Study of EMG Feature Selection for Hand Motions Classification

J. Too[1], A. R. Abdullah[1] and N. M. Saad[2]

[1]*Fakulti Kejuruteraan Elektrik, Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100, Durian Tunggal, Melaka, Malaysia*
[2]*Fakulti Kejuruteraan Elektronik dan Kejuruteraan Komputer, Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100, Durian Tunggal, Melaka, Malaysia*

abdulr@utem.edu.my

Abstract—**In recent days, electromyography (EMG) pattern recognition has becoming one of the major interests in rehabilitation area. However, EMG feature set normally consists of relevant, redundant and irrelevant features. To achieve high classification performance, the selection of potential features is critically important. Thus, this paper employs two recent feature selection methods namely competitive binary gray wolf optimizer (CBGWO) and modified binary tree growth algorithm (MBTGA) to evaluate the most informative EMG feature subset for efficient classification. The experimental results show that CBGWO and MBTGA are not only improves the classification performance, but also reduces the number of features.**

Keywords— Electromyography; feature extraction; time domain feature; feature selection; classification

I. INTRODUCTION

The application of electromyography (EMG) pattern recognition has received much attentions from biomedical researchers. Thanks to current technology, the usage of EMG pattern recognition on myoelectric prosthesis control is becoming viable [1], [2]. However, EMG signal is non-stationary and unstable due to its complex characteristic. Hence, multiple step of processing are needed in order to attain high classification performance [3].

Generally, the processing of EMG signals can be categorized into four parts, which are signal processing, feature extraction, feature selection and classification. However, the signal processing is excluded in this work. This is mainly due to fast processing speed and simplicity [4]. Briefly, feature extraction is a process of extracting the hidden information from the signal. Feature selection attempts to select the most informative features from a large available feature set. Finally, the selected features are fed into the classifier (example: k-nearest neighbor) for performance evaluation [5]–[7].

Recently, feature selection is becoming extremely important in the classification tasks. By selecting the relevant features for classification, the accuracy of the system can be enhanced [8]. In the past study, Krasoulis et al. [9] applied the sequence forward selection (SFS) for evaluating the potential EMG features. Too et al. [6] proposed a new competitive binary grey wolf optimizer (CBGWO) to solve the feature selection problem in EMG signals classification. The authors reported CBGWO can achieve high classification accuracy with a very low computational cost. In the same year, Too et al. [10] introduced another two new feature selection methods namely binary tree growth algorithm (BTGA) and modified binary tree growth algorithm (MBTGA) for EMG feature selection. The authors indicated that MBTGA outperformed BTGA and binary differential evolution (BDE) in evaluating the relevant features. Previous studies have shown the impact of feature selection in EMG signals classification.

However, previous works mainly focus on the analysis of healthy subjects, but not

Article history: Manuscript received 5 February 2019; received in revised form 4 April 2019; Accepted 5 April 2019.

to the trans-radial amputees. Due to the lack of feature selection studies on amputee data sets, we aim to evaluate the performance of feature selection for amputee subjects. In this paper, two recent feature selection methods including CBGWO and MBTGA are employed for EMG signals classification. Initially, the EMG data of amputee subjects are collected from NinaPro database. In the next step, three recent EMG features are extracted and formed the feature vector. The extracted features are then normalized to prevent the numerical issue. Next, CBGWO and MBTGA are used to evaluate the most informative feature subset, and the selected features are fed into the classifier for the recognition tasks. The flow process of proposed EMG recognition system is shown in Fig.1. At the end of this paper, the effectiveness of CBGWO and MBTGA in EMG signals classification are presented.

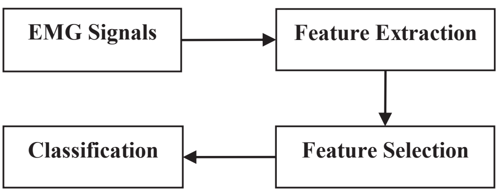

Fig. 1. Flowchart of the proposed EMG recognition system.

II. METHODOLOGY

A. EMG Data

The EMG data are collected from the publicly access NinaPro project. In this study, the surface EMG signals of 17 hand motions, M1-M17 (Exercise B) from 2 amputee subjects from NinaPro database 7 (DB7) are utilized [9]. In the experiment, twelve electrodes were implemented. The subject was asked to perform each hand motion for 5 seconds, followed by a resting phase of 3 seconds. Each hand motion was repeated for six times. Note that the EMG signals were sampled at 2000 Hz [9]. Before any further processing, all the resting phases are removed.

B. Feature Extraction

In this section, the EMG feature extraction is presented. Recently, Samuel et al. [2] proposed three novel EMG features for EMG pattern recognition. In their study, the authors reported that the newly proposed EMG features can achieve better classification performance as compared to traditional EMG features. In this regard, the three novel EMG features are employed in this work. Those features are explained as follow:

Absolute value of the Summation of Square root (ASS) is a recent EMG feature, and it can be defined as [2]:

$$ASS = \left| \sum_{i=1}^{L} (x_i)^{1/2} \right| \tag{1}$$

where x is the EMG signal and L is the number of samples.

Mean value of the Square root (MSR) is defined as the measurement of the total amount of muscle activity, and it can be formulated as [2]:

$$MSR = \frac{1}{L} \sum_{i=1}^{L} (x_i)^{1/2} \tag{2}$$

where x is the EMG signal and L is the number of samples.

Absolute value of the Summation of the exp^{th} root of the given signal and its Mean (ASM) is one of the recent EMG features that comprises of the information related to the amplitude of the rectified signal [2]. Mathematically, ASM can be expressed as:

$$ASM = \left| \frac{\sum_{i=1}^{L} (x_i)^{\exp}}{L} \right| \qquad \exp = \begin{cases} 0.50, \text{ if } i \ge 0.25 * L \text{ and } i \le 0.75 * L \\ 0.75, \text{ otherwise} \end{cases} \tag{3}$$

where x is the EMG signal and L is the number of samples.

C. Feature Selection

Feature selection is an important step in data mining process, which is not only increasing the classification performance, but also reducing the number of features. In the present study, two recent feature selection methods namely Competitive Binary Grey Wolf Optimizer (CBGWO) and Modified Binary Tree Growth Algorithm (MBTGA) are utilized.

CBGWO is an improved version of Binary Grey Wolf Optimization (BGWO), which has been proven to work better than BGWO in evaluating the optimal EMG feature subset. Generally, CBGWO is a population based metaheuristic algorithm. The wolves (solutions) are guided by the leaders (alpha, beta and delta) to move toward optimal prey position (global optimum). In CBGWO, the population is randomly divided into N/2 couples, where N is the population size. Then, the competition is made between two wolves in each couple, and the winner is directly passed into new population. On one side, the losers are updated by learning from the winners and leaders. Furthermore, the leaders are allowed to enhance themselves using the random walk. The detail on CBGWO can be found in [6].

MBTGA is another improved version of Binary Tree Growth Algorithm (BTGA) that has been proven to be outperformed BTGA in EMG feature selection. Like CBGWO, MBTGA is also a population based optimization algorithm. Generally, MBTGA consists of four group of trees. In the first group, the trees updated themselves locally by using the swap operator. In the second group, the crossover and mutation operators are applied for the competition of light. In the third group, the worst trees are removed and the new trees are planted. In the final group, the mask operators are used to create new potential trees. Iteratively, the global best solution is stored to ensure high diversity. The detail on MBTGA can be found in [10].

D. Classification

Once the potential feature subset was selected, the chosen features are then fed into the classifier for the recognition task. In this work, the k-nearest neighbor (KNN) is used to classify the selected features for the recognition of multiple hand motions. According to literature, KNN can achieve promising performance in EMG pattern recognition. Additionally, KNN is simpler together with faster processing speed [11], [12]. Thus, it is believed that KNN can provide satisfactory performance in EMG signals classification.

III. RESULT AND DISCUSSION

A. Pre-Experimental

The EMG data of 17 hand motions of 2 amputee subjects are acquired from DB7. Three recent features are then extracted from each EMG signal. In total, 36 features (3 features × 12 channels) are extracted from each motion from each subject. Afterward, the features are normalized in the range between 0 and 1. Fig.2 illustrates the sample features extracted from channel 9 from one subject. In Fig.2, the figure above displays the extracted features for hand motion type 1 (M1) while the figure below exhibits the extracted features for hand motion type 2 (M2).

In the next step, the CBGWO and MBTGA are employed to evaluate the most informative feature subset. As for feature selection, the classification error rate computed by the KNN algorithm is used for fitness evaluation. Note that 6-folds cross-validation is applied in this work, where the data is randomly divided into 6 equal parts. Each part is used for testing in succession, while the remaining 5 parts are used for training session. The averaged result obtained from 6 folds is used for performance evaluation [9]. Table I outlines the parameter setting of feature selection methods. As can be seen, CBGWO has less parameters than MBTGA. Since the performance of metaheuristic optimization algorithm are greatly influenced by the initial solutions, thus, each feature selection method is executed for 10 runs with different random seed. The averaged result obtained form 10 runs is used for performance comparison.

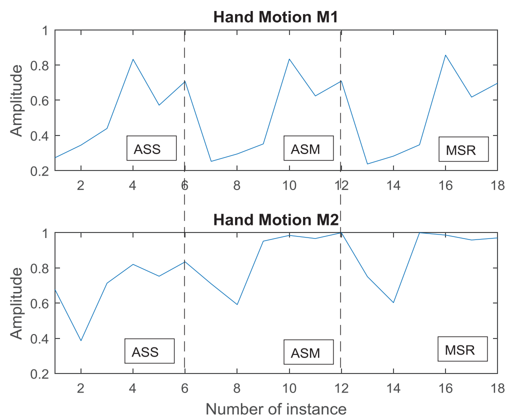

Fig. 2. Extracted features of channel 9 from one subject.

Three statistical parameters namely classification accuracy (CA), number of selected features (NF) and F-measure (FM) are used for performance measurement. Mathematically, the CA and FM can be expressed as follows [13], [14]:

$$\text{CA} = \frac{\textit{Number of corrected classified instances}}{\textit{Total number of instances}} \times 100\% \quad (4)$$

$$\text{FM} = \frac{2TP}{2TP + FP + FN} \quad (5)$$

where TP is the true positive, FP is the false positive and FN is the false negative.

TABLE I. PARAMETER SETTING OF FEATURE SELECTION METHODS

Parameter	Feature Selection Method	
	CBGWO	*MBTGA*
Population size, N	30	30
Maximum number of iteration, T_{max}	100	100
Number of trees in first group, N_1	-	10
Number of trees in second group, N_2	-	15
Number of trees in fourth group, N_4	-	5

B. Experimental Result

Fig.3 shows the results of classification accuracy of two feature selection methods. As can be observe, the worst classification performance was falling on Original. The results indicated that by applying the feature selection (CBGWO or MBTGA), the classification accuracy has been increased. This again verifies the important of feature selection in EMG signals classification. In Fig.3, CBGWO outperformed MBTGA for subject 1. By contrast, MBTGA scores better accuracy for subject 2. This result implies that no universal optimizer can solve all the feature selection problems in the world, which is according to No Free Lunch (NFL) Theorem [15].

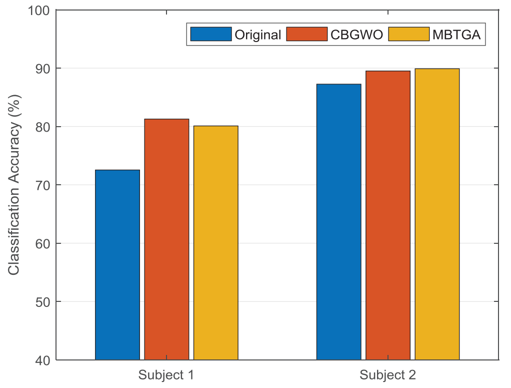

Fig. 3. Classification accuracy of two feature selection methods.

Table II displays the results of feature size. It is worth noting that Original represents the full features (36 features). From Table II, MBTGA achieved the mean feature size of 17.85. On the other hand, CBGWO offered the mean feature size of 18.65, which eliminates more than 50% of features in the process of evaluation. In comparison with MBTGA, CBGWO requires roughly 19 features in providing a high classification result, which leads to low complexity.

TABLE II. RESULTS OF FEATURE SIZE

Subject	Number of selected features (feature size)		
	Original	*CBGWO*	*MBTGA*
1	36	18.70	17.20
2	36	18.60	18.50
Mean	36	18.65	17.85

Table III demonstrates the results of F-measure. It is observed that the performance with feature selection is better than Original (full features). The experimental result again verified the efficiency of feature selection in classification task. In Table III, CBGWO provided higher mean F-measure of 0.8520, followed by MBTGA, 0.8482. Ultimately, the feature selection method that offered the best performance is found to be CBGWO.

Fig.4 presents the results of computational time for CBGWO and MBTGA. Based on the result obtained, it shows that CBGWO provided the fastest processing speed, which leads to very low computational cost. In agreement with previous study, CBWGO made use of only half population in the updating process, thus resulting in fast computational speed [6].

TABLE III. RESULTS OF F-MEASURE

Subject	F-measure		
	Original	*CBGWO*	*MBTGA*
1	0.7268	0.8092	0.7975
2	0.8721	0.8949	0.8989
Mean	0.7994	0.8520	0.8482

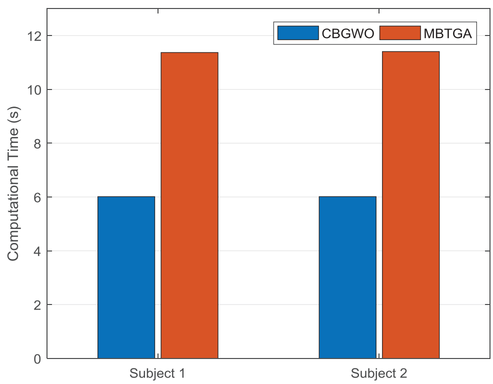

Fig. 4. Computational cost of two feature selection methods.

IV. CONCLUSION

In this paper, the effectiveness of feature selection in EMG signals classification was presented. Occasionally, feature selection not only minimizes the feature size, but also improves the classification performance. Presently, two recent feature selection methods namely CBGWO and MBTGA were employed to evaluate the optimal feature subset. Based on the results obtained, it showed that by applying CBWGO and MBTGA, the classification performance has been improved. In additional, the number of features have been reduced. Furthermore, the performances of CBGWO and MBTGA are found to be similar. However, CBGWO can eliminate more redundant and irrelevant features and it contributes a very low computational cost. In sum, CBGWO and MBTGA are suitable to be applied for rehabilitation and clinical applications.

ACKNOWLEDGMENT

This research is supported and funded by the Skim Zamalah UTeM and Minister of Higher Education Malaysia under grant FRGS/1/2017/TK04/FKE-CeRIA/F00334.

REFERENCES

[1] T. N. S. T. Zawawi, A. R. Abdullah, W. T. Jin, R. Sudirman, and N. M. Saad, “Electromyography Signal Analysis Using Time and Frequency Domain for Health Screening System Task,” Int. J. Hum. Technol. Interact. IJHaTI, vol. 2, no. 1, pp. 35-44–44, Apr. 2018.

[2] O. W. Samuel et al., “Pattern recognition of electromyography signals based on novel time domain features for amputees’ limb motion classification,” Comput. Electr. Eng., vol. 67, pp. 646–655, Apr. 2018.

[3] G. Purushothaman and R. Vikas, “Identification of a feature selection based pattern recognition scheme for finger movement recognition from multichannel EMG signals,” Australas. Phys. Eng. Sci. Med., vol. 41, no. 2, pp. 549–559, Jun. 2018.

[4] J. Too, A. R. Abdullah, T. N. S. T. Zawawi, N. M. Saad, and H. Musa, “Classification of EMG Signal Based on Time Domain and Frequency Domain Features,” Int. J. Hum. Technol. Interact. IJHaTI, vol. 1, no. 1, pp. 25–30, Oct. 2017.

[5] Z. Othman, N. A. Abdullah, K. Y. Chin, F. F. W. Shahrin, S. S. S. Ahmad, and F. Kasmin, “Comparison on Cloud Image Classification for Thrash Collecting LEGO Mindstorms EV3 Robot,” Int. J. Hum. Technol. Interact. IJHaTI, vol. 2, no. 1, pp. 29-34–34, Apr. 2018.

[6] J. Too, A. R. Abdullah, N. Mohd Saad, N. Mohd Ali, and W. Tee, “A New Competitive Binary Grey Wolf Optimizer to Solve the Feature Selection Problem in EMG Signals Classification,” Computers, vol. 7, no. 4, p. 58, Nov. 2018.

[7] A. Phinyomark, R. N. Khushaba, and E. Scheme, “Feature Extraction and Selection for Myoelectric Control Based on Wearable EMG Sensors,” Sensors, vol. 18, no. 5, p. 1615, May 2018.

[8] B. Xue, M. Zhang, and W. N. Browne, “Particle swarm optimisation for feature selection in classification: Novel initialisation and updating

mechanisms," Appl. Soft Comput., vol. 18, no. Supplement C, pp. 261–276, May 2014.

[9] A. Krasoulis, I. Kyranou, M. S. Erden, K. Nazarpour, and S. Vijayakumar, "Improved prosthetic hand control with concurrent use of myoelectric and inertial measurements," J. NeuroEngineering Rehabil., vol. 14, p. 71, Jul. 2017.

[10] J. Too, A. R. Abdullah, N. Mohd Saad, and N. Mohd Ali, "Feature Selection Based on Binary Tree Growth Algorithm for the Classification of Myoelectric Signals," Machines, vol. 6, no. 4, p. 65, Dec. 2018.

[11] H. Sun, X. Zhang, Y. Zhao, Y. Zhang, X. Zhong, and Z. Fan, "A Novel Feature Optimization for Wearable Human-Computer Interfaces Using Surface Electromyography Sensors," Sensors, vol. 18, no. 3, p. 869, Mar. 2018.

[12] W.-T. Shi, Z.-J. Lyu, S.-T. Tang, T.-L. Chia, and C.-Y. Yang, "A bionic hand controlled by hand gesture recognition based on surface EMG signals: A preliminary study," Biocybern. Biomed. Eng., vol. 38, no. 1, pp. 126–135, Jan. 2018.

[13] Y. Gu, D. Yang, Q. Huang, W. Yang, and H. Liu, "Robust EMG pattern recognition in the presence of confounding factors: features, classifiers and adaptive learning," Expert Syst. Appl., vol. 96, pp. 208–217, Apr. 2018.

[14] J. Too, A. R. Abdullah, N. M. Saad, N. M. Ali, and T. N. S. T. Zawawi, "Exploring the Relation Between EMG Pattern Recognition and Sampling Rate Using Spectrogram," J. Electr. Eng. Technol., Jan. 2019.

[15] D. H. Wolpert and W. G. Macready, "No free lunch theorems for optimization," IEEE Trans. Evol. Comput., vol. 1, no. 1, pp. 67–82, Apr. 1997.