IV displays the average testing error for each amputee when using SR with the LDA classifier based on the subject values of the amputees.

The final experiment has been carried out to compare the effect of the used dimensionality reduction methods (SR and PCA) on the accomplished testing error for the resulting models with the LDA classifier. Tables V and VI list obtained results for the same values utilized in the previous two experimentations.

As shown in Table V, the dimensionality reduction method (SR or PCA) can affect the accuracy of the classifier (LDA). The value of classification error decreases when using SR with the LDA classifier, which is 13.11% for six intact limbs. The average testing error for four amputees based on the same dimension reduction (SR with the LDA classifier) is 18.52%, as given in Table VI. This exemplifies the nature of SR dimensionality reduction, in which identifying the optimal feature extraction may work effectively with a large feature dimension.

Comparatively, PCA dimensionality reduction performs less well, with an average test error of 15.75% for six intact limbs. However, the average testing error for four amputees with the same dimension increased to 19.90%. As shown in Tables V and VI, the classification error based on SR decreases as long as the size of feature vectors for all subjects. This is the outcome of the dimensionality reduction method, which searches for patterns among gathered features.

The complexity of the work increases with the number of features.

Classification error results showed a slight difference of 4–5% between amputees and intact-limbed participants (Tables V and VI). This could be related to factors such as getting older, how long it has been since an amputation, or even doing shoulder girdle motions. It was also shown that some amputees did not use the shoulder girdle muscles and did not engage in any sort of exercise to train these muscles. However, it is important to note that SR dimensionality reduction outperformed PCA for both limb-intact and limb-amputated participants.

### C. Comparative analysis of results

Table VII contains a list of the most common methods for classifying different shoulder girdle motions for controlling prosthetics in a comparative examination of EMG and/or acceleration signals collected from the upper limb. The results of the proposed paper were found to have low test errors than previous works.

## IV. CONCLUSION

This study used EMG and accelerometer signals to categorize seven distinct types of shoulder girdle movements performed by high-level upper limb amputees. The results will narrow the focus of dimensionality reduction in the EMG and extract the accelerometer signal feature, which will help determine whether shoulder girdle movements are appropriate as non-invasive and intuitive control signals for upper limb amputees using PR systems. The experimental results showed that SR dimensionality reduction with the LDA classifier was facilitated by extracting regular patterns from biosignals. Experimental results showed that the proposed PR system could identify seven shoulder girdle motions with a classification error of 13.11% for intact-limbed subjects and 18.52% for amputees of SR dimensionality reduction with LDA, and 15.76%–19.90% for PCA dimensionality reduction with the same classifier. Methods from linear algebra that specializes in reducing dimensionality, such as matrix factorization, can be applied in the future studies to decompose a dataset matrix into its component parts. In addition, features can be extracted from biosignals using the TFD to highlight fundamental patterns.

### ACKNOWLEDGMENTS

The authors are grateful to Sharba G.K. for making the dataset available on the internet.



## REFERENCES

- Adadi, A., 2021. A survey on data-efficient algorithms in big data era. *Journal of Big Data*, 8, pp.1-54.
- Al-Timemy, A.H., Khushaba, R.N., Bugmann, G. and Escudero, J., 2016. Improving the performance against force variation of EMG controlled multifunctional upper-limb prostheses for transradial Amputees. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 24(6), pp.650-661.
- Barton, N., 2020. Absence perception and the philosophy of zero. *Synthese*, 197, pp.3823-3850.
- Boashash, B., Khan, N.A. and Ben-Jabeur, T., 2015. Time-frequency features for pattern recognition using high-resolution TFDS: A tutorial review. *Digital Signal Processing*, 40(C), pp.1-30.
- Praciano, F.D.B., Amora, P.R., Abreu, I.C., Pereira, F.L. and Machado, J.C., 2021. Robust cardinality: A novel approach for cardinality prediction in SQL queries. *Journal of the Brazilian Computer Society*, 27, pp.1-24.
- Chen, S., Luo, Z. and Hua, T., 2021. Research on AR-AKF model denoising of the EMG Signal. *Computational and Mathematical Methods in Medicine*, 2021(11), pp.1-10.
- Craelius, W., 2021. *Prosthetic Designs for Restoring Human Limb Function*. Springer, Berlin.
- Choo, Y.J., Kim, D.H. and Chang, M.C., 2022. Amputation stump management: A narrative review. *World J Clin Cases*, 10(13), pp.3981-3988.
- Dong, Y., 2021. A brief review of linear sufficient dimension reduction through optimization. *Journal of Statistical Planning and Inference*, 211, pp.154-161.
- Entezami, A., 2021. Feature extraction in time domain for stationary data. In: *Structural Health Monitoring by Time Series Analysis and Statistical Distance Measures*. Springer, Berlin.
- Groth, D., Hartmann, S., Klie, S. and Selbig, J., 2013. Principal components analysis. *Computational Toxicology*. Springer, Berlin.
- Jang, C.H., Yang, H.S., Yang, H.E., Lee, S.Y., Kwon, J.W., Yun, B.D., Choi, J.Y., Kim, S.N. and Jeong, H.W., 2011. A survey on activities of daily living and occupations of upper extremity amputees. *Annals of Rehabilitation Medicine*, 35(6), pp.907-921.
- Jia, W., Sun, M., Lian, J. and Hou, S., 2022. Feature dimensionality reduction: A review. *Complex and Intelligent Systems*, 8, pp.1-31.
- Jiang, Y., Chen, C., Zhang, X., Chen, C., Zhou, Y., Ni, G., Muh, S. and Lemos, S., 2020. Shoulder muscle activation pattern recognition based on sEMG and machine learning algorithms. *Computer Methods and Programs in Biomedicine*, 197, pp.105721.
- Jolliffe, I.T. and Cadima, J., 2016. Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A Mathematical, Physical and Engineering Sciences*, 374, pp.20150202.
- Karheily, S., Moukadem, A., Courbot, J.B. and Abdeslam, D.O., 2022. sEMG time-frequency features for hand movements classification. *Expert Systems with Applications*, 210, pp.118282.
- Khalid, S., Khalil, T. and Nasreen, S., 2014. A survey of feature selection and feature extraction techniques in machine learning. In: *2014 Science and Information Conference*. IEEE, Piscataway, pp.372-378.
- Li, Q., Liu, Y., Zhu, J., Chen, Z., Liu, L., Yang, S., Zhu, G., Zhu, B., Li, J. and Jin, R., 2021. Upper-limb motion recognition based on hybrid feature selection: Algorithm development and validation. *JMIR mHealth and uHealth*, 9(9), pp.e24402.
- Nanga, S., Bawah, A.T., Acquaye, B.A., Billa, M.I., Baeta, F.D., Odai, N.A., Obeng, S.K. and Nsiah, A.D., 2021. Review of dimension reduction methods. *Journal of Data Analysis and Information Processing*, 9(3), pp.189-231.
- Nsugbe, E. and Al-Timemy, A.H., 2022. Shoulder girdle recognition using electrophysiological and low frequency anatomical contraction signals for prosthesis control. *CAAI Transactions on Intelligence Technology*, 7(1), pp.81-94.
- Padfield, N., 2022. Effective EEG Analysis for Advanced AI-driven Motor Imagery BCI Systems.
- Phinyomark, A., Khushaba, R.N. and Scheme, E., 2018. Feature extraction and selection for myoelectric control based on wearable EMG sensors. *Sensors*, 18(5), pp.1615.
- Pulliam, C.L., Lambrecht, J.M. and Kirsch, R.F. 2011. EMG-based neural network control of transhumeral prostheses. *Journal of rehabilitation research and development*, 48(6), pp.739-754.
- Rampichini, S., Vieira, T.M., Castiglioni, P. and Merati, G., 2020. Complexity analysis of surface electromyography for assessing the myoelectric manifestation of muscle fatigue: A review. *Entropy (Basel)*, 22(5), pp.529.
- Rivela, D., Scannella, A., Pavan, E.E., Frigo, C.A., Belluco, P. and Gini, G. 2015. Processing of surface EMG through pattern recognition techniques aimed at classifying shoulder joint movements. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2015, pp.2107-2110.
- Samuel, O.W., Zhou, H., Li, X., Wang, H., Zhang, H., Sangaiah, A.K. and Li, G., 2018. Pattern recognition of electromyography signals based on novel time domain features for amputees' limb motion classification. *Computers and Electrical Engineering*, 67, pp.646-655.
- Sarker, I.H., 2021. Machine learning: Algorithms, real-world applications and research directions. *SN Computer Science*, 2(3), pp.160.
- Sharba, G.K., Wali, M.K. and Timemy, A.H.A., 2020. Wavelet-based feature extraction technique for classification of different shoulder girdle motions for high-level upper limb amputees. *International Journal of Medical Engineering and Informatics*, 12(6), 609.
- Thudumu, S., Branch, P., Jin, J. and Singh, J.J., 2020. A comprehensive survey of anomaly detection techniques for high dimensional big data. *Journal of Big Data*, 7, pp.1-30.
- Tkach, D., Huang, H. and Kuiken, T.A., 2010. Study of stability of time-domain features for electromyographic pattern recognition. *Journal of Neuroengineering and Rehabilitation*, 7, pp.1-13.
- Toledo-Pérez, D.C., Rodríguez-Reséndiz, J., Gómez-Loenzo, R.A. and Jauregui-Correa, J., 2019. Support vector machine-based EMG signal classification techniques: A review. *Applied Sciences*, 9(20), 4402.
- Zeng, Y., Yang, Z., Cao, W. and Xia, C., 2009. Hand-motion patterns recognition based on mechanomyographic signal analysis. In: *2009 International Conference on Future Biomedical Information Engineering (FBIE)*. Institute of Electrical and Electronics Engineers, Piscataway, pp.21-24.