Check for updates

OPEN ACCESS

EDITED BY Maria Chiara Caschera, National Research Council (CNR), Italy

REVIEWED BY Hong-du Wang, Ocean University of China, China Elizabeth Thomas, Université de Bourgogne, France

*CORRESPONDENCE Muhammad Jawad Khan ⊠ mt.khan@psau.edu.sa

RECEIVED 04 October 2024 ACCEPTED 06 January 2025 PUBLISHED 22 January 2025

CITATION

Nayab M, Waris A, Jawad Khan M, AlQahtani D, Imran A, Gilani SO and Shah UH (2025) Gaussian process latent variable models-ANN based method for automatic features selection and dimensionality reduction for control of EMG-driven systems. *Front. Artif. Intell.* 8:1506042. doi: 10.3389/frai.2025.1506042

COPYRIGHT

© 2025 Nayab, Waris, Jawad Khan, AlQahtani, Imran, Gilani and Shah. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Gaussian process latent variable models-ANN based method for automatic features selection and dimensionality reduction for control of EMG-driven systems

Maham Nayab¹, Asim Waris¹, Muhammad Jawad Khan²*, Dokhyl AlQahtani², Ahmed Imran³, Syed Omer Gilani⁴ and Umer Hameed Shah⁵

¹National University of Science and Technology, Islamabad, Pakistan, ²Department of Electrical Engineering, School of Engineering, Prince Sattam Bin Abdul Aziz University, Al-Kharj, Saudi Arabia, ³Department of Biomedical Engineering and Artificial Intelligence Research Center, College of Engineering and Information Technology, Ajman University, Ajman, United Arab Emirates, ⁴Abu Dhabi University, Abu Dhabi, United Arab Emirates, ⁵Department of Mechanical Engineering and Artificial Intelligence Research Center, College of Engineering and Information Technology, Ajman University, Ajman, United Arab Emirates

Electromyography (EMG) signals have gained significant attention due to their potential applications in prosthetics, rehabilitation, and human-computer interfaces. However, the dimensionality of EMG signal features poses challenges in achieving accurate classification and reducing computational complexity. To overcome such issues, this paper proposes a novel approach that integrates feature reduction techniques with an artificial neural network (ANN) classifier to enhance the accuracy of high-dimensional EMG classification. This approach aims to improve the classification accuracy of EMG signals while substantially reducing computational costs, offering valuable implications for all EMG-related processes on such data. The proposed methodology involves extracting time and frequency domain features from twelve channels of EMG signals, followed by dimensionality reduction using techniques such as PCA, LDA, PPCA, Lasso and GPLVM, and classification using an ANN. Our investigation revealed that LDA is not appropriate for this dataset. The dimensionality reduction models did not have any significant effect on the accuracy, but the computational cost decreased significantly. In individual comparisons, GPLVM had the shortest computational time (29 s), which was significantly less than that of all the other models (p < 0.05), with PCA following at approximately 35 s and Relief at approximately 57 s, while PPCA took approximately 69 s, and Lasso exhibited higher computational costs than all the models but lower computational costs than did the original set. Using the best-performing features, all possible sets of 2, 3, 4 and 5 features were tested, and the 5-feature set exhibited the best performance. This research demonstrates the effectiveness of dimensionality reduction and feature selection in improving the accuracy of movement recognition in myoelectric control.

KEYWORDS

ANN, dimensionality reduction, feature selection, GPLVM, myoelectric control, PCA

1 Introduction

Electromyography (EMG) is a powerful tool that has revolutionized research and healthcare in the field of applied health. It is a diagnostic technique that is used to record and evaluate the electrical activity of skeletal muscle. The importance of EMG signals in research and healthcare cannot be underestimated. They play a vital role in understanding the neuromuscular system and can help diagnose and sometimes even predict various neuromuscular disorders, such as Parkinson's disease (Adem et al., 2022), muscular dystrophy (Quijano-Roy et al., 2004), myasthenia gravis (Baruca et al., 2016), and amyotrophic lateral sclerosis (Sonoo et al., 2009). These signals help clinicians identify the specific muscles affected by the disorder and assess the severity of the conditions (Dai et al., 2022; Nazmi et al., 2016). EMG signals are also used to guide treatments such as physical therapy, electrical stimulation, and surgery. In rehabilitation, EMG signals can be used to evaluate the effectiveness of physical therapy interventions and monitor progress (Dai et al., 2022). EMG signals can provide real-time feedback on muscle activation patterns and help patients learn to activate specific muscles or muscle groups correctly. This feedback can improve motor control and enhance the effectiveness of rehabilitation programs, leading to better outcomes for patients (Nazmi et al., 2016). In myoelectric control, EMG signals are used to control prosthetic devices. These signals are recorded from residual muscles in the amputated limb and transferred to the prosthetic device, allowing the user to control the device's movements. The EMG signals can be processed in real time to translate the user's muscle activity into prosthetic movement, providing a more natural control interface. This can significantly improve the functionality and usability of prosthetic devices, allowing users to perform daily activities more effectively (Chen et al., 2023).

In the field of EMG signal processing, feature extraction is the most crucial step for extracting useful and important information from complex signals. Numerous studies have been conducted to identify new features that can belong to the time domain, frequency domain, or time-frequency domain (Chen et al., 2023; Abbaspour et al., 2020; Phinyomark et al., 2014; Chen et al., 2006; Too et al., 2019b; Danneels et al., 2002; Too et al., 2019c; Naik et al., 2016; Sachin, 2015; Rehman, 2018; Waris et al., 2018; Jochumsen et al., 2018; Zia ur Rehman et al., 2018; Waris et al., 2019; Daffertshofer et al., 2004). However, most of the features are redundant, and using redundant features together could decrease the classification accuracy. Therefore, it is crucial to select appropriate features to achieve optimal performance. Researchers have made significant progress in identifying suitable feature sets and studying their effects on classification accuracy in recent years. The importance of feature reduction lies in its ability to address the issue of high dimensionality, which can lead to overfitting, reduced classification accuracy, and increased computational complexity. Feature reduction techniques can help overcome this problem by reducing the dimensionality of the feature space, thereby improving classification accuracy and reducing computational complexity (Too et al., 2019a; Phinyomark et al., 2012; Phinyomark et al., 2012). Several dimensionality reduction techniques can effectively reduce the number of features while retaining the most informative features. In Zhang et al. (2012), implemented PCA on their EMG data consisting of 512 samples and observed a success rate of 99.8%, indicating the effectiveness of PCA in accurately controlling multiple finger movements using EMG signals. In Yang et al. (2013), employed the wavelet packet transform to extract features from sEMG signals and subsequently applied nonparametric discriminant analysis (NDA) for feature reduction together with a support vector machine (SVM) for classification. Their approach achieved an impressive average accuracy of 98 to 99% per subject. These results demonstrate the potential of using a combination of these methods for accurate EMG pattern recognition. In Junior et al. (2020), compared various techniques of feature reduction for surface electromyography (sEMG) signals obtained from an armband. They employed both feature selection and dimensionality reduction methods, along with seven classifiers, to recognize six different gestures performed by 13 subjects. The researchers found that the highest accuracy of 94% was achieved by combining the support vector machine classifier with dimensionality reduction. In Chu et al. (2007), compared LDA, PCA, NLDA, and self-organizing feature map (SOFM) methods for the classification of features extracted from four-channel EMG signals using WPT. They found that LDA had the best accuracy and real-time performance but required calculating the within-class scatter matrix, which may be singular if there is high feature redundancy. To address this issue, a combination of PCA and SOFM was proposed, which showed improved performance over PCA alone in terms of myoelectric control. This highlights the importance of using appropriate feature reduction methods for improving the accuracy of sEMG signal classification.

Most previous studies have focused on subject-specific EMG classification, whereas this study emphasizes generalized movement classification, which is particularly relevant in robotics and prosthetics. In this study, self-acquired data from an online database is used. The aim of this study is to enhance the performance of classifiers for myoelectric control in generalized devices, including prosthetic limbs, exoskeletons, rehabilitation robots, and assistive robotic systems, while identifying optimal feature sets. Six dimensionality reduction techniques such as Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Probabilistic PCA (PPCA), Gaussian Process Latent Variable Models, Lasso and Relief were applied and their effect on classification accuracies and computational cost is observed. While optimal feature sets are identified using exhaustive feature selection. The most suitable feature sets are identified from 40 well-known electromyography features. The comparison between feature selection methods and dimensionality reduction techniques focused on model accuracy and computational time. Overall, the work emphasizes generalized movement classification rather than subjectspecific approaches. Figure 1 illustrates the study's workflow.

2 Materials and methods

2.1 Data acquisition

This research utilizes two different datasets. Dataset 1 utilizes the online Ninapro database, specifically the second database known as DB2 (Atzori et al., 2014). The dataset consists of EMG signal data from 40 healthy individuals recorded using the Delsys Trigno Wireless EMG system with a sampling rate of 2 kHz. This work is based on the data from 10 of these subjects and examines EMG signals during six distinct movements: abduction and flexion of all fingers (AF, FF), wrist extension (WE), wrist radial deviation (WRD), wrist ulnar deviation (WUD), and wrist extension with a closed hand (WE). Each



movement was repeated six times, with each repetition lasting for 5 s and 3 s of rest between repetitions.

The second dataset was EMG signals collected from ten ablebodied subjects aged between 20 and 35 years (three men/seven women,). The procedures were in accordance with the Declaration of Helsinki and the local ethics committee of the National University of Sciences and Technology, Islamabad, Pakistan (approval number ref. # BMES/REC/22/027). Subjects provided their written, informed consent prior to the experimental procedures. The subjects had no history of upper extremity or other musculoskeletal disorders. Ot Bioelectronica was used for recording EMG signals at a 2 kHz sampling rate. Eight pairs of differential electrodes were positioned below the radio humeral joint at equal intervals; two pairs were positioned on the flexor and extensor digitorium and one pair was placed on the biceps brachii. Six movements were required of the subjects:

- wrist extension (WE)
- wrist radial deviation (WRD)
- wrist ulnar deviation (WUD)
- wrist extension with a closed hand (WE)
- Abduction of all fingrers (AF)
- Fingers Flexed together in a wrist (FF)

Figure 2 illustates the movements performed by subjects. Every movement was performed four times, with a three-second break in between each five-second repetition.

2.2 Pre-processing and feature extraction

Initially, the raw EMG data were preprocessed by applying denoising filters. A notch filter was applied to remove 50 Hz powerline interference, and a fourth-order Butterworth bandpass filter was utilized to permit frequencies within the range of 20–500 Hz (Naik et al., 2016). After filtration, the data were segmented into overlapping segments of 250 ms with 50% overlap. There are three different types of features of EMG signals: the time domain, frequency domain, and time-frequency domain. In this work, thirty-four time domain and six frequency domain features were extracted (Abbaspour et al., 2020; Phinyomark et al., 2014; Chen et al., 2006; Too et al., 2019b; Danneels et al., 2002; Too et al., 2019c; Naik et al., 2016; Sachin, 2015; Rehman, 2018; Waris et al., 2018; Jochumsen et al., 2018; Zia ur Rehman et al., 2018; Waris et al., 2019; Spiewak, 2018). The features are listed in Table 1.

2.3 Feature ranking

After feature extraction, each feature was ranked individually based on its performance. For this, one feature at a time was used as input to the neural network classifier, and its classification accuracy was evaluated. Features that achieved an accuracy above the 80% threshold were considered the best performing. This ranking process was performed using data from a single subject to assess the individual



feature performance. All the other features with accuracies less than the threshold were eliminated from data set.

2.4 Dimensionality reduction techniques

After feature ranking, all features that performed above the threshold were extracted from all subjects across all movements. A generalized approach was applied, where data from all subjects were combined and used to classify the movements, rather than performing classification for each subject individually. The dataset, consisting of 10 subjects, 5 movements, 12 muscles, and 15 features per muscle signal, resulted in a high-dimensional and complex data structure. This led to decreased classification accuracy and increased classifier processing time. Several existing works (Too et al., 2019a; Phinyomark et al., 2012; Phinyomark et al., 2012) have reported the existence of many redundant features in EMG signals, which degrade signal classification accuracy. To increase the classification accuracy, dimensionality reduction techniques can be used. In this work, five dimensionality reduction techniques

were employed, namely, principal component analysis, linear discriminant analysis, probabilistic PCA, least absolute shrinkage and selection, the Relief algorithm and Gaussian process latent variable models (GPLVMs).

PCA is a well-known unsupervised linear dimensionality reduction technique that enables the transformation of highdimensional data into a lower-dimensional space while maintaining crucial information (Daffertshofer et al., 2004). The main goal of PCA is to identify the most important features that have high covariance and use them to transform data into a new vector space with uncorrelated features. The set of new features captures the main trends in key data from the original data, thus making it simpler to visualize and analyze it (Daffertshofer et al., 2004). Linear Discriminant Analysis (LDA) is a widely known supervised statistical technique used for classification and dimensionality reduction. It works by projecting high-dimensional data onto a linearly low-dimensional space that represents discriminative information in the data. This is done by trying to maximize class separability while minimizing within-class scatter. LDA finds applications in many areas such as image processing, speech recognition, bioinformatics etc. due to its

Sr No	Features	Sr No	Features
1	Integrated Emg (iEMG)	21	Maximum Fractal Length (MFL)
2	Root Mean Square (RMS)	22	Log Difference Absolute Mean Value (LDAMV)
3	Variance (VAR)	23	Log of coefficient of variation (LCOV)
4	Waveform Length (WL)	24	Wilson Amplitude (WA)
5	Zero Crossing (ZC)	25	Average Amplitude Change (AAC)
6	Slope Sign Change (SSC)	26	Coefficient Of Variation (CV)
7	Mean Absolute Deviation (MAD)	27	Hjorth Mobility (Hmob)
8	Simple Square Integral (SSI)	28	Absolute Value Of The Summation Of Square Root (ASS)
9	Average Energy (AE)	29	Approximate Entropy (AE)
10	Skewness (Skew)	30	Cardinality (CARD)
11	Modified Mean Absolute Value 1 (MMAV1)	31	Interquartile Range (IQ)
12	Modified Mean Absolute Value 2 (MMAV2)	32	Integrated Absolute Value (IAV)
13	3RD Temporal Moment (TM3)	33	Kurtosis (Kurt)
14	4TH Temporal Moment (TM4)	34	Maximum Energy (ME)
15	5TH Temporal Moment (TM5)	35	Peak Frequency (PF)
16	Standard Deviation (std)	36	Mean Power (MP)
17	V Order (Vo)	37	Mean Frequency (MF)
18	Log Detector (LD)	38	Frequency Ratio (FR)
19	Difference Absolute Mean Value (DAMV)	39	Total Power (TP)
20	Difference Absolute Standard Deviation Value (DAVSR)	40	Modified Mean Frequency (Mmf)

TABLE 1 Features.

simplicity and ease of computation which makes it widespread for various tasks in an analysis of data (Tharwat et al., 2017).

Probabilistic PCA on the other hand is an extension of principal component analysis and is also a linear dimensionality reduction technique. This method uses a probabilistic model to transform data through linear transformations from high dimensions to lower ones unlike the PCA. It assumes that there exists low dimensional subspaces subject to additional Gaussian noises that generate such observations. PPCA estimates this low-dimensional subspace as well as the linear transformation by adding Gaussian noise to a linear transformation of the low-dimensional subspace through expectation maximization algorithm (EM). The noise covariance matrix and model parameters are estimated alternately via the EM technique. The resulting low-dimensional representation of the data preserves its essential features (Tipping and Bishop, 1999).

The GPLVM is a nonlinear technique used for mapping highdimensional data into a low-dimensional latent space. It works similarly to PPCA except that GPLVM is nonlinear. The GPLVM uses a nonlinear function, which is modeled as a Gaussian process (Lalchand et al., 2022). The mapping parameters between the latent and observed variables are treated as random variables with a prior distribution, allowing GPLVMs to capture complex relationships between the variables. The prior distribution is often chosen as a Gaussian distribution, which enables efficient computation via Gaussian process regression. By incorporating a prior distribution on the mapping parameters, GPLVMs can perform Bayesian inference to estimate the posterior distribution of the mapping parameters given the observed data, allowing for probabilistic estimates of the mapping parameters and uncertainty in the model. Kira and Rendell (1992) introduced the Relief algorithm. It functions as an independent assessment method for feature selection. Relief computes a surrogate statistic for each feature, aiding in the estimation of feature 'quality' or 'relevance' to the target concept, such as predicting the endpoint value. These statistics are denoted as feature weights (weights of feature 'A') or, more informally, as feature 'scores', which span a continuum from poor to excellent performance (Kira and Rendell, 1992).

Least absolute shrinkage and selection operator (LASSO) is a popular technique used in machine learning and statistics for selecting a subset of features from a larger set of available features (Muthukrishnan and Rohini, 2016). The functions through adding a penalty toward the standard objective function of linear regression, which penalizes regression coefficients' absolute magnitudes. What this does is to promote sparsity of the solution which means that many of the coefficients are driven to zero and in effect, only a subset of original features are selected. Lasso feature selection is particularly useful in high-dimensional datasets where there are more features than samples, as it can help identify most important predictors while discarding irrelevant or redundant ones. It can also be used for prevention of overfitting and improvement in interpretation by LASSO feature selection (Fonti and Belitser, 2017).

2.5 Exhaustive feature selection

Exhaustive feature selection is a brute-force approach used frequently in machine learning for systematically evaluating all possible feature combinations from a given dataset (Bouzoubaa et al., 2020). Unlike other methods for selecting features that rely on heuristic algorithms or statistical measures, exhaustive feature selection looks exhaustively through all possible feature subsets to identify one that optimizes a predefined criterion like model performance or predictive accuracy. Although exhaustive feature selection guarantees finding the optimal subset of features it becomes computationally expensive as the number of features increases making it impractical for datasets with a large number of features. Despite its computational cost, exhaustive feature selection is valuable in scenarios where model interpretability and the identification of the best-performing feature subset are paramount.

2.6 Classification

A feature set is constructed for movement classification following feature extraction. An ANN was used as a classifier in this evaluation. ANNs are a widespread machine learning method that can figure out intricate linkages and patterns. The composition and operation of the human brain serve as an inspiration for them. An ANN is made up of layers of connected nodes, or neurons. From the input layer, where data are introduced through one or more hidden layers, to the output layer, where predictions or classifications are formed, information travels through the network. Activation functions are used by each neuron to change the input data, and during training, weights-a measure of the strength of connections between neurons-are modified to reduce error between the expected and actual outputs. In this study, a three-layer ANN was used, with each layer containing 20 neurons. The activation function applied was the Rectified Linear Unit (ReLU), and the model was trained for 1,000 iterations. The same ANN was applied after all feature selection and dimensionality reduction methods were performed. All dimensionality reduction and feature selection methods were evaluated using the ANN. Classification accuracies were recorded both before and after dimensionality reduction and then compared to evaluate their performance.

In this study, we employed 5-fold cross-validation to evaluate the model's performance. The dataset was divided into 5 equal folds, and for each iteration, one fold was used for testing while the remaining four folds were used for training. This process was repeated 5 times, ensuring that each fold served as the testing set once. The average performance across all folds was then reported, providing a more robust estimate of the model's generalization capability.

2.7 Statistical analysis

For the dimensionality reduction techniques, the computational cost and accuracies were recorded. The pre-and post-reduction accuracies were compared to observe any significant differences. Similarly, all dimensionality reduction techniques were compared to evaluate which performed better in terms of both accuracy and computational cost. After exhaustive feature selection, the top 10 sets of all the combinations were compared by using statistical analysis. ANOVA was used to analyze and choose the best combination. The best feature sets were choosen based on the classification accuracies. One-way ANOVA was used to analyze the results and determine any statistically significant differences. Since the data followed a nominal distribution, ANOVA was an appropriate choice.

3 Results

To observe the similarities and importance of the extracted features, 3D scatter plots were generated. The scatter plots of DAMV, IAV, IQ, ZC, WL and MAV are shown in Figure 3. The scatter plots show that the majority of the data points are closely cluttered together, which indicates the presence of redundancy in the EMG data (Figure 3).



3.1 Feature ranking and elimination

All forty-one features were evaluated by using ANN. They were ranked based on their classification accuracies. LDAMV had the highest classification accuracy (89%), while cardinality and the WAM were the worst-performing features, with classification accuracies lower than 20%. Figure 4 shows classification accuracy of each individual feature.

Out of 40 features, all the features for which the classification accuracy was less than 85% were eliminated, and only 16 features remained. The remaining features were the modified mean absolute value, interquartile, mean absolute deviation, integral absolute value, integrated EMG, V order, mean absolute value, standard deviation, difference absolute mean value, waveform length, difference absolute standard deviation value, absolute value of the summation of square root, maximum fractal length, and log difference absolute mean value. The trend of feature performance was the same in both datasets.

3.2 Comparison of feature sets with different dimensionality reduction techniques

After conducting feature elimination, a feature vector comprising 16 features from 11 different channels and representing 6 distinct classes extracted from the EMG signals across 10 subjects was generated. Subsequently, various feature selection and reduction techniques were applied to this dataset, and their performance was evaluated based on computational cost and classification accuracy. Given the aim of developing a generalized model, data from all subjects were amalgamated.

Initially, an artificial neural network (ANN) was employed to assess the feature vector post elimination, yielding an accuracy of 88.7%. However, the computational time required for training and evaluating the ANN was approximately 115 s. LDA failed to perform on a large dataset such as this one. Subsequently, PCA was applied, resulting in an accuracy of 88.23% and a substantial reduction in computational time to 35 s (p < 0.05). Similarly, PPCA applied to the original feature vector improved the accuracy to approximately 90%, while reducing the computation time significantly to 69 s (p < 0.05).

When GPLVM was applied to the original feature vector, the accuracy remained relatively constant, yet the computation time decreased significantly to only 29 s (p < 0.05). Relief algorithm application resulted in an accuracy of 89.3% with a computation time reduction to 57 s (p < 0.05). Additionally, the Lasso application yielded an accuracy of 89%, accompanied by a notable reduction in computation time to 90 s (p < 0.05). One-way ANOVA showed no significant difference among the accuracies of all the tested techniques.

The feature selection methodologies had a negligible impact on the classification accuracy but had a considerable influence on the computational expense. Across the spectrum of models, a marked reduction in computational costs was evident compared to the original feature set. Notably, Lasso exhibited significantly diminished costs relative to the original dataset, albeit with higher expenses than other methodologies. PPCA demonstrated a reduction in computational overhead compared to both Lasso and the original set, albeit remaining higher than alternative approaches. Relief exhibited substantially reduced costs in computational expenses. PCA demonstrated diminished computational costs across models, except for GPLVM, which emerged as the frontrunner with significantly reduced expenses compared to all counterparts. Figure 4 illustrates the classification accuracies and computational costs for each dimensionality reduction technique.

3.3 Exhaustive feature selection

Exhaustive feature selection was systematically applied to the datasets, employing sets comprising two, three, four, and five features. Through exhaustive exploration, all possible combinations of these



feature sets were obtained, facilitating a comprehensive evaluation of performance based on classification accuracies. The computational expenditure associated with exhaustive feature selection surpassed that of all other models; however, this was justified by the significantly higher accuracies achieved. Following feature selection, the cost of model training was reduced to a mere second.

Furthermore, the best-performing sets comprising two, three, four, and five features are delineated in the subsequent section.

3.3.1 Two-feature combinations

There are a total of 120 combinations of the 16 features, each of which contains 2 features. All feature sets were evaluated using an ANN. The classification accuracy ranged from 85.39 to 91.87%. The best-performing feature sets were chosen based on the classification accuracies. The top 10 sets of both datasets are depicted in Table 2. Their classification accuracies are plotted in Figures 6, 7.

3.3.2 Three-feature combinations

There are a total of 560 combinations of the 16 features, each of which contains 3 features. All feature sets were evaluated using an ANN. The classification accuracy ranged from 85.58 to 92.40%. The best-performing feature sets were chosen based on the classification accuracies. The top 10 sets from both datasets are shown in Table 3. Their classification accuracies are plotted in Figures 5, 6.

3.3.3 Four-feature combinations

There were a total of 1,820 combinations of the 16 features, each of which contained 4 features. All feature sets were evaluated using an ANN. The classification accuracy ranged from 85.37 to 92.48%. The

best-performing feature set was chosen based on the classification accuracies. The top 10 sets of both datasets are shown in Table 4, and the classification accuracies are plotted in Figures 5, 6.

3.3.4 Five-feature combinations

There were a total of 4,368 combinations of the 16 features, each of which contained 5 features. All feature sets were evaluated using an ANN. The classification accuracy ranged from 83.5 to 92.38%. The best-performing feature sets were chosen based on the classification accuracies. The top 10 sets are shown in Table 5. Their classification accuracies are plotted in Figures 6, 7.

TABLE 2 Optimal two-feature sets.

Sets	Features	
	Dataset 2	Dataset 2
Set 1	IAV, LDAMV	MAD, LDAMV
Set 2	ASS, MFL	IEMG, LDAMV
Set 3	MAV, LDAMV	RMS, LDAMV
Set 4	MAD, LDAMV	DASDV, LDAMV
Set 5	MMAV, LDAMV	VO, LDAMV
Set 6	Vo, LDAMV	IQ, MFL
Set 7	IQ, MFL	LDAMV, MAV
Set 8	MAV, MFL	IQ, LDAMV
Set 9	IEMG, LDAMV	IEMG, MFL
Set 10	ASS, LDAMV	ASS, LDAMV





3.4 Comparison of feature sets

The classification accuracy of feature sets comprising 2, 3, 4, and 5 features was compared to determine the optimal number of features. ANOVA tests were conducted on both datasets to compare the accuracy of different feature sets, which were significantly different (p < 0.002). Both datasets exhibited statistically significant differences among all combinations (p < 0.05). The two-feature set significantly differed from the three-feature set (p < 0.05), and the three-feature set also showed significant differences compared to the four-feature set (p < 0.05), which, in turn, significantly differed from the five-feature set (p < 0.05). The average classification accuracy increased with an increase in the number of features in the set. In conclusion, the findings suggest that using a feature set comprising 5 features yields the highest classification accuracy in both datasets.

4 Discussion

This study highlights the emphasis of dimensionality reduction and feature selection. This study took a comprehensive approach by applying dimensionality reduction methods and exhaustive feature selection to the whole dataset. Unlike other works, which often use subject-specific techniques, our model is intended for use on a wide range of subjects and performs excellently regardless of one's age or sex. This wider scope permits for a more adaptive, more generalized model that can be applied in many different situations and populations. Employing these advanced methodologies across the entire dataset created an all-inclusive robust model that overcame the restrictions set forth in subject-specific analyses.

We have developed an advanced and automated feature extraction method for improving our application. This was because we recognized that different data sets require unique set of features to attain optimum results, hence necessitating the use of a complex technique. It is designed to extract many features from data and subsequently rank them according to their usefulness in a given task. This enabled us to TABLE 3 Optimal three feature sets.

Sets	Features		
	Dataset 1	Dataset 2	
Set 1	VO, DAMV, LDAMV	MAV, DASDV, MFL	
Set 2	MAV, MFL, LDAMV	STD, ASS, LDAMV	
Set 3	IEMG, MFL, LDAMV	RMS, ASS, LDAMV	
Set 4	IEMG, DASDV, LDAMV	MAD, DASDV, MFL	
Set 5	MAD, DAMV, LDAMV	MMAV, DASDV, MFL	
Set 6	MAV, DAMV, LDAMV	MMAV, DAMV, MFL	
Set 7	MAV, STD, MFL	LD, DAMV, MFL	
Set 8	LD, DAMV, LDAMV	RMS, MFL, LDAMV	
Set 9	MMAV, IEMG, LDAMV	WL, ASS, LDAMV	
Set 10	RMS, DASDV, LDAMV	IQ, MFL, LDAMV	

TABLE 4 Optimal 4 feature sets.

Sets	Features	
	Dataset 1	Dataset 2
Set 1	IEMG, Std, WL, LDAMV	MAD,IAV,MFL, LDAMV
Set 2	IAV, WL, DASDV, MFL	MAD,STD, DASDV, LDAMV
Set 3	IEMG, Vo, MFL, LDAMV	WL,DASDV, ASS, MFL
Set 4	RMS, WL, ASS, MFL	IAV,STD,ASS,LDAMV
Set 5	RMS, LDAMV,MAV, DASDV	IAV,STD, DASDV, MFL
Set 6	STD,WL, ASS, MFL	DAMV,WL, ASS, MFL
Set 7	MMAV, Vo, WL, DASDV	IAV,IEMG,WL, LDAMV
Set 8	MAD, DAMV, DASDV, ASS	IAV,RMS, DASDV,LDAMV
Set 9	IAV, Vo, MAV, DAMV	MAD,VO, STD, DAMV
Set 10	RMS, MAV, DAMV, LDAMV	MAV,DAMV,MFL,LDAMV

choose the most appropriate feature reducing time and thus improved over all accuracy and performance of our software intelligently through automation. Such inclusion of this automated approach for selecting features not only made it possible for us to address diverse datasets but also improved the robustness and reliability of our work; hence making our study a significant contribution toward development of feature extraction techniques (Figure 7).

Our investigation suggested that employing dimensionality reduction techniques such as PCA, GPLVM, lasso, relief, and PPCA can notably enhance both accuracy and reduce processing time. However, LDA may not be suitable for this type of analysis. Although these reduction techniques did not notably increase the classification accuracy, they significantly reduced the computational costs. They performed well, particularly improving classification accuracy when applied to larger datasets. Comparing each model individually, we found significantly lower computational times. The GPLVM had the shortest computational cost at 29 s, which was significantly less than that of all the other models (p < 0.05). PCA took approximately 35 s, which was significantly greater than that of GPLVM but notably less than that of the other models. Relief took approximately 57 s, which was significantly greater than that of GPLVM and PCA but less than that of the other models. PPCA took approximately 69 s, which was significantly greater than that of Relief, GPLVM, and PCA but less than that of the other models. Although the computational cost of Lasso was greater than that of all other models, it was lower than that of the original set.

All these models are viable for dimensionality reduction, with GPLVM emerging as the best model according to our study. However, for the highest accuracy, exhaustive feature selection remains the best model despite its high cost. The observed outcome might be attributed to the utilization of unsupervised dimensionality reduction techniques, which inherently aim to map high-dimensional data onto lower-dimensional spaces. Consequently, the selected features, while deemed optimal, may not possess substantial variance or discriminatory power in the reduced space. This could explain the observed limitation in capturing diverse patterns or distinguishing characteristics within the data (Figure 8).



dataset 2.

The number of features used and the types of considered features play a crucial role in the classification of hand gestures. This is why the features were initially ranked, and those performing poorly were subsequently eliminated from the set. Therefore, the number of features used depends on different factors, such as the type of activity, muscle, sampling frequency, and type of feature in a combination. Figure 9 illustrates the effect of the number of features on the classification accuracy. The classification accuracies generally show an upward trend from 2 to 10 features, followed by a slight decrease when there are more than 9 features in a set. This trend could be attributed to feature redundancy. In this work, we used exhaustive feature selection to obtain 10 high-performing sets of 2, 3, 4 and 5 combinations. Even though the protocols for both dataset acquisition and processing were the same, they still had different highperforming sets, as shown in Tables 2–4.

TABLE 5	Optimal	five	feature	sets.

Sets	Features	
	Dataset 1	Dataset 2
Set 1	MMAV, WL, DASDV,ASS, LDAMV	MMAV, RMS, MAV,ASS, LDAMV
Set 2	MFL, RMS, DAMV,LDAMV, MMAV	IQ, IEMG, MAV, MFL, LDAMV
Set 3	MMAV, MAD, IEMG, DASDV, LDAMV	IQ, MAD, RMS, MFL, LDAMV
Set 4	IQ, IAV,RMS, MFL, LDAMV	MMAV,LD,RMS, MFL, DAMV
Set 5	IQ,IEMG, RMS,DASDV, LDAMV	MMAV, MAV, WL, ASS, MFL
Set 6	IQ, Vo, MAV, DAMV, LDAMV	IQ, LD, DASDV, ASS, LDAMV
Set 7	IQ, WL, ASS, MFL, LDAMV	IQ, MAD, WL, ASS, MFL
Set 8	IQ, MAD, DASDV, MFL, LDAMV	IQ, LD, IEMG, DAMV, LDAMV
Set 9	IQ, Vo, DASDV, ASS, LDAMV	MMAV,IQ, RMS, ASS, LDAMV
Set 10	MMAV, MAD, WL, DASDV, LDAMV	MMAV, LD, STD, MFL, LDAMV





This study utilized data from 12 muscles to analyze five different movements. Among these, some muscles were identified as dominant for specific movements, while others were not. Even if a muscle is not dominant for a particular movement, it often plays a supporting or stabilizing role. The interplay between muscles is complex, as most movements require some degree of co-activation to ensure smooth and coordinated motion. To further reduce computational time, future studies could investigate whether classification accuracies remain consistent when the number of muscles analyzed is reduced or when only a subset of muscles is selected for the analysis.

5 Conclusion

This work provides evidence that by using dimensionality reduction techniques, higher-dimensional features extracted from multiple EMG channels can improve the accuracy of classification in EMG signal analysis. The use of such feature selection also significantly improved the classification accuracy, with the 5-feature set exhibiting the best performance. These findings have implications for developing more accurate and efficient algorithms for myoelectric control applications.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://doi.org/10.6084/m9.figshare.23898012.v1.

Ethics statement

The studies involving humans were approved by local ethics committee of the National University of Sciences and Technology, Islamabad, Pakistan (approval number ref. # BMES/REC/22/027). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

MN: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. AW: Funding acquisition, Project administration, Resources, Supervision, Writing – review & editing. MJ: Validation, Writing – review & editing. DA: Funding acquisition, Writing – review & editing. AI: Writing – review & editing. SG: Writing – review & editing. US: Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was funded by Ajman University Internal Research Grant (No. 2023-IRG-ENIT-32).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that Gen AI was used in the creation of this manuscript. Generative AI was only used to edit and refine the manuscript and for removing any grammatical errors

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abbaspour, S., Lindén, M., Gholamhosseini, H., Naber, A., and Ortiz-Catalan, M. (2020). Evaluation of surface EMG-based recognition algorithms for decoding hand movements. *Med. Biol. Eng. Comput.* 58, 83–100. doi: 10.1007/s11517-019-02073-z

Adem, H. M., Tessema, A. W., and Simegn, G. L. (2022). Classification of Parkinson's disease using EMG signals from different upper limb movements based on multiclass support Vector Machine. *Int. J. Bioautom.* 26, 109–125. doi: 10.7546/ ijba.2022.26.1.000849

Atzori, M., Gijsberts, A., Castellini, C., Caputo, B., Hager, A.-G. M., Elsig, S., et al. (2014). Electromyography data for noninvasive naturally controlled robotic hand prostheses. *Sci. Data* 1, 1–13. doi: 10.1038/sdata.2014.53

Baruca, M., Leonardis, L., Podnar, S., Hojs-Fabjan, T., Grad, A., Jerin, A., et al. (2016). "Single fiber EMG as a prognostic tool in myasthenia gravis. *Muscle Nerve* 54, 1034–1040. doi: 10.1002/mus.25174

Bouzoubaa, Kawtar, Taher, Youssef, and Nsiri, Benayad. "Dos attack forecasting: a comparative study on wrapper feature selection." 2020 international conference on intelligent systems and computer vision (ISCV), (2020).

Chen, Z., Min, H., Wang, D., Xia, Z., Sun, F., and Fang, B. (2023). A review of myoelectric control for prosthetic hand manipulation. *Biomimetics* 8:328. doi: 10.3390/biomimetics8030328

Chen, W.-t., Wang, Z.-z., and Ren, X.-m. (2006). Characterization of surface EMG signals using improved approximate entropy. *J Zhejiang Univ Sci B* 7, 844–848. doi: 10.1631/jzus.2006.B0844

Chu, J.-U., Moon, I., Lee, Y.-J., Kim, S.-K., and Mun, M.-S. (2007). A supervised feature-projection-based real-time EMG pattern recognition for multifunction myoelectric hand control. *IEEE/ASME Trans. Mechatron.* 12, 282–290. doi: 10.1109/TMECH.2007.897262

Daffertshofer, A., Lamoth, C. J. C., Meijer, O. G., and Beek, P. J. (2004). PCA in studying coordination and variability: a tutorial. *Clin. Biomech.* 19, 415–428. doi: 10.1016/j.clinbiomech.2004.01.005

Dai, Y., Wu, J., Fan, Y., Wang, J., Niu, J., Gu, F., et al. (2022). MSEva: a musculoskeletal rehabilitation evaluation system based on EMG signals. *ACM Trans. Sens. Networks* 19:1. doi: 10.1145/3522739

Danneels, L., Coorevits, P. L., Cools, A. M., Vanderstraeten, G. G., Cambier, D. C., Witvrouw, E. E., et al. (2002). Differences in electromyographic activity in the multifidus muscle and the iliocostalis lumborum between healthy subjects and patients with sub-acute and chronic low back pain. *Eur. Spine J.* 11, 13–19. doi: 10.1007/s005860100314

Fonti, V., and Belitser, E. (2017). Feature selection using lasso. VU Amsterdam Res. Paper Bus. Anal. 30, 1–25.

Jochumsen, M., Waris, A., and Kamavuako, E. N. (2018). The effect of arm position on classification of hand gestures with intramuscular EMG. *Biomed. Signal Proces. Control* 43, 1–8. doi: 10.1016/j.bspc.2018.02.013

Junior, J. J. A. M., Freitas, M. L., Siqueira, H. V., Lazzaretti, A. E., Pichorim, S. F., Stevan Jr, S. L. (2020). Feature selection and dimensionality reduction: An extensive comparison in hand gesture classification by sEMG in eight channels armband approach. *Biomedical Signal Processing and Control*, 59:101920.

Kira, Kenji, and Rendell, Larry A.. "A practical approach to feature selection." Machine learning proceedings 1992. Morgan Kaufmann, (1992). 249–256

Lalchand, V., Ravuri, A., and Lawrence, N. D. (2022). Generalized Gaussian process latent variable models (GPLVM) with stochastic variational inference. *arXiv*.

Muthukrishnan, Ramakrishnan, and Rohini, R., "LASSO: a feature selection technique in predictive modeling for machine learning." 2016 IEEE international conference on advances in computer applications (ICACA), (2016). Naik, G. R., Selvan, S. E., Gobbo, M., Acharyya, A., and Nguyen, H. T. (2016). Principal component analysis applied to surface electromyography: a comprehensive review. *IEEE Access* 4, 4025–4037. doi: 10.1109/ACCESS.2016.2593013

Nazmi, N., Abdul Rahman, M., Yamamoto, S.-I., Ahmad, S., Zamzuri, H., and Mazlan, S. (2016). A review of classification techniques of EMG signals during isotonic and isometric contractions. *Sensors* 16:1304. doi: 10.3390/s16081304

Phinyomark, A., Phukpattaranont, P., and Limsakul, C. (2012). Feature reduction and selection for EMG signal classification. *Expert Syst. Appl.* 39, 7420–7431. doi: 10.1016/j. eswa.2012.01.102

Phinyomark, A., Quaine, F., Charbonnier, S., Serviere, C., Tarpin-Bernard, F., and Laurillau, Y. (2014). Feature extraction of the first difference of EMG time series for EMG pattern recognition. *Comput. Methods Prog. Biomed.* 117, 247–256. doi: 10.1016/j. cmpb.2014.06.013

Quijano-Roy, S., Renault, F., Romero, N., Guicheney, P., Fardeau, M., and Estournet, B. (2004). EMG and nerve conduction studies in children with congenital muscular dystrophy. *Muscle Nerve* 29, 292–299. doi: 10.1002/mus.10544

Rehman, Zia Ur, Muhammad, et al."multiday EMG-based classification of hand motions with deep learning techniques." Sensors. (2018). 18:2497, doi: 10.3390/s18082497

Sachin, D. (2015). Dimensionality reduction and classification through PCA and LDA. *Int. J. Comp. App.* 122, 4–8. doi: 10.5120/21790-5104

Sonoo, M., Kuwabara, S., Shimizu, T., Komori, T., Hirashima, F., Inaba, A., et al. (2009). Utility of trapezius EMG for diagnosis of amyotrophic lateral sclerosis. *Muscle Nerve* 39, 63–70. doi: 10.1002/mus.21196

Spiewak, C. "A comprehensive study on EMG feature extraction and classifiers." Open Access Journal of Biomedical Engineering and Biosciences 1.1 (2018).

Tharwat, A., Gaber, T., Ibrahim, A., and Hassanien, A. E. (2017). Linear discriminant analysis: a detailed tutorial. *AI Commun.* 30, 169–190. doi: 10.3233/AIC-170729

Tipping, M. E., and Bishop, C. M. (1999). Probabilistic principal component analysis. J. Royal Statis. Soc. Series B: Statis. Methodol. 61, 611–622. doi: 10.1111/1467-9868.00196

Too, J., Abdullah, A. R., Mohd Saad, N., and Tee, W. (2019a). EMG feature selection and classification using a Pbest-guide binary particle swarm optimization. *Computation* 7:12. doi: 10.3390/computation7010012

Too, J., Abdullah, A. R., and Saad, N. M. (2019b). Classification of hand movements based on discrete wavelet transform and enhanced feature extraction. *Int. J. Adv. Comput. Sci. Appl.* 10:6. doi: 10.14569/IJACSA.2019.0100612

Too, J., Abdullah, A. R., and Saad, N. M. (2019c). Study of EMG feature selection for hand motions classification. *Int. J. Human Technol. Interaction (IJHaTI)* 3, 19–24.

Waris, A., Mendez, I., Englehart, K., Jensen, W., and Kamavuako, E. N. (2019). On the robustness of real-time myoelectric control investigations: a multiday Fitts' law approach. *J. Neural Eng.* 16:026003. doi: 10.1088/1741-2552/aae9d4

Waris, A., Niazi, I. K., Jamil, M., Englehart, K., Jensen, W., and Kamavuako, E. N. (2018). Multiday evaluation of techniques for EMG-based classification of hand motions. *IEEE J. Biomed. Health Inform.* 23, 1526–1534. doi: 10.1109/JBHI.2018.2864335

Yang, P., Xing, K., Huang, J., and Wang, Y., "A novel feature reduction method for real-time EMG pattern recognition system," 2013 25th Chinese Control and Decision Conference (CCDC), Guiyang, China, (2013), pp. 1500–1505.

Zhang, Daohui, et al. "PCA and LDA for EMG-based control of bionic mechanical hand." 2012 IEEE international conference on information and automation, (2012).

Zia ur Rehman, M., Gilani, S. O., Waris, A., Niazi, I. K., Slabaugh, G., Farina, D., et al. (2018). Stacked sparse autoencoders for EMG-based classification of hand motions: a comparative multi day analyses between surface and intramuscular EMG. *Appl. Sci.* 8:1126. doi: 10.3390/app8071126