AN HYBRID PARTICLE SWARM OPTIMIZATION WITH MULTILAYER PERCEPTRON FOR EEG CLASSIFICATION

Dr.G.V.Sridhar
Professor, Department of ECE,
Raghu Engg College.,Visakhapatnam, A.P.
Email: Sridhar1209@gmail.com

Dr.M.Murali
Professor, Department of ECE,
Raghu Engg College.,Visakhapatnam, A.P.

Dr.A.Vamsidhar
Asoc Professor, Department of ECE,
Raghu Engg College.,Visakhapatnam, A.P.

Abstract— Brain Computer Interface (BCI) research includes recording and analyzing Electro Encephalo Graphic (EEG) data and recognizing EEG patterns in different mental states. Multilayer Perceptron Neural Networks (MLPNN) ability to model non-linear relationship can model EEG signals complex nature. MLPNN's convergence is relatively slow, often yielding sub-optimal solutions. In this paper, it is proposed to enhance the classification ability of the MLPNN incorporating hybrid Particle Swarm Optimization (PSO). This optimization helps to overcome the slow convergence rate and to avoid local minimum. The features of motor imagery in the frequency domain are extracted using Hilbert transform. To reduce the dimensionality of the feature set, it is proposed to use a hybrid PSO – Genetic Algorithm (PSO-GA) for feature selection. The selected features are classified by MLPNN modified with hybrid Particle Swarm Optimization – Genetic Algorithm (PSO-GA). Keywords— Brain Computer Interface (BCI), EEG, Multilayer Perceptron Neural Networks (MLPNN), Particle Swarm Optimization (PSO), Genetic Algorithm (GA), IV A Dataset.

1. INTRODUCTION

Brain Computer Interface (BCI) enables use of brain's neural activity to communicate or control machines, artificial limbs or robots without physical movements [1-3]. BCI applications strength is how neural patterns are extracted from EEG and input into machine commands. EEG is recording the brain's electrical activity. There are two types of EEG based on the location of electrodes on the head: scalp and intracranial. For scalp EEG, electrodes are placed on a scalp with mechanical/electrical contacts. But, intracranial EEG is got through special electrodes implanted during brain surgery. Scalp EEG, is focused on this research and is a common diagnostic method to detect brain's electrical activity abnormalities. "Electroencephalography" (EEG) measures brain's neural activity as electrical voltage fluctuations along a scalp due to current flows in brain neurons [4]. In an EEG test, electrodes are fixed to a scalp monitor/record a brain's electrical activity [5]. EEG recordings enable understanding of epilepsy. Detection of seizures seen in EEGs is an important component in epilepsy diagnosis and treatment [6].

EEG is recorded from electrodes arranged in a specific pattern or montage. The common standard is the International 10/20 System. These are cheap methods which provide a continuous record of brain activity with better than millisecond resolution. This ensures high temporal resolution and so detailed discoveries of dynamic cognitive processes were reported using EEG and ERP (Event Related Potentials) methods. Most EEG waves range from 0.5-500Hz, but the following 4 frequency bands are clinically relevant: (i) delta, (ii) theta, (iii) alpha and (iv) beta [7].

MLPNNs are commonly used feed forward neural networks (NN) due to fast operation, implementation ease and reduced training set requirements. MLPNN have three sequential layers: input, hidden and output layers. Hidden layer processes and transmits input information to output layer. A MLPNN model with insufficient/excessive neurons in hidden layer mostly causes problems of bad generalization and over fitting. There is no analytical method to determine neurons in a hidden layer. Hence, it is only through trial and error [8]. This work proposes a hybrid PSO and GA where it is used for momentum and learning rates parameters, section 2 explains the literature survey about the work, section 3 explains methodology, section 4 explains experimental results and discussion with section 5 concluding the work.

2. RELATED WORKS

The feasibility of Bens Spike Algorithm (BSA) to encode continuous EEG spatio-temporal data into input spike streams for a classification in a spiking NN classifier was investigated by Nuntalid et al [9]. A new Evolving Probabilistic Spiking Neural Network Reservoir (epSNNr) architecture learns and classifies EEG signals after BSA transformation. A new Adaptive Neural Network Classifier (ANNC) of EEG-P300 signals from mental activities was proposed by Turnip & Hong [10]. To overcome classifier overtraining due to noisy and non-stationary data, EEG signals are filtered and their Auto Regressive (AR) properties extracted with an AR model prior to being passed to ANNC. With/without AR property extraction, the new ANNC achieved 100% accuracy for all subjects. Key EEG data features are considered while transforming through linear projection techniques by Sarkar and Ganguly [11]. Linear discriminant functions then classify EEG data in two broad classes: right hand imagery movement and left hand imagery movement, but classification accuracy emerges lower than anticipated 51.9%. Probabilistic NN is implemented next due to its quick training and lowered complexity, yielding 73.076% accuracy for two patterns classification. Classification of a 3-class mental task-based BCI using Hilbert-Huang transform (HHT) for feature extractor and Fuzzy Particle Swarm Optimization with Cross Mutated-based Artificial Neural Network (FPSOCM-ANN) for classifier was presented by Chai et al [12]. Results revealed a dominant alpha wave during eyes closure with above 90% average classification accuracy. Accuracies for tetraplegia patients were lower compared to able-bodied subjects; but, this improved by increasing time-windows duration. FPSOCM-ANN ensured improved accuracies compared to GA based Artificial Neural Network (GA-ANN) for 3 mental tasks-based BCI classification with best classification accuracy being for a 7s time-window: 84.4% (FPSOCM-ANN) compared to 77.4% (GA-ANN).

A new approach for P300-BCI classification based on a convolutional NN presented by Cecotti and Graser [13] had accuracy equal to best current method on third BCI competition's Data set II. It outperforms the best method in two situations: first, when electrodes were restricted to 8; second, when number of epochs considered was 10. An ANN-based method, Genetic Neural Mathematic Method (GNMM), was applied to two EEG channel selection/classification problems by Yang et al [14]. This had three successive steps: (1) a GA-based input selection process; (2) MLP-based modelling; and (3) successful training dependent rule extraction. GNMM performs effective channel selections/reductions, which reduces data collection difficulty and improves classifier generalization.

An improved method based on single trial EEG data for online classification of motor imagery tasks for BCI applications was presented by Ko and Sim [15]. This research aims to develop a new classification method to control an interactive robot agent platform through a BCI system. A recognition system for Motor Imagery (MI) EEG data single-trial analysis was proposed by Hsu [16]. Quantum-behaved Particle Swarm Optimization (QPSO) selects features from feature combination. Finally, such selected sub-features are classified by Support Vector Machine (SVM). Compared to Genetic Algorithm (GA), feature classification with Fisher's Linear Discriminant (FLD) on MI data from two data sets for eight subjects, reveals that the new method to be promising in BCI applications. Hassani & Lee [17] utilized Incremental Quantum PSO (IQPSO) algorithm for incremental classification of EEG data stream. IQPSO builds the classification model as a set of explicit rules which benefits from semantic symbolic knowledge representation and enhanced comprehensibility. Harmony Search (HS)-based BP NN is used for epileptic EEG signals classification as espoused by Gao et al [18]. It is known that gradient descent-based learning method leads to local optima in BP NN training which affects their approximation performances. Two HS methods, an original version and a new variation proposed by authors of the present paper, are applied to optimize weights in BP NN for epileptic EEG signals classification. Simulations proved that BP NN classification accuracy is improved by HS method-based training.

3. METHODOLOGY

3.1 Data Set

To investigate the new method publicly available dataset in [19] was used. The IV A dataset used in the BCI competition provided by Intelligent Data Analysis Group is the dataset for experimentation. EEG recordings from five healthy subjects sitting on a chair with their arms resting on armrests are the dataset. Visual cues for 3.5 s were shown for subject to perform 3 motor imageries: (L) left hand, (R) right hand, (F) right foot. The dataset has continuous signals from 118 EEG channels and markers showing time points of 280 cues for each of 5 subjects (aa, al, av, aw, ay). This study used subject aa.

3.2 Feature Extraction - Hilbert Transform

Hilbert transforms are involved in signal processing. Bandpass sampling, Analytic signal, minimum phase networks and spectral analysis are based on Hilbert transform relationships.

The Hilbert transform [20] of a function $x(t)$ is given by equation (1):

$$h(t) = H\{x(t)\} = \frac{1}{\pi}\int_{-\infty}^{+\infty}\frac{x(\tau)}{t-\tau}d(\tau) \quad (1)$$

Using the Fourier identities, the Fourier transform of the Hilbert transform of $x(t)$ is in equation (2):

$$h(t) \Leftrightarrow H(f) = -j\,\mathrm{sgn}(f)X(f) \quad (2)$$

Where $x(t) \Leftrightarrow X(f)$ is a Fourier transform pair and

$$\mathrm{sgn}(f) = \begin{cases} 1 & f > 0 \\ 0 & f = 0 \\ -1 & f < 0 \end{cases}$$

Frequencies got with Hilbert transform have artifacts removed with a band pass Chebyshev filter [21] so that all frequencies below 5 Hz and above 20Hz are removed. Chebyshev response achieves quicker roll-off by allowing ripple in frequency response. Usually, a ripple depth between 0.1 dB and 3 dB is selected. When ripple is set at 0% it is known as maximally flat or Butterworth filter.

The dB ripple for a Chebyshev filter is the peak-to-peak pass band ripple. The parameter ε is determined as in equation (3):

$$dBripple = 10\log\left(1+\varepsilon^2\right) \quad (3)$$

This can be solved for ε to obtain as in equation (4):

$$\varepsilon = \sqrt{10^{dB/10} - 1} \quad (4)$$

Parameter h is required for obtaining Chebyshev transfer functions, and is determined as in equation (5):

$$h = \tanh\left(\frac{1}{n}\sinh^{-1}\frac{1}{\varepsilon}\right) \quad (5)$$

where n is the order of low-pass filter.

3.3 Feature Selection

An EEG feature selection technique developed for classification selects features with maximum mutual information with specified classes of interest (two classes here). The simplest way is considering all feature subsets (M out of N). But, with small features number, this is computationally impossible and impractical. In a classification context, feature selection techniques are organized into 3 categories, based on how they combine feature selection search with classification model construction: filter methods, wrapper methods and embedded methods. In this study, it is proposed to use a hybrid PSO – Genetic Algorithm (PSO-GA) for feature selection.

3.4 Multilayer Perceptron (MLP)

Multilayer perceptron (MLP) networks have an input layer (*Xi*), one/more intermediary or Hidden Layers (*HL*) and output layer (*Y*). Weight matrix (*W*) is defined for every layer. ANN topology solves classification problems with non-linearly separable patterns and is used as universal function generator. MLP has 2 phases: training and execution. Backpropagation with variants is a MLP network training algorithm which is more complex than that of a perceptron network and is of the supervised variety [22].

The MLP learning algorithm [23]:

1. Initializes network with weights set to random numbers between -1 and +1.
2. Presents first training patterns/obtain outputs.
3. Compares network output with target output.
4. Propagates error backwards.

(a) Correct output weights layer with the formula in equation (6).

$$w_{ho} = w_{ho} + \left(\eta \delta_o o_h\right) \quad (6)$$

where w_{ho} is the weight connecting hidden unit h with output unit o, η is the learning rate, o_h is the output at hidden unit h. δ_o is given as in equation (7):

$$\delta_o = o_o\left(1 - o_o\right)\left(t_o - o_o\right) \quad (7)$$

Where o_o is output at node o of output layer, and t_o is target output for that node.

(b) Correct input weights using formula in equation (8).

$$w_{ih} = w_{ih} + \left(\eta \delta_h o_i\right) \quad (8)$$

where w_{ih} is weight connecting node i of input layer with node h of hidden layer, o_i is input at node i of input layer, δ_h is calculated as in equation (9).

$$\delta_h = o_h\left(1-o_h\right)\sum_o\left(\delta_o w_{ho}\right) \tag{9}$$

5. Calculate error, by taking average difference between target and output vector. The equation (10) function can be used as an example.

$$E = \frac{\sqrt{\sum_{n=1}^{p}\left(t_o - o_o\right)^2}}{p} \tag{10}$$

Where p is number of units in output layer.

6. Repeat from 2 for each pattern in training set to complete one epoch.
7. Shuffle training set randomly, to prevent network being influenced by the data order.
8. Repeat from step 2 for a specific number of epochs, or till the error ceases to change.

Transfer function is selected after which parameters are adjusted by a learning rule to ensure that neuron input/output relationship has a specific goal. This study uses sigmoid and tanh transfer function.

3.5 Proposed Hybrid Particle Swarm Optimization

In this work, the proposed Hybrid PSO used for two concepts; first for feature selection, where initial population is taken from the features extracted and second for optimizing MLPNN weights during training where the initial populations consists of weights of neurons.

Similarity between PSO and GA is system initialization with a random solutions population. Instead of using genetic operators, evolution of population generations of such individuals in a system is through cooperation and competition among individuals. Also, a randomized velocity is assigned to every potential solution/particle that it is flows through hyperspace. While stochastic factors ensure a thorough search of inter region space that are relatively good, modifications momentum effect of existing velocities lead to potential regions exploration of problem domain. This way, adjustment by PSO is similar to crossover operation in GA while stochastic processes are closer to evolutionary programming.

PSO is a population based stochastic optimization technique, simulating bird, bees and fish schooling social behavior. By randomly initializing algorithm with candidate solutions, PSO leads to a global optimum achieved by iterative procedure based on movement processes and intelligence in an evolutionary system. In PSO, every potential solution represents a particle. Two properties (position x and velocity v) are known to each particle. If x and v of ith particle are given as in equation (11):

$$\begin{aligned} x &= (x_{i1}, x_{i2}, ..., x_{iN}) \\ v &= (v_{i1}, v_{i2}, ..., v_{iN}) \end{aligned} \tag{11}$$

where N stands for problem dimensions. In iterations, a fitness function is evaluated for all swarm particles. Each particle's velocity is updated by tracking two best positions. One is best position a particle traverses till then called "*pBest*". The other is best position that a particle neighbor traverses till then. It is neighborhood best called "*nBest*". When a particle takes entire population as its neighborhood, neighborhood best becomes global best called "*gBest*".

Genetic Algorithm (GA) is a parallel mathematical algorithm transforming a set of individual mathematical objects with regard to time with operations patterned to the Darwinian principle of reproduction and survival of fittest after naturally arising from a genetic operations series which highlights sexual recombination. Each mathematical object is a string of characters (letters or numbers) of fixed length fitting the model of chains of chromosomes and associated with a specific mathematical functions reflecting their ability. GA has an operator determining search capability and algorithm convergence. A genetic operator embraces selection, crossover and mutation on a population, generating a new population.

A hybrid PSO based on GA is proposed to solve an electromagnetic optimization problem. The method includes strong co-operation between GA and PSO, as it maintains integration of two techniques for entire run. At iterations, the population is divided into two parts evolved with two techniques. They then recombine in updated population, again randomly divided into two parts in next iteration for another run of genetic or PSO. Population update concept is understood thinking that a part of individuals are substituted by the newly generated through a GA, while remaining are similar to the previous generation but moved on solution space by PSO.

This paper proposes a hybrid genetic algorithm–PSO method. The hybrid approach executes two systems simultaneously selecting P individuals from every system to exchange after designated N iterations. Individual with larger fitness has more chances of being selected.

In MLP training by PSO, representation of connection weight matrix is initial population. As the algorithm iterates, PSO operator and GA operator are applied on initial population and better population shows fitness function. Fitness of initial population for PSO and GA is expressed in equation (12):

$$fitness\ \mathrm{f}(\mathrm{FS}_i / w_i) = \frac{1}{S}\sum_{k=1}^{s}\left[\sum_{l=1}^{o}\{t_{kl} - p_{kl}(\mathrm{FS}_i / w_i)\}^2\right] \tag{12}$$

Where f is the fitness value, t_{kl} is the target output; p_{kl} is the predicted output based on FS_i / w_i; S is the number of training set samples; and, O is the number of output neurons.

The steps in a hybrid approach are:

1. Initializing GA and PSO subsystems.
2. Executing GA and PSO simultaneously.
3. Memorizing best solution as final solution and stopping if best individual in one of two subsystems satisfies termination criterion.
4. Performing hybrid process when generations are divided by designated number of N iterations. Selecting P individuals randomly from sub-systems according to fitness and exchanging. Go to step 3.

4. RESULTS AND DISCUSSION

The features of motor imagery in the frequency domain are extracted using Hilbert transform, and proposed PSO and Hybrid PSO is used for feature selection. The selected features are classified by MLPNN modified with PSO and hybrid PSO. The classification accuracy obtained from the proposed hybrid method and MLP as the classifier is shown in Figure 1.

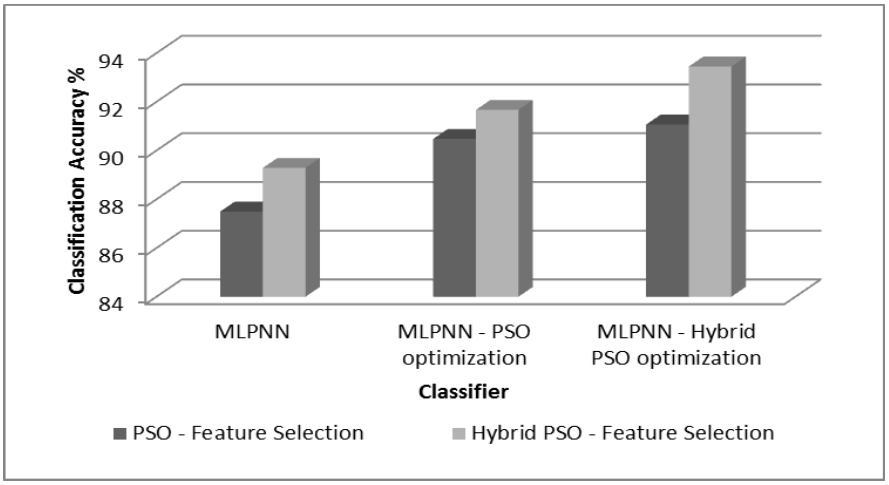

Fig 1: Classification accuracy

It is observed that the proposed hybrid PSO feature selection improves the classifier efficacy by 1.3% to 2.58%. The proposed optimization of the MLPNN using hybrid PSO achieves 4% higher classification accuracy than MLPNN and 0.65% than MLPNN with PSO for PSO feature selection method. Similarly, MLPNN using hybrid PSO achieves 4.55% higher classification accuracy than MLPNN and 1.92% than MLPNN with PSO for hybrid PSO feature selection method. Figure 2 and 3 shows the precision and recall respectively.

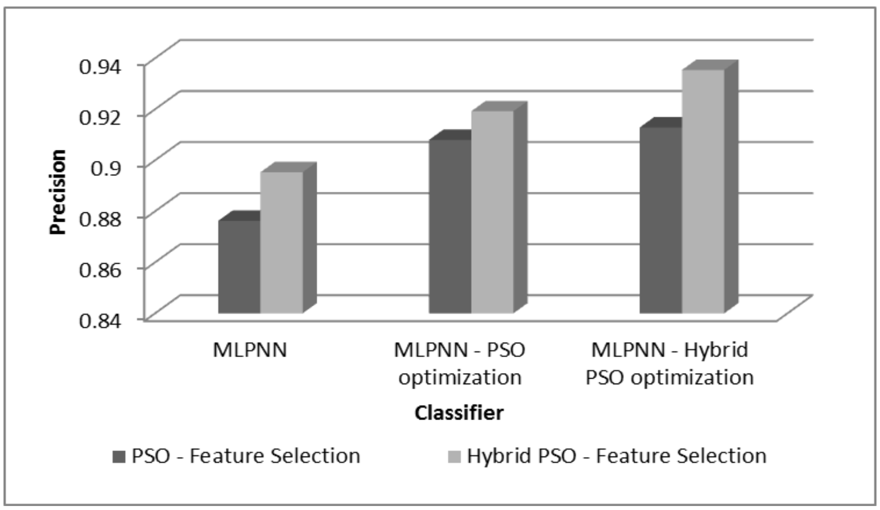

Fig 2: Precision

The precision is higher for the proposed optimization as seen from figure 2. The proposed optimization of the MLPNN using hybrid PSO achieves 4.09% higher precision than MLPNN and 0.54% than MLPNN with PSO for PSO feature selection method. Similarly, MLPNN using hybrid PSO achieves 4.38% higher precision than MLPNN and 1.74% than MLPNN with PSO for hybrid PSO feature selection method.

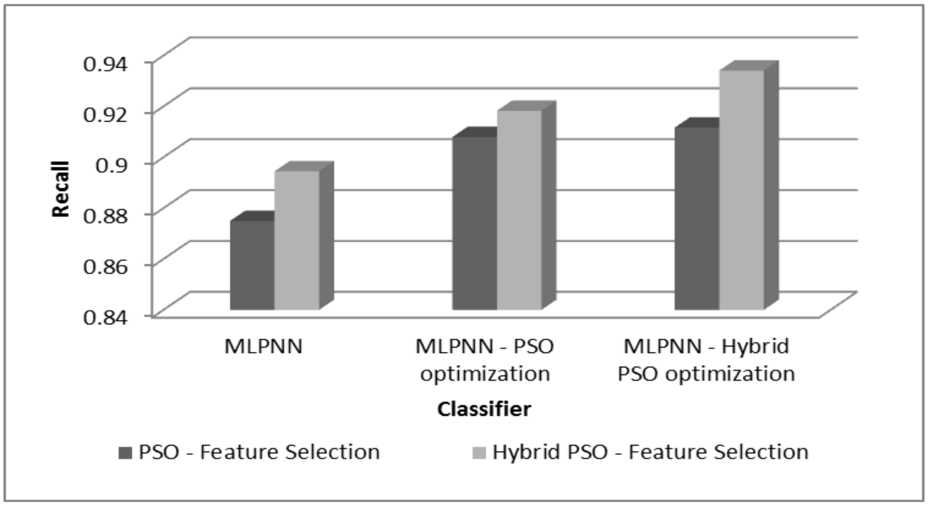

Fig 3: Recall

It is observed that the proposed hybrid PSO feature selection improves the recall by 1.13% to 2.43%. The proposed optimization of the MLPNN using hybrid PSO achieves 4.11% higher recall than MLPNN and 0.41% than MLPNN with PSO for PSO feature selection method. Similarly, MLPNN using hybrid PSO achieves 4.33% higher recall than MLPNN and 1.71% than MLPNN with PSO for hybrid PSO feature selection method.

5. CONCLUSION

In this study, features are extracted by using Hilbert Transform, for feature selection MLPNN is used. These get optimized by hybrid PSO-GA. IV A Dataset was used for experiments. Experimental results show that the proposed hybrid PSO feature selection improves the classifier efficacy by 1.3% to 2.58%. The proposed optimization of the MLPNN using hybrid PSO achieves 4% higher classification accuracy than MLPNN and 0.65% than MLPNN with PSO for PSO feature selection method. Similarly, MLPNN using hybrid PSO achieves 4.55% higher classification accuracy than MLPNN and 1.92% than MLPNN with PSO for hybrid PSO feature selection method. Similarly, proposed method obtains better precision and recall.

REFERENCES

[1] Donoghue, J. P. (2002). Connecting cortex to machines: recent advances in brain interfaces. *nature neuroscience*, *5*, 1085-1088.
[2] Vallabhaneni, A., Wang, T., & He, B. (2005). Brain—computer interface. In*Neural Engineering* (pp. 85-121). Springer US.
[3] Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain–computer interfaces for communication and control. *Clinical neurophysiology*, *113*(6), 767-791.
[4] Niedermeyer, E., & da Silva, F. L. (Eds.). (2005). *Electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins.
[5] Sleight, J., Pillai, P., & Mohan, S. (2009). Classification of executed and imagined motor movement EEG signals. *Ann Arbor: University of Michigan*, 1-10.
[6] Subasi, A. (2005). Automatic recognition of alertness level from EEG by using neural network and wavelet coefficients. *Expert Systems with Applications*, *28*(4), 701-711.
[7] Sanei, S., & Chambers, J. A. (2008). *EEG signal processing*. John Wiley & Sons.
[8] Aliabadi, R., Keynia, F., & Abdali, M. (2013). Epilepsy Seizure Diagnosis in EEG by Artificial Neural Networks. *Majlesi Journal of Multimedia Processing*, *2*(2).
[9] Nuntalid, N., Dhoble, K., & Kasabov, N. (2011, January). EEG classification with BSA spike encoding algorithm and evolving probabilistic spiking neural network. In Neural Information Processing (pp. 451-460). Springer Berlin Heidelberg.
[10] Turnip, A., & Hong, K. S. (2012). Classifying mental activities from EEG-P300 signals using adaptive neural network. Int. J. Innov. Comp. Inf. Control, 8(9), 6429-6443.
[11] Turnip, A., & Hong, K. S. (2012). Classifying mental activities from EEG-P300 signals using adaptive neural network. Int. J. Innov. Comp. Inf. Control, 8(9), 6429-6443.
[12] Chai, R., Ling, S., Hunter, G., Tran, Y., & Nguyen, H. (2013). Brain Computer Interface Classifier for Wheelchair Commands using Neural Network with Fuzzy Particle Swarm Optimization.
[13] Cecotti, H., & Graser, A. (2011). Convolutional neural networks for P300 detection with application to brain-computer interfaces. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 33(3), 433-445.

[14] Yang, J., Singh, H., Hines, E. L., Schlaghecken, F., Iliescu, D. D., Leeson, M. S., & Stocks, N. G. (2012). Channel selection and classification of electroencephalogram signals: An artificial neural network and genetic algorithm-based approach. Artificial intelligence in medicine, 55(2), 117-126.

[15] Ko, K. E., & Sim, K. B. (2013). Harmony search-based hidden Markov model optimization for online classification of single trial eegs during motor imagery tasks. International Journal of Control, Automation and Systems, 11(3), 608-613.

[16] Hsu, W. Y. (2013). Application of quantum-behaved particle swarm optimization to motor imagery EEG classification. International journal of neural systems, 23(06).

[17] Hassani, K., & Lee, W. S. (2014, May). An incremental framework for classification of EEG signals using quantum particle swarm optimization. In Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA), 2014 IEEE International Conference on (pp. 40-45). IEEE.

[18] Gao, X. Z., Wang, J., Tanskanen, J., Bie, R., Wang, X., Guo, P., & Zenger, K. (2014). Optimal Classification of Epileptic EEG Signals Using Neural Networks and Harmony Search Methods. Journal of Software, 9(1), 230-239.

[19] Müller, K. R., Tangermann, M., Dornhege, G., Krauledat, M., Curio, G., & Blankertz, B. (2008). Machine learning for real-time single-trial EEG-analysis: from brain–computer interfacing to mental state monitoring. *Journal of neuroscience methods*, *167*(1), 82-90.

[20] Pei, S. C., & Jaw, S. B. (1989). Computation of discrete Hilbert transform through fast Hartley transform. Circuits and Systems, IEEE Transactions on, 36(9), 1251-1252.

[21] Smith, S. W. (1997). The scientist and engineer's guide to digital signal processing.

[22] de Albuquerque, V. H. C., de Alexandria, A. R., Cortez, P. C., & Tavares, J. M. R. (2009). Evaluation of multilayer perceptron and self-organizing map neural network topologies applied on microstructure segmentation from metallographic images. *NDT & E International*, *42*(7), 644-651.

[23] Noriega, L. (2005). Multilayer perceptron tutorial. *School of Computing. Staffordshire University*.