# Adaptive Wavelet Extreme Learning Machine (AW-ELM) for Index Finger Recognition Using Two-Channel Electromyography

Khairul Anam[1,2] and Adel Al-Jumaily[2]

[1] University of Jember, Indonesia
Khairul.Anam@student.uts.edu.au
[2] University of Technology Sydney, Australia
adel@uts.edu.au

**Abstract.** This paper proposes a new structure of wavelet extreme learning machine i.e. an adaptive wavelet extreme learning machine (AW-ELM) for finger motion recognition using only two EMG channels. The adaptation mechanism is performed by adjusting the wavelet shape based on the input information. The performance of the proposed method is compared to ELM using wavelet (W-ELM0 and sigmoid (Sig-ELM) activation function. The experimental results demonstrate that the proposed AW-ELM performs better than W-ELM and Sig-ELM.

**Keywords:** Wavelet extreme learning machine, adaptive.

## 1 Introduction

A wavelet neural network (WNN) is a special case of a feed-forward neural network which its activation function is wavelets [1]. A standard gradient descend can be used to train the weight of WNN. However, drawbacks of the gradient descent method such as long training time and easy trapped to local minima have hampered the implementation of WNN in the real-time application [2]. On the other hand, an extreme learning machine (ELM) was introduced to train a single-hidden layer feed-forward networks (SLFNs) resulting in a system which is fast and able to avoid a local minima [3]. Inevitably, WNN can be constructed using SLFNs.

The combination of ELM and WNN can be conducted by simply replacing the activation function of ELM with wavelets [4] [5]. This is the simplest unification of both networks as has been done in [5]. Cao et al. [6] introduced a new combination of these two algorithms by proposing a composite function of WNN with ELM. In this method, they implemented two activation functions, a wavelet function and any piecewise function which are done in order.

Another new unification of ELM and WNN was proposed by Javed et al. [7] who proposed a summation wavelet extreme learning machine (SW-ELM). Same as Cao, Javed et al. utilized two activation functions but employed them in different ways.

C.K. Loo et al. (Eds.): ICONIP 2014, Part I, LNCS 8834, pp. 471–478, 2014.

These two activation functions were done in parallel and their outputs were averaged to be the output of the hidden nodes.

This paper proposes an adaptive wavelet extreme learning machine (AW-ELM), a new unification of ELM and WNN. According to WNN structure, the proposed system utilizes a wavelet function as the activation function in the hidden node. However, the activation functions are not fixed but they are adjusted regarding to the changing in the input. The sigmoid function is used to process the input information and produce translation parameters of the wavelets in the related hidden-node. In this paper, the performance of AW-ELM will be tested to classify the finger motions from the surface Electromyography signal (EMG) extracted from two-channel sources on the forearm. In addition, its classification performance will be compared with two types of ELM, ELM with wavelet activation function (W-ELM) and sigmoid activation function (Sig-ELM).

The organization of the paper is as follows: section 2 describes the theory of W-ELM and AW-ELM, and the implementation of AW-ELM for finger motion classification. Then section 3 and 4 presents the results and the discussion. Finally section 4 will conclude this paper.

## 2 Methods

### 2.1 Wavelet Extreme Learning Machine (W-ELM)

W-ELM can be considered as a special case of extreme learning machine which its activation function is wavelets. The output function of W-ELM for arbitrary samples $(\boldsymbol{x}_k,\boldsymbol{t}_k) \in \boldsymbol{R}^n \text{ x } \boldsymbol{R}^o$ with M hidden nodes is

$$f_i^{\,k}(\mathbf{x}) = \sum_{j=1}^{M} V_{ij}\psi_{a_j bj}(w_j, c_j, \mathbf{x}_k) \qquad i = 1,2,\ldots,O \tag{1}$$

where

$$\psi_{a_j bj}(x) = \frac{1}{\sqrt{a_j}}\psi\left(\frac{x - b_j}{a_j}\right), \qquad j = 1,2,\ldots\ldots,M \tag{2}$$

in which $a_j$ and $b_j$ are dilatation and translation parameters of the wavelets, respectively. An initialization of dilatation and translation parameters, $a_j$ and $b_j$, in WNN is an important issue. The initialization should consider the input information in order to let the time domain of the wavelet covering the input domain. According to [1], suppose the input vector $x_k$ has the domain $[x_{kmin}, x_{kmax}]$, $t^*$ and $\sigma^*$ are the centre and the radius of the mother wavelet $\psi_{a_i b_i}$, then domain of $\psi_{a_i b_i}$ is given by:

$$[b_j + a_j(t^* - \sigma^*) \ , \ b_j + a_j(t^* + \sigma^*)]$$

Meanwhile, the input information range for $i$th hidden layer can be calculated as:

$$\left[\sum_{i=1}^{N} w_{ji} x_{i\min} \quad , \quad \sum_{i=1}^{N} w_{ji} x_{i\max}\right]$$

where $w_{ji}$ is the weight connecting the $j$th hidden layer the $i$th input. The wavelet can cover the input space if

$$b_j + a_j (t^* - \sigma^*) = \sum_{i=1}^{N} w_{ji} x_{i\min} \tag{3}$$

and

$$b_j + a_j (t^* + \sigma^*) = \sum_{i=1}^{N} w_{ji} x_{i\max} \tag{4}$$

From equation (9) and (10), we can calculate $a_i$ and $b_i$ as:

$$a_j = \frac{1}{2\sigma^*}\left(\sum_{i=1}^{N} w_{ji} x_{i\max} - \sum_{i=1}^{N} w_{ji} x_{i\min}\right) \tag{5}$$

$$b_j = \frac{1}{2\sigma^*}\left(\sum_{i=1}^{N} w_{ji} x_{i\max} (\sigma^* - t^*) + \sum_{i=1}^{N} w_{ji} x_{i\min} (\sigma^* + t^*)\right) \tag{6}$$

### 2.2 Adaptive Wavelet Extreme Learning Machine (AW-ELM)

**The Proposed Structure**

The proposed AW-ELM is depicted by Fig. 1. If M is the number of hidden node and N is the number of input, then the input of the hidden layer $P_j$ is given by

$$P_j(x) = \sum_{i=1}^{N} x_i w_{ji} + c_j \qquad j = 1, 2, ..., M \tag{7}$$

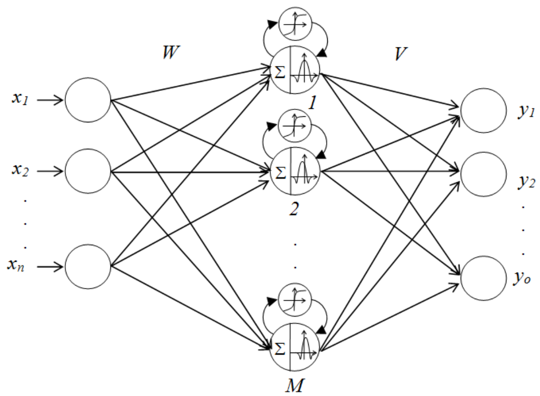



**Fig. 1.** The proposed adaptive wavelet extreme learning machine

where $x_i$ are the input variables, $w_{ji}$ are the weights of the connection between $i$th input and $j$th hidden nodes, and $c_j$ denotes the bias of $j$th hidden layer. Using equation (8), the output of the hidden node is given by:

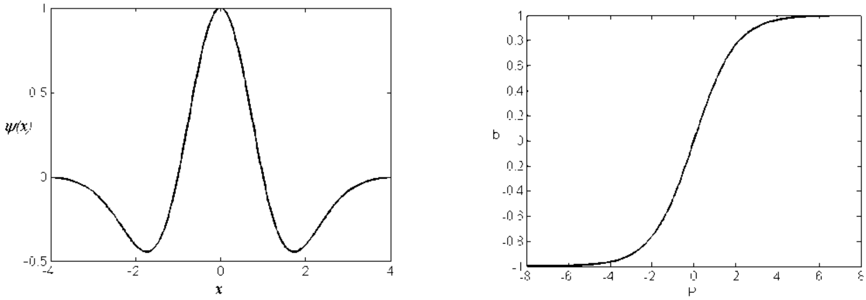
$$\psi_{a_j b_j}(P_j(x)) = \psi\left(\frac{P_j(x) - b_j}{a_j}\right), \qquad j = 1, 2, ..., M \tag{8}$$

In this proposed work, the Mexican Hat function [6] is used as the mother wavelet $\psi_{a_i b_i}$ as described in fig. 2a, and defined as

$$\psi(x) = e^{-x^2/2}(1 - x^2) \tag{9}$$

Therefore, the wavelet activation function of AW-ELM is:

$$\psi_{a_j b_j}(P_j) = e^{-0.5\left(\frac{P_j - b_j}{a_j}\right)^2}\left(1 - \left(\frac{P_j - b_j}{a_j}\right)^2\right) \tag{10}$$




( a) (b)

**Fig. 2.** Two difference functions used in this work: (a) The mother wavelet of the Mexican hat (b) A nonlinear function to produce $b_j$

In this proposed AW-ELM, the dilatation parameters $a_j$ are fixed and initialized using Equation (5). As for the translation parameters $b_j$, they are varied according to the input information and driven by a nonlinear function f(.) as follows:

$$b_j = f(P_j) \tag{11}$$

where

$$f(P_j) = \frac{2}{1 + e^{-P_j}} - 1 \tag{12}$$

as depicted in Fig. 2b. Eventually, a new structure of an adaptive W-ELM is presented in Fig. 1. A small circle on the top of each hidden node is used to adjust the b parameters in order to change the shape of the wavelet. Thus, the output of AW-ELM is:

$$f_i^k(\mathbf{x}) = \sum_{j=1}^{M} V_{ij}\psi_{a_j b_j}(w_j, c_j, \mathbf{x}_k) = \sum_{j=1}^{M} V_{ij}\psi_{a_j b_j}(P_j(\mathbf{x}_k)) \quad i = 1,2,\dots,O \tag{13}$$

**The Learning Algorithm**
For the desired output:

$$\mathrm{D} = (\boldsymbol{d}_1^T \quad \boldsymbol{d}_2^T \quad \dots \quad \boldsymbol{d}_L^T)_{LxO} \tag{14}$$

The AW-ELM described in (13) can be written as a linear system as follows:

$$\mathrm{HV} = \mathrm{D} \tag{15}$$

where

$$\mathrm{H} = \begin{bmatrix} \psi_{a_1 b_1}(P_1(\mathbf{x}_1)) & \cdots & \psi_{a_M b_M}(P_M(\mathbf{x}_1)) \\ \vdots & \vdots & \vdots \\ \psi_{a_1 b_1}(P_1(\mathbf{x}_L)) & \vdots & \psi_{a_M b_M}(P_M(\mathbf{x}_L)) \end{bmatrix}_{LxM} \tag{16}$$

$$\mathrm{V} = (\boldsymbol{v}_1^T \quad \boldsymbol{v}_2^T \quad \cdots \quad \boldsymbol{v}_O^T)_{MxO}^T \tag{17}$$

V can be obtained by solving the least-square solution of (15) and given by:

$$\hat{\mathrm{V}} = \mathrm{H}^{\dagger}\mathrm{D} \tag{18}$$

where $\mathrm{H}^{\dagger}$ is the Moore-Penrose generalized inverse of the matrix H.

The training algorithm of AW-ELM can be implemented as follows:

**Algorithm of AW-ELM**. Given a training set $\aleph = \{(\boldsymbol{x}_k, \boldsymbol{t}_k) \mid \boldsymbol{x}_k \in \boldsymbol{R}^n, \boldsymbol{t}_k \in \boldsymbol{R}^o, k = 1,2\dots, L\}$, the hidden node output function $\psi_{ai,bi}(\boldsymbol{w}, \boldsymbol{c}\ ; \boldsymbol{x})$ and the hidden node number $M$:

(1) Randomly assign vector matrix W and initialize the hidden node parameters $a_j$, j = 1,2,.., M according to (5)
(2) Calculate the input hidden layer $P_j$ in (7)
(3) Calculate $\underline{b_j}$ in (11) and (12)
(4) Calculate the hidden layer output H in (16)
(5) Calculate the output weight $\hat{\mathrm{V}}$ in (18)

### 2.3 AW-ELM for Finger Motion Recognition

The proposed recognition system consists of the same stages as depicted in Figure 3. Firstly, signals from two-channel EMG located on the forearm were acquired by a data acquisition device from eight subjects. The experimental procedures for the data acquisition could be referred to [8]. Then the filtering and windowing was applied to the collected data before being extracted using a time domain (TD) and autoregressive (AR) features.

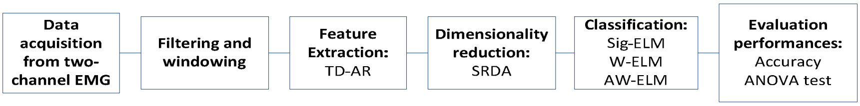



**Fig. 2.** The motion finger classification using AW-ELM

The features were extracted from the time domain feature set which consists of Waveform Length (WL), Slope Sign Changes (SSC), Number of Zero Crossings (ZCC), and Sample Skewness (SS). In addition, some parameters from Hjorth Time Domain Parameters (HTD) and Auto Regressive (AR) Model Parameters were included. To reduce the dimension of the features, SDRA was employed. All features were concatenated and reduced using SRDA. SRDA is an extension of LDA that can deal with singularity and a large data set. The 200 ms window length was applied to the signal to comply with the real time application along with a 25 increment.

The reduced feature set resulted in the previous stage is utilized in the classification. The objective of the classification that was performed using AW-ELM and other ELM classifiers is to recognize ten classes of the individual and combined finger movements consisting of the flexion of individuated fingers. They consisted of Thumb (T), Index (I), Middle (M), Ring (R), Little (L) and the pinching of combined Thumb–Index (T–I), Thumb–Middle (T–M), Thumb–Ring (T–R), Thumb–Little (T–L), and the hand close (HC). Finally, statistical analyses were performed to validate the result.

## 3 Results and Discussion

In this section, the performance of the proposed AW-ELM was compared to wavelet extreme learning machine (W-ELM) and sigmoid extreme learning machine (sig-ELM). All classifiers classified ten finger motions using EMG signal from two channel electrodes. The four-fold cross validation was used to validate the classification results. Simulation was done in the MATLAB 8.3 environment running on 2.8 GHz PC.

Table 1 shows the classification results of three classifiers in recognizing ten finger motions classes defined in 2.4. In all ELMs, the number of hidden nodes varied from 50 up to 500. The results indicate that the average accuracy the proposed of AW-ELM was higher than standard W-ELM in all cases. Likewise, the AW-ELM performance is better than Sig-ELM in all hidden node numbers except 50 and 75. In these two hidden numbers, the Sig-Elm achieved better accuracy that AW-ELM. Overall, the adaptation of wavelet shape using a sigmoid function in AW-ELM could enhance the performance of the original wavelet extreme learning machine and in several condition, could attain better performance than Sig-ELM.

**Table 1.** The average classification accuracy of AW-ELM across eight subjects using four-fold cross validation compared with W-ELM and Sig-ELM

| # Hidden Node | Accuracy (%) | | |
|---|---|---|---|
| | W-ELM | AW-ELM | Sig-ELM |
| 50 | 91.07 ± 0.17 | 91.57 ± 0.08 | **91.65 ± 0.08** |
| 75 | 91.56 ± 0.14 | 91.93 ± 0.14 | **91.97 ± 0.10** |
| 100 | 91.79 ± 0.10 | **92.05 ± 0.09** | 92.01 ± 0.08 |
| 125 | 91.90 ± 0.08 | **92.08 ± 0.09** | 92.03 ± 0.10 |
| 150 | 91.94 ± 0.11 | **92.06 ± 0.10** | 92.04 ± 0.10 |
| 175 | 91.98 ± 0.09 | **92.06 ± 0.09** | 92.04 ± 0.08 |
| 200 | 91.99 ± 0.08 | **92.04 ± 0.08** | 92.01 ± 0.06 |
| 500 | 91.79 ± 0.08 | **91.56 ± 0.06** | 91.37 ± 0.06 |

**Table 2.** Processing time of different ELM classifiers

| #Hidden | Training Time (s) | | | Testing Time (s) | | |
|---|---|---|---|---|---|---|
| Node | W-ELM | AW-ELM | Sig-ELM | W-ELM | AW-ELM | Sig-ELM |
| 50 | 0.16 ± 0.01 | 0.19 ± 0.02 | 0.12 ± 0.00 | 0.03 ± 0.00 | 0.06 ± 0.00 | 0.01 ± 0.00 |
| 75 | 0.28 ± 0.01 | 0.33 ± 0.01 | 0.19 ± 0.00 | 0.06 ± 0.01 | 0.07 ± 0.00 | 0.02 ± 0.00 |
| 100 | 0.38 ± 0.02 | 0.45 ± 0.02 | 0.27 ± 0.01 | 0.07 ± 0.00 | 0.10 ± 0.00 | 0.02 ± 0.00 |
| 125 | 0.59 ± 0.05 | 0.70 ± 0.06 | 0.48 ± 0.07 | 0.09 ± 0.00 | 0.14 ± 0.00 | 0.03 ± 0.00 |
| 150 | 0.71 ± 0.01 | 0.81 ± 0.01 | 0.51 ± 0.01 | 0.12 ± 0.00 | 0.19 ± 0.01 | 0.04 ± 0.00 |
| 175 | 0.93 ± 0.06 | 1.06 ± 0.05 | 0.69 ± 0.05 | 0.14 ± 0.00 | 0.22 ± 0.00 | 0.04 ± 0.00 |
| 200 | 1.07 ± 0.08 | 1.22 ± 0.05 | 0.85 ± 0.06 | 0.17 ± 0.00 | 0.26 ± 0.00 | 0.05 ± 0.00 |
| 500 | 4.20 ± 0.08 | 5.42 ± 0.08 | 2.82 ± 0.10 | 0.75 ± 0.01 | 1.28 ± 0.01 | 0.12 ± 0.00 |

In terms of processing time, the ELM using a sigmoid function (Sig-ELM) spent less training time than W-ELM and AW-ELM in as shown in Table 2. Table 2 shows that the larger the number of the hidden node, the longer the time difference between AW-ELM and Sig-ELM. Likewise, in the testing time, AW-ELM is the slowest system. The adaptive mechanism adds the processing time in both the training and testing trials.

**Table 3.** The p-value of anova test on the classification accuracy between AW-ELM and other tested classifiers

| #Hidden | *p*-value | |
|---|---|---|
| Node | AW-ELM & W-ELM | AW-ELM & Sig-ELM |
| 50 | **0.0000** | **0.0000** |
| 75 | **0.0000** | 0.1283 |
| 100 | **0.0000** | **0.0006** |
| 125 | **0.0000** | 0.0610 |
| 150 | **0.0000** | 0.3477 |
| 175 | **0.0021** | 0.5746 |
| 200 | **0.0098** | 0.0552 |
| 500 | **0.0000** | **0.0000** |

The one-way ANOVA test was done to evaluate the improvement significance of AW-ELM compared to W-ELM and Sig-ELM as presented in Table 4. Table 4 shows that p-values on the comparison of AW-ELM and W-ELM are less than 0.05. In other words, the performance improvement in recognizing ten finger motions by AW-ELM is significantly achieved. Furthermore, the performance of AW-ELM and Sig-ELM in some cases is significantly different in the hidden number node 50, 100 and 500 whereas it is significantly similar in other hidden node numbers. Nevertheless, the AW-ELM produced better accuracy in most trials than Sig-ELM.

## 4 Conclusion

This paper proposed a novel ELM i.e. an adaptive wavelet extreme learning (AW-ELM) for recognizing finger motions using two-channel EMG signals. The adaptation mechanism of the proposed method is conducted by adjusting the shape of the wavelet based on the information provided in the input. The experimental results showed that the proposed AW-ELM improved the performance of the original wavelet ELM in all cases tested and performed better than Sig-ELM in most cases observed. In the future, the performance of AW-ELM should be compared with other well-known classifiers such as support vector machine (SVM) and Linear Discriminant Analysis (LDA).

## References

1. Zhou, B., Shi, A., Cai, F., Zhang, Y.-S.: Wavelet neural networks for nonlinear time series analysis. In: Yin, F.-L., Wang, J., Guo, C. (eds.) ISNN 2004. LNCS, vol. 3174, pp. 430–435. Springer, Heidelberg (2004)
2. Lin, C.-J., Tsai, H.-M.: FPGA implementation of a wavelet neural network with particle swarm optimization learning. Mathematical and Computer Modelling 47(9-10), 982–996 (2008)
3. Huang, G.B., et al.: Extreme learning machine for regression and multiclass classification. IEEE Trans. Syst. Man Cybern. B, Cybern. 42(2), 513–529 (2012)
4. Cao, J., Lin, Z., Huang, G.-B.: Composite function wavelet neural networks with extreme learning machine. Neurocomputing 73(7), 1405–1416 (2010)
5. Salih, D.M., et al.: Wavelet network based online sequential extreme learning machine for dynamic system modeling. In: 2013 9th Asian Control Conference (ASCC) (2013)
6. Ling, S.-H., et al.: Improved hybrid particle swarm optimized wavelet neural network for modeling the development of fluid dispensing for electronic packaging. IEEE Transactions on Industrial Electronics 55(9), 3447–3460 (2008)
7. Javed, K., Gouriveau, R., Zerhouni, N.: SW-ELM: A summation wavelet extreme learning machine algorithm with a priori parameter initialization. Neurocomputing 123, 299–307 (2014)
8. Anam, K., Al-Jumaily, A.A.: Swarm-based extreme learning machine for finger movement recognition. In: 2014 Middle East Conference on Biomedical Engineering (MECBME) (2014)