International Journal on Advanced Science Engineering Information Technology

Vol.10 (2020) No. 5
ISSN: 2088-5334

Performance Evaluation of SRELM on Bio-signal Pattern Recognition Using Two Electromyography Channels

Khairul Anam[a], Adel Al-Jumaily[b]

[a]*Faculty of Electrical Engineering, Universitas Jember, Jl. Kalimantan 37, Jember, 68121, Indonesia*
E-mail: khairul@unej.ac.id

[b]*School of Biomedical Engineering, University of Technology Sydney, 15 Broadway, Ultimo NSW 2007, Australia*
E-mail: adel.al-Jumaily@uts.edu.au

Abstract*— The classification accuracy of pattern recognition is determined by the extracted features and the utilized classifiers. Many efforts have been conducted to obtain the best features either by introducing a new feature or proposing a new projection method to increase class separability. Recently, spectral regression extreme learning machine (SRELM) has been introduced to improve the class separability of the features. However, the evaluation of SRELM was only focused on the myoelectric or electromyography pattern recognition from many EMG channels. In practical application, the user is more convenient with less number of channels. Then, the problem is whether the SRELM would be able to work efficiently for less EMG channels. The objective of this paper is to examine the performance of SRELM for bio-signal pattern recognition using two EMG channels. The EMG electrodes were located on flexor policies lounges and flexor digitorium superficial muscles of ten healthy subjects. Various time domain features were involved with various sizes. SRELM will project these features to more recognize features before being feed to multiple classifiers. Those classifiers are randomized Variable Translation Wavelet Neural Networks (RVT-WNN), extreme learning machine(ELM), support vector machine (SVM), and linear discriminant analysis (LDA). The performance of SREM was compared to other feature methods, such as LDA, uncorrelated LDA (ULDA), orthogonal fuzzy neighborhood dimensionality reduction (OFNDA), and spectral regression discriminant analysis (SRDA). The experimental results show that SRELM performed well when dealing with different class numbers by classification accuracy of around 95.67% for ten class movements and performed better than SRDA.

***Keywords*— myoelectric; pattern recognition; dimensionality reduction.**

I. INTRODUCTION

The artificial neural network using backpropagation has been used widely for decades in many applications. The weights on the hidden layers, as well as on the output layers, are trained using a famous backpropagation algorithm. For speeding up the learning time, a new idea was emerged by randomizing the weights on the hidden layer. For the first time, this idea was introduced by Schmidt et al. [1]. The idea was proved experimentally. A couple of years later, Huang et al. [2] demonstrated the efficacy of the random weights theoretically and experimentally by introducing an extreme learning machine (ELM) [3].

ELM is a single hidden layer feed-forward neural network. The randomization is applied to the weights of the hidden layer only. Meanwhile, the weights of the output layer are determined analytically using least square methods. Therefore, the training process runs very fast compared to the gradient descent method. ELM has entered various areas, including classification, regression, and dimensionality reduction. For classification, ELM has been applied for bio-signal pattern recognition [4], [5], character recognition [6], [7], face recognition [8], protein structure detection, and cancer detection [9]. As for regression, ELM has shown its efficacy for the estimation of the physical parameters [10] and the electrical power system [11].

Besides classification and regression, ELM can be employed to overcome a curse dimensionality of the features. This idea was proposed by Huang et all. [12] who introduced an unsupervised ELM. This ELM reduces the feature's dimension without labels, as in principal component analysis (PCA). In the case of the existence of the label, the linear discriminant analysis is preferable. Martinez et all. [13] showed that, in many cases, LDA is better than PCA. To accommodate the known labels in this unsupervised ELM, spectral regression extreme learning machine (SRELM) was proposed [14].

SRELM is a combination of spectral regression (SR) and ELM. As in a normal ELM, the hidden weights of SRELM are set randomly. Meanwhile, the role of SR is to calculate the output weights using spectral analysis. Phukpattaranont et all. [15] has evaluated SRELM for finger movement recognition using 6 electromyography (EMG) channels. The experimental results indicated that SRELM outperformed over other tested methods.

The good performance of SRELM [15] was achieved when using 6 EMG channels. Sometimes, using many channels is not convenient for amputees. Therefore, the myoelectric pattern recognition employing fewer channels is preferable. Few publications considering less EMG channels have been reported [16, 17]. Therefore, it is interesting and challenging to examine the performance of SRELM myoelectric pattern recognition using fewer channels. In this paper, two EMG channels were selected. The main contribution of this paper is the evaluation of SRELM for myoelectric pattern recognition using two-channel. It is interesting to examine its performance when the availability of the bio-signal is not adequate. The organization of this paper is as follows. The first section presents the introduction. The coming section provides the basic theory of SRELM and the experimental methodology. Section III presents the experimental results and discussions. Finally, Section IV is the conclusion.

II. MATERIALS AND METHOD

A. Theory of SRELM

Spectral Regression Extreme learning machine or SRELM is a combination of spectral regression and extreme learning machine for feature projection or also known by dimensionality reduction. SRELM is constructed from unsupervised extreme learning developed by Huang et all. [12] for unsupervised dimensionality reduction. Basically, this unsupervised ELM is similar to a famous unsupervised dimensionality projection, i.e. principal component analysis (PCA). PCA reduces the dimension of the features without a label. Different from PCA, linear discriminant analysis (LDA) reduces the feature dimension with known labels. Martinez et all. [13] show that, in many cases, LDA outperforms PCA. Starting from that point, spectral regression was introduced to the unsupervised ELM to incorporate the known labels to the unsupervised ELM.

As for ELM, it is single layer feed-forward neural networks (SLFNs) that the hidden weights are set randomly. On the other hand, the output weights are calculated based on the least square methods. Assume, there are N samples data $\{(x_i, t_i)\}_{i=1}^{N} \in R^n \times R^m$, and L hidden nodes, the ELM output is

$$f(\mathbf{x}_i) = \sum_{j=1}^{L} \beta_j G\left(\mathbf{a}_j, b_j, \mathbf{x}_i\right) = \mathbf{h}(\mathbf{x}_i)\boldsymbol{\beta} = t_i, \;\; i = 1, ..., N \tag{1}$$

where $\mathbf{h}(\mathbf{x}_i) \in R_{NxL}$ and $\beta \in R_{Lxm}$. Here, $\mathbf{h}(\mathbf{x}_i)$ is determined randomly. Meanwhile, the output weights are calculated by minimization of the sum of squared prediction error. It is described by:

$$\begin{aligned} &\text{Minimize } : L_{ELM} = \tfrac{1}{2}\|\boldsymbol{\beta}\|^2 + C\tfrac{1}{2}\sum_{i=1}^{N}\|e\|^2 \\ &\text{Subject to} : \mathbf{h}(\mathbf{x}_i)\boldsymbol{\beta} = t_i^T - e_i^T \qquad i = 1, ..., N \end{aligned} \tag{2}$$

The substitution of the constraint into the objective function yields:

$$\text{Minimize } : L_{ELM} = \tfrac{1}{2}\|\boldsymbol{\beta}\|^2 + C\tfrac{1}{2}\sum_{i=1}^{N}\|\mathbf{T} - \mathbf{H}\boldsymbol{\beta}\|^2 \tag{3}$$

in which $H = [\mathbf{h}(\mathbf{x}_1), \ldots, \mathbf{h}(\mathbf{x}_N)^T]^T \in R^{NxL}$ and $T \in R^{Nxm}$. From Eq. (3), we can obtain the gradient system with respect to β:

$$\nabla L_{ELM} = \boldsymbol{\beta} + C\mathbf{H}^T\left(\mathbf{T} - \mathbf{H}\boldsymbol{\beta}\right) = 0 \tag{4}$$

From Eq. (4), two solutions of β can be obtained, subject to the H. The first solution is when **H** has fewer columns than rows:

$$\boldsymbol{\beta} = \left(\mathbf{H}^T\mathbf{H} + \frac{I_L}{C}\right)^{-1}\mathbf{H}^T\mathbf{T} \tag{5}$$

where I_L is an identity matrix. The second solutions is when H has fewer rows than columns, i.e.:

$$\boldsymbol{\beta} = \mathbf{H}^T\left(\mathbf{H}\mathbf{H}^T + \frac{I_N}{C}\right)^{-1}\mathbf{T} \tag{6}$$

ELM can be extended for a dimensionality reduction by considering unknown labels [12]. For that aim, the objective function in Eq. (2) can be modified as:

$$\begin{aligned} &\text{Minimize } : L_{ELM} = \tfrac{1}{2}\|\boldsymbol{\beta}\|^2 + \lambda\tfrac{1}{2}\mathrm{Tr}(\mathbf{F}^T\mathbf{L}\mathbf{F}) \\ &\text{Subject to} : \boldsymbol{f}_i = \mathbf{h}(\mathbf{x}_i)\boldsymbol{\beta} \qquad i = 1, ..., N \end{aligned} \tag{7}$$

where L is a graph laplacian and N is the number of samples.. Substitution of the constraint to the objective function gives:

$$\begin{aligned} &\text{Minimize } : L_{ELM} = \tfrac{1}{2}\|\boldsymbol{\beta}\|^2 + \lambda\tfrac{1}{2}\mathrm{Tr}(\boldsymbol{\beta}^T\mathbf{H}^T\mathbf{L}\boldsymbol{\beta}\mathbf{H}) \\ &\text{Subject to } : \boldsymbol{\beta}^T\mathbf{H}^T\mathbf{L}\boldsymbol{\beta}\mathbf{H} = \mathbf{I}_m \end{aligned} \tag{8}$$

Huang et all. [12] has proved that the optimal solution of Eq. (8) is eigenvalue of.:

$$\left(\mathbf{I}_L + \lambda\mathbf{H}^T\mathbf{L}\mathbf{H}\right)\mathbf{u} = \gamma\mathbf{H}^T\mathbf{L}\mathbf{H}\mathbf{u} \tag{9}$$

In spectral regression [18, 19], a graph is mapped to real line y by a linear function:

$$y = \mathbf{H}\mathbf{u} \tag{10}$$

Thus, Eq. (9) becomes:

$$\left(\mathbf{I}_L + \lambda\mathbf{H}^T\mathbf{L}\mathbf{H}\right)\mathbf{u} = \gamma\mathbf{H}^T\mathbf{L}\mathbf{y} \tag{11}$$

To find the optimal y, we have to minimize:

$$\sum_{i,j}\left(y_i - y_j\right)2\mathbf{W}_{ij} = 2\mathbf{y}^T\mathbf{L}\mathbf{y} \tag{12}$$

The value of L is obtained by subtracting W from D. D is a diagonal matrix, while W is an N x N matrix. Finally, the solution is by solving the maximum eigenvalue problem [18]:

$$\mathbf{W}\mathbf{y} = \lambda\mathbf{D}\mathbf{y} \tag{13}$$

To summarize, the solution to Eq. (13) is done in two steps. The first one is solving the eigenvalue problem and then finding **u** that satisfies **Hu = y** by employing:

$$\mathbf{u} = \arg\min_u\left(\sum_{i=1}^{N}\left(\mathbf{u}^T h(x_i) - y_i\right)^2 + \alpha\sum_{j=1}^{L}u_j\right) \tag{14}$$

Here, u_j is the **u** component and α is a regression parameter. Therefore, the output weigh is given by:

$$\boldsymbol{\beta} = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{c-1}] \in \mathrm{RL} \tag{15}$$

B. Method

The primary goal of this research is to evaluate SRELM for myoelectric pattern recognition. Myoelectric signal (MES) or electromyography (EMG) utilized in this paper were collected in [14]. The MES was acquired from two channels from 10 subjects. In this research, one channel is added from the summation of these two original channels to get a new MES. From these three channels, time-domain features (TD) were extracted. It involved zero waveform lengths (WL) (3 features), slope sign changes (SSC) (3 features), number of zero crossings (ZCC) (3 features), sample skewness (SS) (3 features), mean absolute value (MAV) (3 features), Hjorth-time domain parameters (HTD) (9 features) and 6-order autoregressive parameters (AR) (18 features).

The total number of features extracted was 42. However, later, more features are added to observe the performance of SRELM dealing with a wide range of features. For segmentation, we apply disjoint windowing with a window length of 100 ms every 100 ms. In this test, we examined the performance of SRELM to reduce the features from MES from various features.

TABLE I describes the various features used in the experiment.

TABLE I
VARIOUS FEATURES USED IN THE EXPERIMENT

Name	#features	Features	Group
F1	12	SSC,ZC, WL, MAV	Small dimension
F2	15	SSC,ZC,WL,MAV, MAVS	
F3	24	SSC,ZC,WL,MAV, SKW, HJORTH	Medium dimension
F4	36	SSC,ZC,WL, MAV, MAVS, RMS,6AR	
F5	42	SSC,ZC,WL, MAV,SKW,HJORTH,6AR	
F6	48	SSC,ZC,WL,SKW,MAV,MAVS, HJORTH,RMS,6AR	
F7	195	Power autoregressive	Large dimension

Furthermore, several classifiers were employed, along with the majority vote with n=4. They are RVT-WNN (randomized Variable Translation Wavelet Neural Networks)[5], RBF-ELM (radial basis extreme learning machine), SVM (support vector machine), and LDA (linear discriminant analysis). This experiment conducted 3-fold cross-validation.

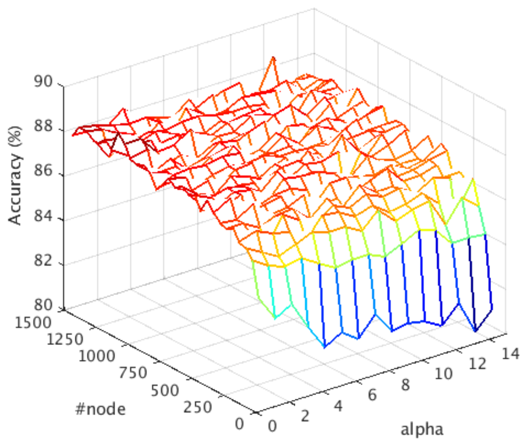
As for movements, there are ten fingers movements, i.e., Thumb (T), Index finger (I), Middle finger (M), Ring finger (R), Little finger (L), Thumb finger – Index finger (T-I), Thumb finger – Middle finger (TM), Thumb finger – Ring finger (T-R), Thumb finger - Little finger (T-L), and Hand-Closed (HC)

III. Results and Discussion

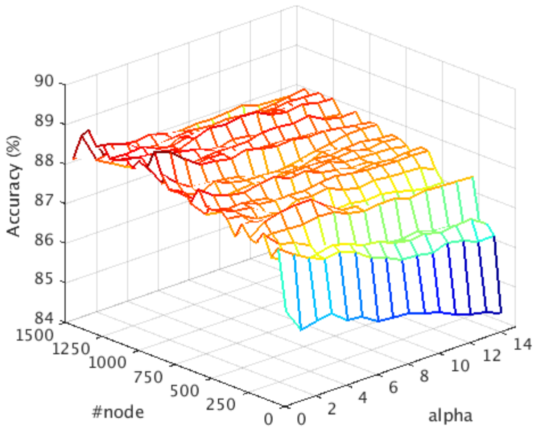
A. The optimization Parameters

SRELM has two parameters that should be chosen properly. These are the number of hidden nodes and alpha α. They are optimized using a grid search method. The classification accuracy is employed to measure the performance of the myoelectric pattern recognition using two EMG channels. Two classifiers were used, RVT-WNN and RBF-ELM. The number of hidden nodes of VRT-WNN, was 100. Meanwhile, the parameters of RBF-ELM were 2^0 and 2^{-5} for C and gamma, respectively. The results are presented in Fig. 1.

Fig. 1 shows that the big number of nodes produces better accuracy than the small one. However, the accuracy does not increase significantly for the number node more than 500. Controversy, the small number of alpha yields better accuracy than the big one. Considering these two trends, finally, the 1000 nodes and 0.05 of alpha were chosen for the next experiment.



(a)



(b)

Fig. 1. The correlation of alpha (α) and the number of nodes using classifier RVT-WNN (a) and RBF-ELM (b)

B. Class separability

We compare the performance of the dimensionality reduction from a scatter plot of the data. The performance of

SRELM is compared to linear discriminant analysis (LDA), uncorrelated LDA (ULDA), orthogonal fuzzy neighborhood dimensionality reduction (OFNDA), and spectral regression discriminant analysis (SRDA), and. Fig. 2 describes the scatter plot of the three first features of the original feature set before being projected. The picture indicates that the data are scattered completely. The projection methods could improve the class separability of the data, as shown in Fig. 3 to Fig. 6.

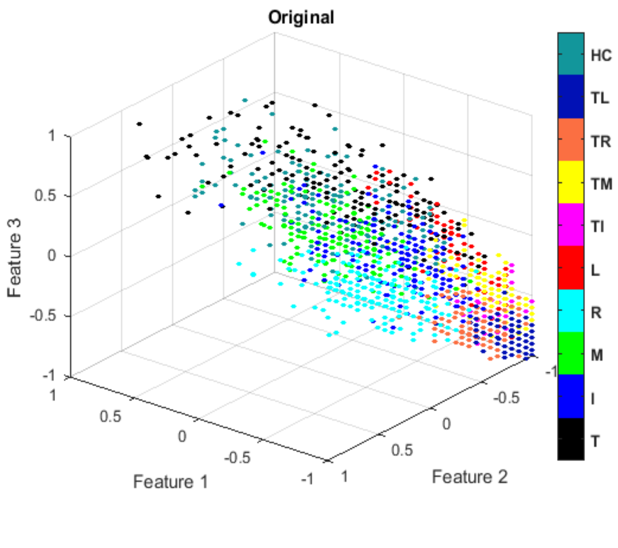


Fig. 2 Scatter plot of the original features before projected

Fig. 3 and Fig. 4 presents the scatter plot of ULDA and OFNDA, respectively. Both methods could enhance the class separability of features by grouping the data according to the class. Similarly, Fig. 5 and Fig. 6 describing the scatter plot of the features using SRDA and SRELM indicates that SRDA and SRELM could improve the class separability of the features, compared with **Fig. 2**. We can compare the scatter plot of SRDA and SRELM in Fig. 6. The figure shows that features projected using S-ELM are slightly better than SRDA.

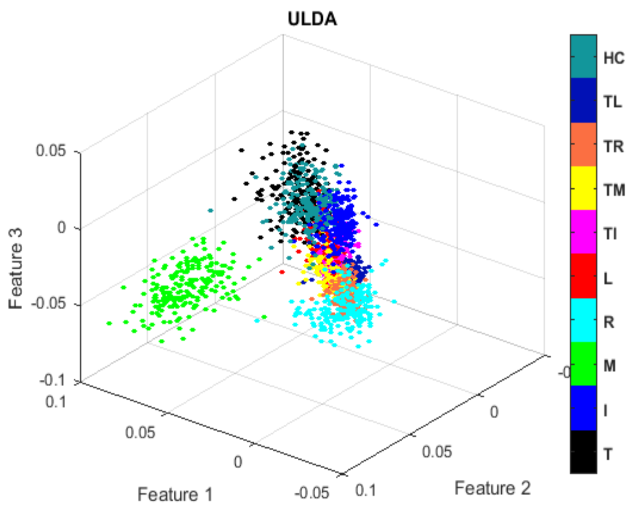


Fig. 3 Scatter plot of the projected features using ULDA

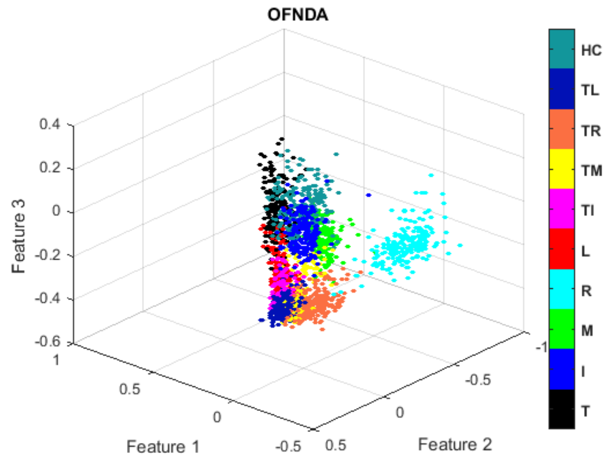


Fig. 4 Scatter plot of the projected features using OFNDA

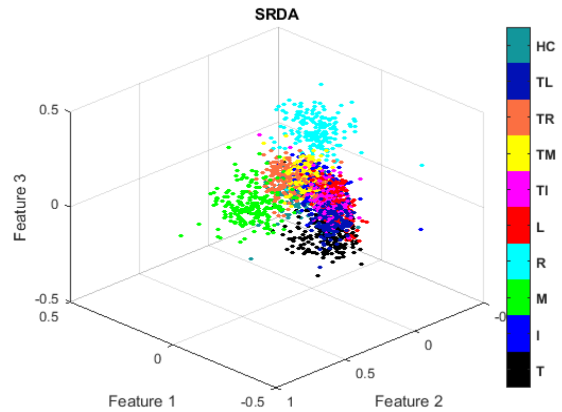


Fig. 5 Scatter plot of the projected features using SRDA

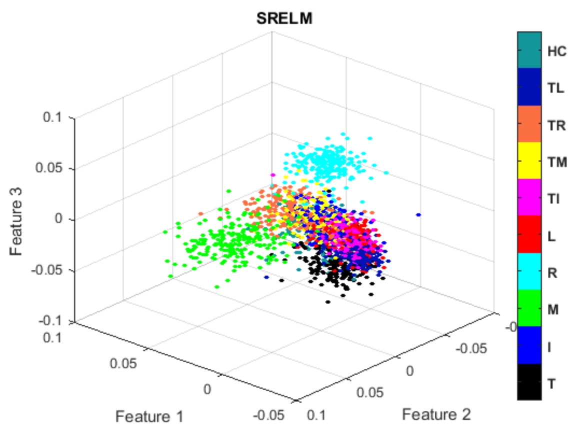


Fig. 6 Scatter plot of the projected features using SRELM

C. Different feature

In this experiment, few methods were added, such as PCA (principal component analysis), USELM (unsupervised extreme learning machine), and BASELINE (without dimensionality reduction method). TABLE II provides the experimental results when using RVT-WNN as a classifier to classify ten finger motions from the features reduced using various dimensionality reduction methods.

TABLE II
THE ACCURACY ACHIEVED EMPLOYING RVT-WNN ON THE FEATURE TEST USING 3-FOLD CROSS-VALIDATION

Set	# features	Accuracy (100%)						
		UL-DA	SR-DA	SRELM	OFN-DA	PCA	US-ELM	BASE-LINE
F1	12	88.8	88.4	**90.5**	86.2	87.6	80.5	89.5
F2	15	89.7	88.8	**91.4**	90.8	81.8	78.7	89.5
F3	24	93.6	93.3	**94.0**	93.5	90.4	87.9	93.0
F4	36	93.8	93.7	93.7	**93.9**	85.4	79.5	91.5
F5	42	**94.3**	94.2	94.2	94.2	88.6	84.5	92.2
F6	48	94.8	94.8	94.7	**94.9**	88.9	85.1	92.5
F7	195	63.9	94.0	92.5	**94.9**	88.8	85.1	92.5

According to results in TABLE II, SRELM achieved the highest accuracy for the feature set F1 up to F3. For the feature set F4 - F6, the accuracy of the system is not the highest, but it is very close to the highest one. Another interesting fact is revealed when we compare SRDA and SRELM. In all features, SRELM attained better accuracy than SRDA except on the feature set F6 and F7. Especially on the feature set F7, whose dimension is very large, the difference of SRDA and SRELM is noticeable. Moreover, SRELM is better than unsupervised dimensionality reduction, PCA, and US-ELM. The comparison of SRELM and Baseline shows that the SRELM could reduce the dimension of data and, at the same time, could increase the class separability of the features. The evidence is on the accuracy of the Baseline. Its accuracy is lower than SRELM except on the feature set F7. It seems that SRELM does not perform well on a large dimension of data.

In addition to RVT-WNN, LDA also classified the ten-finger movements along with various dimensionality reduction methods. TABLE III presents the experimental results. The table shows that SRELM is the most accurate method across six features sets: F1 to F6. However, SRELM is worse than SRDA when projecting the feature set F7, but it is still better than ULDA and OFNDA. This new fact confirms the previous assumption about SRELM that the accuracy is slightly low when it works on the large dimension of features. In general, compared to PCA, US-ELM, and the Baseline, SRELM attained better accuracy.

TABLE III
THE ACCURACY ACHIEVED EMPLOYING LDA ON THE FEATURE TEST USING 3-FOLD CROSS-VALIDATION

Set	#fea-tures	Accuracy (100%)						
		ULDA	SRDA	SR-ELM	OFN-DA	PCA	US-ELM	BASELINE
F1	12	85.7	84.2	**88.0**	82.1	82.8	73.1	85.7
F2	15	86.9	85.1	**88.8**	86.9	78.2	70.5	86.9
F3	24	92.5	92.3	**93.3**	92.5	88.5	82.4	92.5
F4	36	93.0	92.7	**93.6**	93.0	84.6	75.2	93.0
F5	42	93.3	93.3	**94.1**	93.3	87.7	80.9	93.3
F6	48	94.2	93.9	**94.6**	94.2	88.1	81.2	94.2
F7	195	92.1	**93.4**	92.8	92.1	91.3	87.2	92.1

By looking at TABLE IV, TABLE V, and facts mentioned in the previous discussion in this section, we can conclude that SRELM is very good for reducing the feature with low up to medium dimension. Compared to SRDA, the performance of SRELM is lower than SRDA. However, its accuracy is still higher than the unsupervised dimensionality reduction (PCA and USELM) and the baseline.

TABLE IV
THE ACCURACY ACHIEVED EMPLOYING RBF-ELM ON THE FEATURE TEST USING 3-FOLD CROSS-VALIDATION

Set	# fea-tures	Accuracy (100%)						
		ULDA	SR-DA	SRELM	OFN-DA	PCA	US-ELM	BASE-LINE
F1	12	90.6	87.3	91.1	85.5	80.2	81.2	80.8
F2	15	91.8	89.2	92.2	89.1	78.1	79.8	80.2
F3	24	94.4	93.1	94.4	92.3	87.2	88.5	87.8
F4	36	94.7	93.8	94.3	93.4	78.6	80.5	83.2
F5	42	94.8	93.8	94.5	93.3	82.4	85.1	86.6
F6	48	95.5	94.6	94.8	94.4	83.5	85.9	87.3
F7	195	92.5	93.1	93.0	17.6	87.6	87.3	88.0

TABLE V
THE ACCURACY ACHIEVED EMPLOYING SVM ON THE FEATURE TEST USING 3-FOLD CROSS-VALIDATION

Set	# features	Accuracy (100%)						
		UL-DA	SR-DA	SRELM	OFN-DA	PCA	US-ELM	BASELINE
F1	12	89.3	87.3	89.6	89.2	85.8	75.6	84.9
F2	15	90.6	88.1	90.4	90.5	81.6	73.7	85.6
F3	24	93.5	93.4	93.6	93.6	90.4	84.5	91.9
F4	36	94.2	93.9	93.9	94.2	84.9	77.8	91.8
F5	42	94.3	94.3	94.3	94.3	88.7	82.4	92.9
F6	48	95.2	94.8	94.6	95.2	89.5	83.1	93.3
F7	195	92.4	93.8	92.7	92.4	91.4	87.7	92.4

D. The Experiment on the Class Number

The performance of SRELM was examined for different class numbers. Five until ten classes were involved, as mentioned in II.B. The SRELM's performance is compared to ULDA, SRDA, OFNDA, PCA, USELM, and BASELINE (without dimensionality reduction method). Furthermore, all experiments utilized RVT-WNN as a classifier. For post-processing, a majority vote method was utilized [20]. Fig. 7 presents the experimental result.

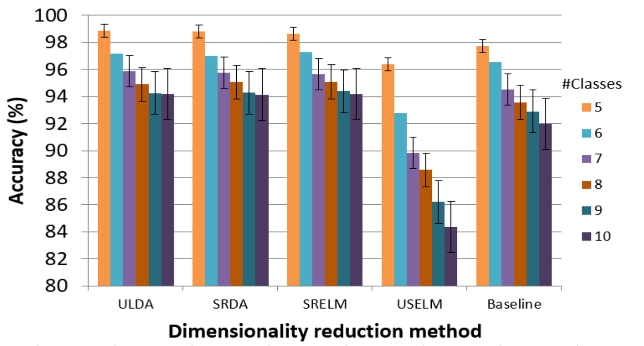


Fig. 7 The comparison of SRELM and other methods in bio-signal pattern recognition using two EMG channels without a majority vote

Fig. 7 shows that the system accuracy is decreasing as the number of classes is increasing because the patterns are getting complex. Here, SRELM worked as well as other linear discriminant analysis such as ULDA, SRDA, SRELM, and OFNDA. However, SRELM performed better than all methods when it was tested in the system that does not use the post-processing method (the majority vote). Its accuracy ranges from 95.67 % to 86.73 % for 5 to 10 classes of movement, as seen in Fig. 7. However, when the system utilized the majority vote, the performance of SRELM and other methods is comparable (see Fig. 8) with accuracy ranging from 98.64% to 94.16% for five to ten motion classes.

SRELM and SRDA employ the same spectral regression. They are different in treating spectral regression (SR). SRDA uses SR to find the eigenvectors for the projection while SR-LEM utilizes SR to obtain the weight output of the extreme learning machine. Another difference is that SRELM involves a random projection as additional to the SR projection. Experimentally, the random projection could improve the performance of the SRDA, as seen in Fig. 7 and Fig. 8. These figures show that SRELM is more accurate than SRDA and even better than other methods across five different classes. It indicates that SRELM enhances the performance of SRDA. However, when the classifiers employed the majority vote, the improvement of the SRELM is not seen significantly. Despite having better performance, the processing time of SRELM is longer than SRDA, as described in Fig. 9. Fortunately, it is still worthy. Even, it is less than the processing time of LDA.

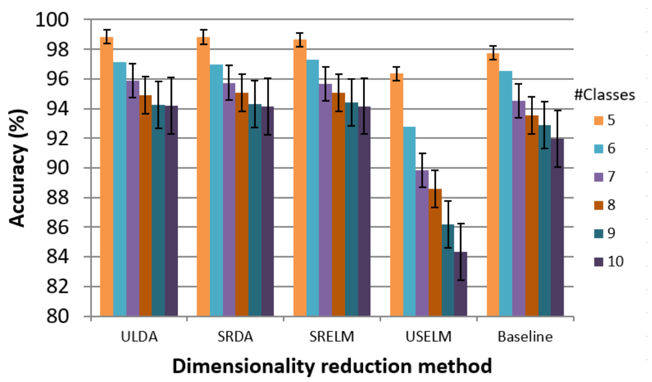


Fig. 8 The comparison of SRELM and other methods in bio-signal pattern recognition using two EMG channels a with a majority vote

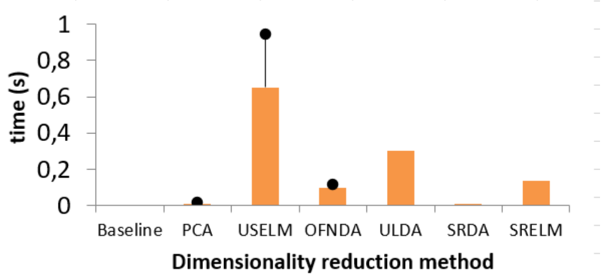


Fig. 9 The processing time

E. Classifier experiments

Different classifiers were involved in examining the performance of SRELM. Fig. 10, and Fig. 11 displays the result. There are two main experiments: with the majority vote and without a majority vote. Fig. 10 showed the performance of SRELM when the classifier dropped the majority vote. The accuracy of the pattern recognition system using SRELM is the highest when using classifier kNN, RBF-ELM, and RVT-WNN. However, when the classifiers utilized the majority vote (see Fig. 11), the influence of SRELM is not noticeable. Involving the majority vote could improve the performance, but it can increase the processing time [21].

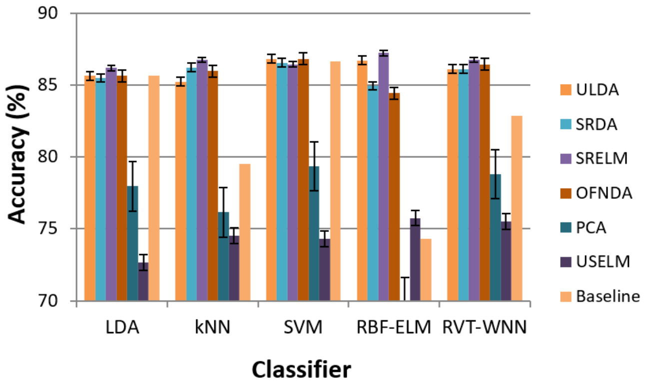


Fig. 10 SRELM performance on different classifiers without majority vote across eight subjects

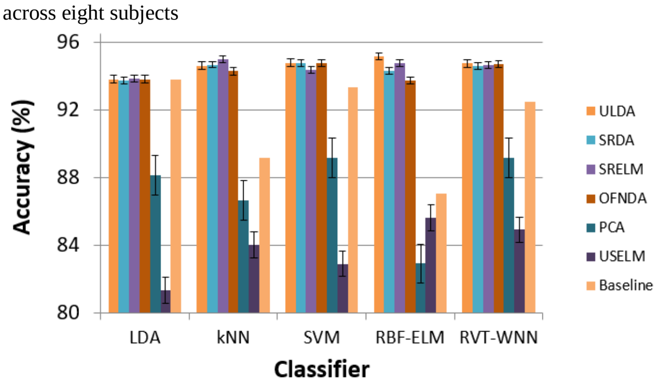


Fig. 11 SRELM performance on different classifiers plus majority vote across eight subjects

Finally, an analysis of variance (ANOVA) test was conducted. The confidence level p was set at 0.05, to understand the significance of SRELM. The results indicate that the improvement of SRELM over SRDA is significant ($p = 0.043 < 0.05$). However, the difference between SRELM and other LDA models like ULDA and OFNDA is not significant ($p=0.142 > 0.05$). This statistic analysis highlights the advantage of SRELM over SRDA.

IV. Conclusion

This paper evaluates the performance of SRELM for bio-signal pattern recognition using two EMG channels. Although using only two channels, the performance of SRELM can still deal with it. Compared to other dimensionality reduction methods, the performance of SRELM is better than SRDA, but it is not significantly different from ULDA and OFNDA. As for the processing time of, SRELM took more time than SRDA, but better than ULDA and OFNDA. Furthermore, SRELM was able to perform well when dealing with various class numbers. The accuracy was around 95.67% for 10 class movements across eight subjects, using only two EMG channels. This result is promising, especially for real-time myoelectric pattern recognition.

REFERENCES

[1] Schmidt, W.F., M.A. Kraaijveld, and R.P. Duin. *Feedforward neural networks with random weights*. in *Pattern Recognition, 1992. Vol. II. Conference B: Pattern Recognition Methodology and Systems, Proceedings., 11th IAPR International Conference on*. 1992. IEEE.

[2] Huang, G.-B., Q.-Y. Zhu, and C.-K. Siew, *Extreme learning machine: Theory and applications.* Neurocomputing, 2006. **70**(1–3): p. 489-501.

[3] Huang, G.B., et al., *Extreme learning machine for regression and multiclass classification.* IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 2012. **42**(2): p. 513-529.

[4] Anam, K. and A. Al-Jumaily, *Evaluation of extreme learning machine for classification of individual and combined finger movements using electromyography on amputees and non-amputees.* Neural Networks, 2017. **85**: p. 51-68.

[5] Anam, K. and A. Al-Jumaily. *Evaluation of Randomized Variable Translation Wavelet Neural Networks*. in *International Conference on Soft Computing in Data Science*. 2017. Springer.

[6] Chacko, B.P., et al., *Handwritten character recognition using wavelet energy and extreme learning machine.* International Journal of Machine Learning and Cybernetics, 2012. **3**(2): p. 149-161.

[7] Zheng, W., Y. Qian, and H. Lu, *Text categorization based on regularization extreme learning machine.* Neural Computing and Applications, 2013. **22**(3-4): p. 447-456.

[8] Mohammed, A.A., et al., *Human face recognition based on multidimensional PCA and extreme learning machine.* Pattern Recognition, 2011. **44**(10): p. 2588-2597.

[9] Saraswathi, S., et al., *ICGA-PSO-ELM Approach for Accurate Multiclass Cancer Classification Resulting in Reduced Gene Sets in Which Genes Encoding Secreted Proteins Are Highly Represented.* IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2011. **8**(2): p. 452-463.

[10] Javed, K., R. Gouriveau, and N. Zerhouni, *SW-ELM: A summation wavelet extreme learning machine algorithm with a priori parameter initialization.* Neurocomputing, 2014. **123**(0): p. 299-307.

[11] Nizar, A.H., Z.Y. Dong, and Y. Wang, *Power Utility Nontechnical Loss Analysis With Extreme Learning Machine Method.* IEEE Transactions on Power Systems, 2008. **23**(3): p. 946-955.

[12] Huang, G., et al., *Semi-supervised and unsupervised extreme learning machines.* IEEE Transactions on Cybernetics, 2014. **44**(12): p. 2405-2417.

[13] Martínez, A.M. and A.C. Kak, *Pca versus lda.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001. **23**(2): p. 228-233.

[14] Anam, K. and A. Al-Jumaily. *A novel extreme learning machine for dimensionality reduction on finger movement classification using sEMG*. in *The 7th International IEEE/EMBS Conference on Neural Engineering (NER)*. 2015. Montpellier, France.

[15] Phukpattaranont, P., et al., *Evaluation of feature extraction techniques and classifiers for finger movement recognition using surface electromyography signal.* Medical & biological engineering & computing, 2018: p. 1-13.

[16] Anam, K., R. Khushaba, and A. Al-Jumaily. *Two-channel surface electromyography for individual and combined finger movements.* in *Proceeding of the 35th Annual International IEEE Engineering Medicine and Biology Society Conference (EMBC)*. 2013.

[17] Al-Timemy, A.H., et al. *Single channel-based myoelectric control of hand movements with Empirical Mode Decomposition*. in *Engineering in Medicine and Biology Society,EMBC, 2011 Annual International Conference of the IEEE*. 2011.

[18] Cai, D., X. He, and J. Han. *Spectral regression: A unified approach for sparse subspace learning*. in *The Seventh IEEE International Conference on Data Mining (ICDM)*. 2007. IEEE.

[19] Cai, D., X. He, and J. Han, *SRDA: An efficient algorithm for large-scale discriminant analysis.* IEEE Transactions on Knowledge and Data Engineering, 2008. **20**(1): p. 1-12.

[20] Simon, A.M. and L.J. Hargrove. *A comparison of the effects of majority vote and a decision-based velocity ramp on real-time pattern recognition control*. in *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE.* 2011. IEEE.

[21] Farrell, T. and R.F. Weir. *Analysis window induced controller delay for multifunctional prostheses*. in *Proceeding of MyoElectric Controls/Powered Prosthetics Symposium*. 2008. Myoelectric Symposium.