Swarmed Discriminant Analysis for Multifunction Prosthesis Control

Rami N. Khushaba *, Ahmed Al-Ani, and Adel Al-Jumaily,

Abstract—One of the approaches enabling people with amputated limbs to establish some sort of interface with the real world includes the utilization of the myoelectric signal (MES) from the remaining muscles of those limbs. The MES can be used as a control input to a multifunction prosthetic device. In this control scheme, known as the myoelectric control, a pattern recognition approach is usually utilized to discriminate between the MES signals that belong to different classes of the forearm movements. Since the MES is recorded using multiple channels, the feature vector size can become very large. In order to reduce the computational cost and enhance the generalization capability of the classifier, a dimensionality reduction method is needed to identify an informative yet moderate size feature set. This paper proposes a new fuzzy version of the well known Fisher’s Linear Discriminant Analysis (LDA) feature projection technique. Furthermore, based on the fact that certain muscles might contribute more to the discrimination process, a novel feature weighting scheme is also presented by employing Particle Swarm Optimization (PSO) for estimating the weight of each feature. The new method, called PSOFLDA, is tested on real MES datasets and compared with other techniques to prove its superiority.

Keywords—Discriminant Analysis, Pattern Recognition, Signal Processing.

I. INTRODUCTION

The myoelectric signal (MES), also known as the Electromyogram (EMG), is a non-stationary signal that carries the distinct signature of the voluntary intent of the central nervous system. It is usually recorded in a noninvasive scheme utilizing a set of surface electrodes mounted on the human forearm. One of the most important applications of the MES is its use in controlling prosthetic devices functioning as artificial alternatives to missing limbs [1]. Advances in myoelectric signal studies revealed that the MES exhibits different temporal structure for different kinds of the arm movements. This in turn facilitates the use of pattern recognition in myoelectric control for prosthetic control. To this end, a wide set of pattern recognition methods were proposed in the literature to accurately classify the MES into one of a predefined set of movements [2].

In order to capture the complete muscles activity and maximize the amount of available information, a multi channel approach is usually utilized when measuring the MES signal. However, this will increase the number of extracted features (variables that describe these movements) and hence it will increase the learning parameters of the classifier and may degrade its performance. A straight forward solution to these problems is to project the data onto low-dimensional subspaces to extract the most significant features. Many feature projections techniques were used in myoelectric control with the aim to produce a statistically uncorrelated or independent feature set, a desirable goal in any pattern recognition system. Various approaches of dimensionality reduction were utilized in myoelectric control, these include: principal component analysis (PCA) [3], a combination of PCA and self organizing feature map (SOFM) [4] and linear discriminant analysis (LDA) [5].

Despite being a well known projection technique, the classical LDA suffers from a number of limitations [6]. The first is that it requires the scatter matrices to be nonsingular, while in real world problems they can be singular. The second limitation with LDA is that it treats all data points equivalently, whereas in real world problems each sample may belong to each of the different classes with a certain degree. Finally, classical LDA pays no attention to the decorrelation of the data, which is a desirable property in many applications. One possible approach to overcome the first and the third problems is to use the uncorrelated linear discriminant analysis (ULDA) that requires the reduced features to be statistically uncorrelated with one another [7].

As a variation to the ULDA approach which is based on Singular Value Decomposition (SVD) that is known to be expensive in terms of time and memory requirements for large datasets, this paper proposes a new mixture of fuzzy logic and discriminant analysis as a novel dimensionality reduction technique. The proposed method aims to reduce the dimensionality of the extracted feature set and cluster features, such that the classification accuracy is improved. Due to the fact that most of the biosignals generated by the human body tend to produce patterns that are fuzzy in nature (i.e., belongs to different classes with certain degrees), then the incorporation of the concept of fuzzy memberships is required to reduce the effect of overlapping and outliers points. Unlike the current available variations to Fisher’s linear discriminant analysis (LDA), the new method, called PSOFLDA, accounts for the different contribution of different muscles into the discrimination process. Thus, it assumes that the extracted features vary in their importance. In order to implement this muscle importance concept, a novel feature weighting scheme that employs Particle Swarm Optimization (PSO) for the estimation of features’ weights is introduced. Also in order to overcome the singularity problem, a regularization parameter is included within each particle (where each particle represents one member of the population).

This paper is structured as follows: Section 2 explains the
proposed methodology. Section 3 presents the swarm-based weight optimization. The experimental results are given in section 4. Finally the conclusion is given in section 5.

II. METHODOLOGY

A variation of fisher’s classical LDA is the fuzzy discriminant analysis (FLDA) [8] that emerged as a classification tool, proving to present better performance than LDA. The goal of FLDA is to determine the linear discriminant function that provides the maximum separation of fuzzy groups in a real space. Although the initial work on FLDA dates back to 1986 [8], there were only few attempts in the literature to propose variations to the original FLDA classifier. Inspired by the work of [8], an interesting approach in fuzzy linear discriminant analysis was proposed for use with chemical datasets [9]. In their approach a K-nearest neighbor (KNN) rule was utilized in order to estimate the required memberships. The approach was compared with classical LDA and proved to outperform classical LDA on chemical datasets. The authors mentioned explicitly that several runs should be made in order to decide the best value of K and the other membership parameter utilized in their approach. A different approach in estimating the FLDA memberships using the KNN rule was proposed in [10]. The main difference between the two approaches presented in [9] and [10] is that a preprocessing step employing PCA is utilized by the latter. This in turn forms a fuzzy variation to classical subspace LDA [11]. It is generally known from the literature that the limitation of this approach is that the application of PCA may leads to loosing useful information. In another attempt and inspired by fuzzy support vector machine principles, another technique for estimating the memberships in FLDA was presented first, then a kernel technique was introduced to perform the nonlinear mapping [12]. The presented comparison between the kernel based FLDA (KFDA) and FLDA indicated that better results were achieved by the former, but with the additional computational requirements of the kernel matrix.

In spite of the good performance achieved by FLDA in a number of applications, there are two main issues that both LDA and FLDA suffer from. Firstly, both techniques pay no attention to the decorrelation of the data. Hence they may not always give the optimal results especially when the feature set contains a large degree of redundancy. Secondly, both LDA and FLDA require the total scatter matrix to be nonsingular, while in real problems the scatter matrix can be singular. In order to overcome these problems with LDA and FLDA, a number of approaches were proposed, such as the subspace LDA (mentioned earlier) and regularized LDA [13]. In the regularized LDA a constant $\xi$ is added to the diagonal elements of $SW$, $SW = SW + \xi I$, for some $\xi > 0$, where $I$ is an identity matrix. It is easy to see that the new $SW$ is nonsingular, but the main problem is finding the optimal value of $\xi$.

This paper proposes a new weighted fuzzy linear discriminant analysis technique (term as PSOFLDA) for feature projection. A novel approach is presented within PSOFLDA to overcome the earlier mentioned problems by assigning a weighting factor $w_i$ for each feature. In such approach, the features that are most relevant to the problem will be weighted with higher importance than other features. Thus rather than applying PCA to minimize redundancies, we employ a weighting process that aims to maximize the interaction between features. In order to identify these weights, particle swarm optimization is utilized, as explained in the next section.

A. Weighted Fuzzy Discriminant Analysis

Consider a classification problem with $c$ classes, in which the data set of labelled training samples is given as:

$$S = \{(x_1, y_1), (x_2, y_2), \ldots, (x_l, y_l)\} \subseteq (X, Y)^l$$  \hspace{1cm} (1)

Where $X$ is the input space and $Y$ is the output space. $X \subseteq \mathbb{R}^n$, $Y \subseteq \mathbb{R}^c$, and $l$ is the number of samples. Each training point $x_i$ originally belongs to one of the $c$ classes and is given a label $y_i \in \{1, 2, 3, \ldots, c\}$ for $i = \{1, 2, 3, \ldots, l\}$. The goal is to find an optimal hyper-plane using the training samples that can recognize the test points, i.e., the classifier will have a good generalization capability. In PSOFLDA each point, $x_i$, belongs to each of the $c$ classes with a certain degree of membership. The fuzzy within class scatter matrix $SW$, fuzzy between class scatter matrix $SB$, and the fuzzy total class scatter matrix $ST$ are given as follows:

$$SW = \sum_{i=1}^{c} \sum_{k=1}^{l} u_{ik}^m (x_k - v_i)(x_k - v_i)^T \otimes (ww^T)$$  \hspace{1cm} (2)

$$SB = \sum_{i=1}^{c} \sum_{k=1}^{l} u_{ik}^m (v_i - \bar{x})(v_i - \bar{x})^T \otimes (ww^T)$$  \hspace{1cm} (3)

$$ST = \sum_{i=1}^{c} \sum_{k=1}^{l} u_{ik}^m (x_k - \bar{x})(x_k - \bar{x})^T \otimes (ww^T)$$  \hspace{1cm} (4)

Where $u_{ik}$ is the membership of pattern $k$ in class $i$, $m$ (given that $m > 1$) is the fuzzification parameter, $x_k$ is the value of the $k$’th sample across the $j$’th dimension, $\bar{x}_i$ is the mean of the patterns belonging to class $i$, and $v_i$ is its value across the $j$’th dimension. $\otimes$ refers to the Kronecker product operation, $w$ is the weight vector associated with all features, i.e., $w = \{w_1, w_2, \ldots, w_f\}$, where $f$ is the total number of features. $\bar{x}$ is the mean of the training samples which is given in Eq.(5) below.

$$\bar{x} = \frac{1}{f} \sum_{k=1}^{l} x_k$$  \hspace{1cm} (5)

In this paper, the value of the membership $u_{ik}$ is calculated using a possibilistic fuzzy clustering approach. The cost function of the possibilistic clustering approach is adopted from [14], as given in Eq.(6) below.

$$J(\theta, U) = \sum_{i=1}^{l} \sum_{k=1}^{c} u_{ik}^m (x_k - \theta_i)^2 + \sum_{i=1}^{c} \sum_{k=1}^{l} \eta_i (1 - u_{ik})^m$$  \hspace{1cm} (6)

where $\theta_i$ is the $i$’th cluster center, $\eta_i$ are positive constants that are suitably chosen. The first term in Eq. (6) is the same objective function used in probabilistic clustering approach, while the second term is added to reduce the effect of outliers. In order to find the membership values from the above
equation, then the values of the clusters centers are needed. A
direct way would be to differentiate Eq. (6) with respect to \( \theta \),
but this in turn would cancel the second term leaving only the
first term. A general look at the first term of Eq. (6) reveals
that it represents the classical within class scatter matrix \( S_W \)
given in Eq. (2) if the weight is removed. Thus applying the
values of the clusters means ensures that the objective function
given by Eq. (6) would settle at a global optimum value. Then
in order to compute the membership values, a differentiation
of the resultant function with respect to \( u_{ik} \) needs to be done as
follows.

\[
\frac{\partial J(\theta, U)}{\partial u_{ik}} = mu_{ik}^{m-1}(x_k - v_i)^2 - m\eta_i(1 - u_{ik})^{m-1} = 0
\]  

(7)

This would in turn result in the following function

\[
u_{ik} = \frac{1}{1 + \frac{(x_k - v_i)^2}{\eta_i}}
\]

(8)

The values of \( \eta_i \), where \( i = \{1, 2, 3, \ldots, c\} \) were chosen to be
equal to the maximum distance between the samples belonging
to that class and the class center.

After computing all the variables, PSOFLDA finds the
vector \( G \) that would maximize the ratio of the between class scatter matrix
to the within class scatter matrix by solving the following equation:

\[
G = \arg \max_G \text{trace}((G^T S_W G)^{-1}G^T S_B G)
\]  

(9)

The solution can be readily computed by applying an eigen-
decomposition on \( S_W^{-1} S_B \), provided that the within class scatter matrix \( S_W \) is nonsingular. In this paper, we are using a
regularized version of \( S_W \) given by \( S_W = S_W + zI \), for some
\( z > 0 \) that is included in the particle representation of the
weights, where \( I \) is an identity matrix. In this way the scatter
matrix is guaranteed to be nonsingular. Since the rank of the
between class scatter matrix is bounded from above by \( c - 1 \),
there are at most \( c - 1 \) discriminant vectors by PSOFLDA.

III. SWARM BASED WEIGHT OPTIMIZATION

One possible solution for finding the best values of the
weights is to employ evolutionary algorithms, or EAs. Powerful
EA algorithms include genetic algorithm (GA) and Particle
Swarm Optimization (PSO). PSO is an effective continuous
function optimizer that encodes the parameters as floating-
point numbers and manipulate them with arithmetic operators.
By contrast, GAs are often better suited for combinatorial
optimization because they encode the parameters as bit strings
and modify them with logical operators. There are many
variants to both approaches, but because PSO is primarily a
numerical optimizer, thus PSO is considered in this paper.

A. Particle Swarm Optimization

Particle swarm optimization, is a population based stochas-
tic optimization technique developed by Eberhart and Kennedy
in 1995 [15]. It represents an example of a modern search
heuristics belonging to the category of Swarm Intelligence
methods. PSO mimics the behavior of a swarm of birds or a
school of fish. The swarm behavior is modelled by particles in
multidimensional space that have two characteristics: position
(\( p \)) and velocity (\( s \)). These particles wander around the hyper
space and remember the best position that they have discovered.
A particle’s position in the multi-dimensional problem
space represents one solution for the problem. They exchange
information about good positions to each other and adjust
their own position and velocity with certain probabilities based
on these good positions. The original formula developed by
Kennedy and Eberhart was improved by Shi and Eberhart with
the introduction of an inertia weight \( \varpi \) that decreases over
time, (typically from 0.9 to 0.4), to narrow the search that
would induce a shift from an exploratory to an exploitative
mode. Though the maximum velocity of a particle \( (s_{max}) \)
was no longer necessary for controlling the explosion of the
particles, Shi and Eberhart continued to use it, often setting
\( s_{max} = p_{max} \) that is the maximum velocity is equalled to the
maximum value along the specific dimension, in order to
keep the system within the relevant part of the search space.
This was found to be a good idea that significantly improves
the PSO performance and at the same time it costs very
little computationally. During iterations each particle adjusts
its own trajectory in the space in order to move towards its
best position and the global best according to the following equations:

\[
s_{ij}(t + 1) = \varpi s_{ij}(t) + c_1 r_1 (p_{besti} - p_{ij}) + c_2 r_2 (g_{bestj} - p_{ij})
\]

(10)

\[
p_{ij}(t + 1) = p_{ij}(t) + s_{ij}(t + 1)
\]

(11)

Where

\( i \): is the particle index  
\( j \): is the current dimension under consideration  
\( p_i \): is the current position,  
\( s_i \): is the current velocity  
\( \varnothing \): is the inertia weight  
\( t \): is the current time step  
\( r_1 \) and \( r_2 \) are two random numbers uniformly distributed in
the range \((0,1), c_1 \) and \( c_2 \) are cognitive and social parameters
respectively. \( p_{besti} \) is the local best position, the one
associated with the best fitness value the particle has achieved
so far, and \( g_{bestj} \) is the global best position, the one
associated with the best fitness value found among all of the
particles.

The personal best of each particle is updated according to
the following equation:

\[
p_{besti}(t + 1) = \begin{cases} p_{besti}(t) & \text{if } (p_{besti}(t)) \leq f(p_i) \\ p_i(t) & \text{if } (p_{besti}(t)) > f(p_i) \end{cases}
\]

(12)

Finally, the global best of the swarm is updated using the following equation:

\[
g_{best}(t + 1) = \arg \min_{p_{besti}} f(p_{besti}(t + 1))
\]

(13)

Where \( f(.) \) is a function that evaluates the fitness value for a
given position. This model is referred to as the \( g_{best} \) (global
best) model. In this paper, each particle will represent a vector whose elements are the weights assigned to each feature plus the regularization parameter $z$. The idea here is to generate a new weight vector by utilizing a set of particles that wander through the solution space searching for the best possible representation achieving the minimum error rates. In such a system, the fitness function was chosen to be the error rates achieved by a suitable classifier. The details of the classifiers chosen will be given in the experiments section.

### B. Application of PSO in Fuzzy Discriminant Analysis

The aim of this section is to help the reader understand how to apply PSO in discriminant analysis. As mentioned before, the equation for both the between-class scatter matrix $S_B$ and the within-class scatter matrix $S_W$ require the multiplication by the weight vector $w$ that consists of the weights values associated with all features, i.e., $w = \{w_1, w_2, \ldots, w_f\}$. This paper proposes to utilize PSO to find the optimal values for $w$, as shown in Fig. 1.

In the first steps, the particles and their associated velocities are initialized with random numbers. Each of the particles will hold one possible representation for the weight vector based on which the scatter matrices are computed. After computing the scatter matrices, the datasets (training and validation) are projected. The optimality of each particle, i.e., the fitness, is evaluated using a classical wrapper approach in which a suitable classifier is chosen as will be mentioned later in the experiments section. Then using the fitness values for all particles, the local best ($pbest_i$) and global best ($gbest$) for the whole swarm are updated. The stopping criterion chosen in this paper was to stop the PSO optimization after reaching a certain number of iterations. In such a procedure, the weight vector that will correspondingly produce the minimum training and validation errors will be the optimal one. When the procedure finishes the testing dataset is employed to test the generalization capability of this technique on completely unseen data.

### IV. EXPERIMENTS AND RESULTS

In order to present a fair comparison with the available techniques, we include many of them in the experiments. The details of the experiments carried on are listed below:

- **Comparison with other methods**: The PSOFLDA will be compared against two groups of other techniques: the first has already been applied into myoelectric control like ULDA [6], and PCA [3]. The second group include technique that were not used within the myoelectric control problems, like Orthogonal Linear Discriminant Analysis (OLDA) [16], and Fuzzy Discriminant Analysis (FLDA) [9]. These were included because they represent new variations to Fisher’s LDA.

- **Datasets employed**: Since this work aims to present a novel variation to the existing techniques, a comparison with the existing techniques is necessary on different datasets before employing it on an MES dataset. For this reason, two sets of experiments are conducted. In the first, the proposed PSOFLDA is tested on datasets acquired from the Machine Learning Repository (www.ics.uci.edu/~mlearn/mlrepository.html) with different number of samples and numbers of features. In the second experiment, an MES dataset collected from thirty subjects is employed.

  - **Testing method employed**: The general testing scheme employed is a three way data split. The dataset utilized is divided into three sets: training, validation, and testing. An initial projection matrix is calculated based on the training set. Then a validation set is used in order to optimize the weights to produce the optimum projection matrix that can minimize the mean of the training and validation errors. Finally a completely unseen testing set is utilized to measure the generalization capability of the proposed system.

  - **Parameters of PSO**: Specifically the following parameters values were used: maximum number of generations, 30; maximum velocity $s_{max}$: 20% of the range of the corresponding variable; maximum value along a specific dimension $p_{max}$ = 1 and minimum $p_{min}$ = 0; $w$ decreases linearly from 0.9 to 0.4; and acceleration constants $c_1, c_2$ are set to 2.0.

  - **Classifier Type**: Three different classifiers will be utilized to validate the results. The first is a K-Nearest Neighbor classifier (k-NN). The second is Support Vector Machine classifier (SVM) for which the LIBSVM package available online at (http://www.csie.ntu.edu.tw/~cjlin/libsvm/) is utilized. The third classifier is the Linear Discriminant Analysis classifier (LDA).

#### A. Experiments on UCI datasets

Each dataset taken from the UCI Repository is subdivided into three parts, with the percentage of the data forming each of the training, validation and testing given as 20%, 20% and 60% respectively. The classification error results shown in Table 1 were acquired by utilizing a k-NN classifier with the number of neighbors ($k$) being 5, while Table II gives the classification error results obtained using the SVM classifier. The results indicate that there are significant variations between the results obtained by FLDA, ULDA, and OLDA on the different datasets, in comparison with PCA. Meanwhile, the proposed PSOFLDA proves to outperform all of those techniques on all of the datasets utilized in a significant manner. This also proves the effectiveness of the PSOFLDA while training and validating on a very small number of samples and testing on a bigger number of samples, thus better generalization is achieved by PSOFLDA.

#### B. Experiments on MES datasets

The MES dataset utilized in this research was originally collected and used by Chan et al [7]. Eight channels of surface MES were collected from the right arm of thirty normally limbed subjects (twelve males and eighteen females). Each session consisted of six trials. Seven distinct limb motions were used, hand open (HO), hand close (HC), supination (S), pronation (P), wrist flexion (WF), wrist extension (WE), and rest state (R). Data from the first two trials were used
as training set and data from the remaining four trials were divided equally into two trials for validation (trails 1 and 2) and two trails for testing (trails 3 and 4).

As a first part of the MES pattern recognition system, two sets of features were extracted from the raw data in order to test the performance of the proposed method with different feature extraction techniques. The first set of extracted features included a combination of the first four autoregressive (AR) coefficients and the root mean square value (Time-Domain (TD) feature), i.e., number of features = 40 (8 channels × 5 features/channel). This feature set was referred to as the TDAR feature set. The second feature set extracted included the mean of the square values of the wavelet coefficients using a Symmlet wavelet family with five levels of decomposition, i.e., number of features = 48 (8 channels × 6 features/channel). This feature set was referred to as the WT feature set. The analysis window size was 256 msec. Data that were 256 msec before or after a change in limb motion were removed from the training set to avoid transitional data. As a dimensionality reduction part, all of the following five methods: PSOFLDA, ULDA, OLDA, FLDA, and PCA were utilized to compare their performance. The final step of the MES recognition system involves a suitable classifier that can be chosen at the disposal of the designer. In the current experiments a Linear Discriminant Analysis (LDA) classifier was chosen. The advantage of this classifier is that it does not require iterative training, avoiding the potential for under- or over-training [7].

The classification accuracy results averaged across thirty subjects (with one standard deviation) using both the TDAR and the WT feature sets reduced in dimensionality with PSOFLDA, ULDA, OLDA, FLDA, and PCA are shown in Fig.2. The number of extracted features from all methods was set to $c - 1$, where $c$ is the number of classes, as the discriminant analysis based techniques usually ends up with $c - 1$ features. The results shown for both the validation and testing sets were given first without post processing (referred to as Initial), then with a majority vote (MV) as a post processing step, followed by the transitional data between classes removed (NT), and finally with both majority vote and the removal of the transitional data (MV+NT). The results for both the validation and testing sets are given in the Table-III.

It is clear from the results that the PSOFLDA was able to outperform all other methods. This is due to the fact the PSOFLDA is assigning higher importance to good features compared to those that are less useful. At the same time, the
use of the classification accuracy as a judgment criterion on the weight values moved the projection matrix closer toward the optimal projection matrix than all other techniques. Also the PSOFLDA assigns lower fuzzy membership values to the outlier points, thus reducing their effect. Another issue to be mentioned here is that with all of the feature projection techniques, the WT features achieved higher accuracies than that achieved using the simple TDAR features. But from computational cost point of view the performance of the system with the TDAR features is still highly accepted.

In order to provide a rigorous validation or comparison with existing techniques for dimensionality reduction, the confusion matrix for all the subjects was also computed for the different feature sets. A plot of the diagonal values of the confusion matrices (class wise classification accuracy) validation and testing sets are presented in Fig. 3. All the results indicate that there were more significant enhancements when applying the PSOFLDA method than that of the other techniques. Finally, we also provide the statistical significance test results for the proposed PSOFLDA against all other method.

A two-way-analysis of variance (ANOVA) test was carried out with the significance level set to 0.05 and the test results

---

**TABLE I**

Results on data obtained from the UCI Repository averaged across 10 runs, using KNN classifier, with K=5

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Divisions</th>
<th>PCA</th>
<th>FLDA</th>
<th>ULDA</th>
<th>OLDA</th>
<th>NPE</th>
<th>LPP</th>
<th>PSOFLDA</th>
</tr>
</thead>
<tbody>
<tr>
<td>German</td>
<td>Train</td>
<td>26.0101</td>
<td>18.8889</td>
<td>11.7172</td>
<td>29.3435</td>
<td>18.5354</td>
<td>7.1747</td>
<td>6.2096</td>
</tr>
<tr>
<td>Dermatology</td>
<td>Train</td>
<td>3.5870</td>
<td>0.7126</td>
<td>0.5737</td>
<td>0.4288</td>
<td>1.1353</td>
<td>0.4288</td>
<td>1.2863</td>
</tr>
</tbody>
</table>

---

**TABLE II**

Results on data obtained from the UCI Repository averaged across 10 runs, using SVM classifier (LIBSVM package)

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Divisions</th>
<th>PCA</th>
<th>FLDA</th>
<th>ULDA</th>
<th>OLDA</th>
<th>NPE</th>
<th>LPP</th>
<th>PSOFLDA</th>
</tr>
</thead>
<tbody>
<tr>
<td>German</td>
<td>Train</td>
<td>17.0161</td>
<td>18.8889</td>
<td>29.3435</td>
<td>18.5354</td>
<td>13.9899</td>
<td>18.5354</td>
<td>18.0808</td>
</tr>
<tr>
<td>Ionosphere</td>
<td>Train</td>
<td>19.2647</td>
<td>2.2059</td>
<td>0.8823</td>
<td>8.2353</td>
<td>35.2941</td>
<td>8.2353</td>
<td>5.2941</td>
</tr>
<tr>
<td>WDBC</td>
<td>Train</td>
<td>27.1338</td>
<td>15.7795</td>
<td>48.0157</td>
<td>15.7008</td>
<td>3.2283</td>
<td>15.7952</td>
<td>15.9055</td>
</tr>
<tr>
<td>Wine</td>
<td>Train</td>
<td>4.3750</td>
<td>0.3125</td>
<td>0.0000</td>
<td>0.0000</td>
<td>1.5257</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>Validate</td>
<td>6.9369</td>
<td>5.6687</td>
<td>5.6687</td>
<td>5.9251</td>
<td>6.4378</td>
<td>6.6943</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>Test</td>
<td>5.7943</td>
<td>3.2710</td>
<td>5.6687</td>
<td>5.9251</td>
<td>6.4378</td>
<td>6.6943</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

---

**TABLE III**

Classification results achieved by different methods

<table>
<thead>
<tr>
<th>Feature</th>
<th>Divisions</th>
<th>PSOFLDA</th>
<th>FLDA</th>
<th>ULDA</th>
<th>OLDA</th>
<th>PCA</th>
</tr>
</thead>
<tbody>
<tr>
<td>TDAR</td>
<td>Validate</td>
<td>95.02</td>
<td>92.25</td>
<td>92.38</td>
<td>92.38</td>
<td>83.51</td>
</tr>
<tr>
<td></td>
<td>Test</td>
<td>93.68</td>
<td>91.67</td>
<td>91.79</td>
<td>91.79</td>
<td>81.93</td>
</tr>
<tr>
<td>WT</td>
<td>Validate</td>
<td>96.26</td>
<td>93.28</td>
<td>93.35</td>
<td>93.35</td>
<td>89.90</td>
</tr>
<tr>
<td></td>
<td>Test</td>
<td>94.60</td>
<td>92.32</td>
<td>92.44</td>
<td>92.44</td>
<td>88.21</td>
</tr>
</tbody>
</table>

---
V. CONCLUSION

In this paper, a novel feature projection technique based on a mixture of fuzzy logic and Fisher’s LDA was developed. Unlike the typical variations to LDA, the new technique assigned higher importance to good features compared with others. The importance was based on a weighting scheme that was optimized with PSO technique. This in turn caused the PSOFLDA’s projection matrix to be closer to the optimal. The proposed PSOFLDA technique was fairly compared with other techniques like FLDA, ULDA, OLDA, and PCA proving to present better results on real MES datasets. This in turn proves the ability of the proposed technique in enhancing the performance of the multifunction myoelectric hand control system.

ACKNOWLEDGMENT

The Authors would like to acknowledge the support of Dr. Adrian D. C. Chan from Carleton University for providing us with the MES datasets.

REFERENCES


Linear Discriminant Analysis for Chemical Data Sets. *Chemometrics and Intelligent Laboratory Systems*, 45(1-2), 295-302.


---

**TABLE IV**

<table>
<thead>
<tr>
<th>Method</th>
<th>TDAR</th>
<th>WT</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSOFLDA vs.</td>
<td>0.00E+00</td>
<td>1.75E-04</td>
</tr>
<tr>
<td>PCA</td>
<td>3.17E-05</td>
<td>0.0013</td>
</tr>
<tr>
<td>OLDA</td>
<td>3.17E-05</td>
<td>0.0013</td>
</tr>
<tr>
<td>ULDA</td>
<td>1.44E-05</td>
<td>9.56E-04</td>
</tr>
<tr>
<td>FLDA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>