Indonesian Journal of Electrical Engineering and Computer Science
Vol. 9, No. 2, February 2018, pp. 403~409
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v9.i2.pp403-409

Performance of Support Vector Machine in Classifying EEG Signal of Dyslexic Children using RBF Kernel

AZA. Zainuddin[1], W. Mansor [2], Khuan Y. Lee[3] , Z. Mahmoodin[4]
[1,2]Faculty of Electrical Engineering, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia
[1,2,3,4]Computational Intelligence Detection RIG, Pharmaceutical Life Sciences CORE, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia
[2,3]Medical Engineering Technology Section, Universiti Kuala Lumpur, 53100 Gombak, Selangor, Malaysia

Article Info

Article history:

Received Oct 19, 2017
Revised Dec 22, 2017
Accepted Jan 14, 2018

Keywords:

Dyslexia
Electroencephalogram
RBF kernel
Support Vector Machine

ABSTRACT

Dyslexia is referred as learning disability that causes learner having difficulties in decoding, reading and writing words. This disability associates with learning processing region in the human brain. Activities in this region can be examined using electroencephalogram (EEG) which record electrical activity during learning process. This study looks into performance of Support Vector Machine (SVM) using RBF kernel in classifying EEG signal of Normal, Poor and Capable Dyslexic children during writing words and non-words. Discrete Wavelet Transform (DWT) with Daubechies order 2 was employed to extract the power of beta and theta waves of EEG signal. Beta and Theta/Beta ratio form the input features for classifier. Multiclass one versus one SVM was used in the classification where RBF kernel parameters and box constraint values were varied with the factor of 10 to analyze performance of the classifier. It was found that the best performance of SVM with 91% overall accuracy was obtained when both kernel scale and box constraint are set to one.

Copyright © 2018 Institute of Advanced Engineering and Science.
All rights reserved.

Corresponding Author:

AZA. Zainuddin,
Faculty of Electrical Engineering,
Universiti Teknologi MARA
40450 Shah Alam, Selangor,
Malaysia
Email: zuber@unikl.edu.my

1. INTRODUCTION

Dyslexia is neurobiological inefficiency of some part in the brain that makes the people experience difficulty in acquiring fluent skills in reading although they have received appropriate academic education at the same level as normal children [1]. Despite this learning disability, dyslexic children possess the same or above IQ level compared with normal children [2].

Several studies have been conducted to identify cognitive strengths and weaknesses of the children using computer model analysis from Gibson test [3]. While Malaysia Ministry of Education uses the Dyslexia check list as the instrument to identify the probability of the children having learning disability specific to dyslexia by measuring their capability in spelling, reading, and writing.

Beside visual, auditory, processing and word test to examine the etiology of dyslexia, further studies were carried out using imaging techniques such as functional Magnetic Resonance Imaging fMRI [4], Positron Emission Tomography PET [5], Magnetoencephalogram MEG [6] which examine cognitive process associated with learning disabilities. However, EEG analysis is the subject of interest in this study due to its practicality and cost-effective with high temporal resolution.

Electrical activities of the brain can be recorded and monitored noninvasively using EEG electrodes attached to the scalp. This signal shows activities of the brain region during executing a task such as

decoding, reading, and writing. EEG signal consists of several frequencies bands. Delta waves δ (1-4Hz), Theta waves θ (4-7Hz), Alpha waves α (8-12Hz), Beta waves β (13-30hz) and Gamma waves γ (31Hz and above) that indicate different activities and level of awareness in the brain.

Hence, several studies were conducted to extract and classify EEG signal in identifying Intelligent Quotient (IQ) [7] as well EEG related problem such as sleep studies [8], epileptic [9,10] mental task [11], mental imaginary [12], motor imaginary [13], brain-computer interface [14,15] and learning disability [16] to name a few. Later, the extracted EEG signal was subjected to classification for identification.

Various classification techniques have been investigated to identify dyslexia accurately. One of them is SVM, which is known as good performance classifier compared to other classifiers. SVM is a supervised binary classification algorithm that finds the optimal separating boundary in hyperplane by maximising the margin of two classes/training data. SVM has great ability in solving high dimension and nonlinear features. However, the performance of SVM in classifying dyslexia using the optimum value obtained by varying the scale of kernel parameter has not been reported.

It is anticipated that by tuning the kernel parameter of the SVM, the classifier can produce high accuracy in classifying dyslexia and perform better than other classifiers. This paper describes the classification of EEG signals of normal, poor dyslexic and capable dyslexic children using multiclass SVM binary learner through one versus one coding design. Varying scale of SVM and RBF kernel parameter is carried out to find the optimum parameters.

2. RESEARCH METHOD

In this work, the examination of the SVM performance in classifying dyslexia was carried out through several stages which include subject identification, EEG signal acquisition, notch and high pass filtering, power feature extraction, kernel parameter scale tuning, cross validation and classification as shown in Figure 1.

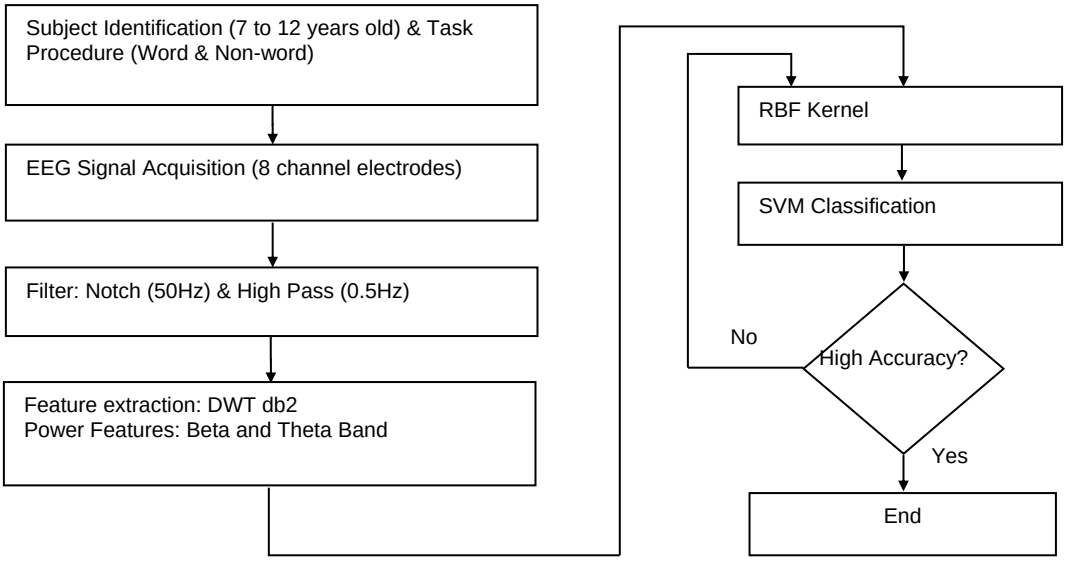


Figure 1. Flow Chart of EEG Signal Analysis

2.1 Subject Identification and Task Procedure

Wireless bio signal acquisition system g.nautilus was used to capture EEG signal from the scalp of the children. Head cover consists of 8 channel electrodes that are complied with international 10 to 20 electrode placement system was used during the recording. These electrodes were positioned at C3, P3, T7 and FC5 in the left side of the brain and C4, P4, T8 and FC6 at the right side of the brain as shown in Figure. 2. The system acquired EEG signal, amplified and sampled it using a sampling frequency 256Hz before transmitting the signal wirelessly to a personal computer for recording and analyzing.

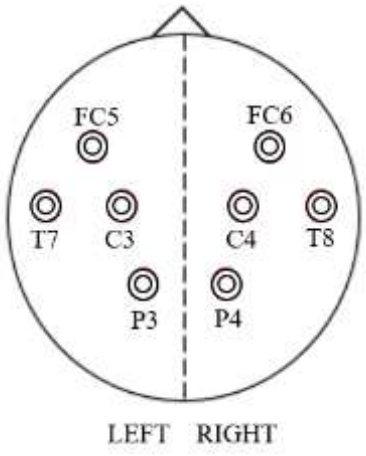


Figure 2. Electrode Placement in Left and Right Hemisphere of Brain.

In this study, the EEG data were recorded from 33 subjects with the age ranging from 7 to 12 years old. From the total subjects, the distribution is 8 normal, 17 poor dyslexics and 8 capable dyslexics. This data was acquired with the assistant from Dyslexia Association of Malaysia and Rakan Dyslexia Malaysia group.

Two categories of word were prepared for the subject; known word or word that was familiar to the subject with which can be visualized in their mind or have a specific meaning. Another category is non-word which has not seen before by the subject or word that have no specific meaning in particular and is not referring to anything. Three sets of word and non-word were prepared based on their age appropriate to their academic level. Set A was for subject of age 7 to 8, set B was for subject of age 9 to 10 and set C was for subject of age 11 to 12. Table 1 shows five tasks performed by the subject while their brain activities are recorded.

Table 1. Tasks That Were Performed During EEG Signal Recording

Task	Description
Task 1: Baseline	Subject was asked to relax and try to think of nothing in particular for 40 seconds.
Task 2: Simple Word	Three simple words were shown one by one and the subject was asked to write the word on a piece of paper then relax.
Task 3: Complex Word	Then another three complex words were shown one by one and the subject was asked to write the word they saw on a piece of paper then relax.
Task 4: Simple Non-Word	Three simple non-words were shown one by one and the subject was asked to write the word they saw on a piece of paper then relax.
Task 5: Complex Non-Word	Three complex non-words were shown one by one and the subject was asked to write the word they saw on a piece of paper then relax.

Altogether 170 datasets were collected where each dataset contains 8-electrode recording. Hence, the total number of data recorded was 1360. Out of this, sixty-five percent (65%) of the dataset was used for training data and the remaining thirty-five percent (35 %) of the dataset was used for testing data.

2.2 EEG Signal Pre-processing and Features Extraction

The recorded EEG signals were filtered using a notch filter to eliminate power line noise at 50Hz and a high pass filter with a cutoff frequency of 0.5Hz to remove dc offset. The data were analyzed using a program written in Matlab. Since EEG signal is non-stationary, time-scale analysis is more suitable for extracting the underlying information than other methods. The raw EEG signals were extracted using DWT to decompose the signal into frequency sub-bands as shown in Figure 3. In this work, input features were not normalized because the output variation was small.

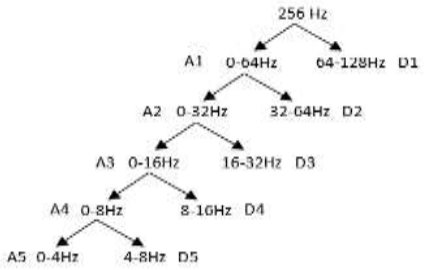


Figure 3. DWT Decomposition of EEG Signal

Out of several wavelet family, Daubiechies of order 2 (db2) was employed to provide EEG signal time-frequency scale representation as its ability to localize features and smoothing over EEG signal [17]. The detail coefficient D5 is theta band that indicates drowsiness and the detail coefficient D3 is beta band, which refers to active attention and was the subject of interest in this study. When a task is performed by the subject, the brain waves will shift towards increasing beta band frequency while the rest of the band frequency will be reduced.

Theta-Beta ratio is an indication of the relationship between internal**,** (slow activity) and sequential, (fast activity) [18,19]. Theta band represents the subconscious mind and beta band represents the conscious mind. Brain activation through theta-beta ratio was examined to analyze the brain state at a particular site between logical and spontaneous processing. Higher ratio indicates theta is dominant while lower ratio indicates beta is dominant.

2.3 Classification

In this stage, multiclass classification with one versus one was employed to classify normal, poor dyslexic and capable dyslexic. SVM with RBF kernel was then applied to the extracted band power features of Beta and Theta-Beta ratio. SVM classification is based on finding maximum margin separation boundary between two classes. In linear form, the separation can be done straight forward but for nonlinear condition, the data has to be placed in features space where the separation is performed in hyperspace. Kernel is a string that specifies the kernel function and is used to map the data from input space into a new space. There are three types of kernel function that can be used. They are known as Linear, Polynomial and RBF. Polynomial and RBF kernel are used for mapping non-linear data into hyperspace. The SVM classifier can be written as in Equation (1) and the RBF kernel function is shown in Equation (2).

$$f(x)=\sum_{i}^{N}\alpha_i y_i k(x_i,x)+b \quad (1)$$

$$k(x,x')=\exp\left(-\frac{||x-x'||^2}{2\sigma^2}\right) \quad (2)$$

$$\gamma=\frac{1}{2\sigma^2} \quad (3)$$

Equation (3) shows γ or kernel width that is a positive number specifying the kernel scale factor which is used to specify the shape of "peak" either broader or pointed bump. The SVM classifier with RBF kernel is given by Equation (4).

$$f(x)=\sum_{i}^{N}\alpha_i y_i \exp\left(-\frac{||x-x_i||^2}{2\sigma^2}\right)+b \quad (4)$$

The SVM classifier with RBF kernel has two parameters; kernel scale (γ) and box constraint (C). Box constraint is a regulation parameter which controls tradeoff between margin maximization and errors of training data. SVM with (C) is shown in Equation (5)

$$f(x)=C\sum_{i}^{N}\alpha_i y_i \exp\left(-\frac{||x-x_i||^2}{2\sigma^2}\right)+b \quad (5)$$

To obtain the optimal parameters, varying scale on SVM with RBF kernel was carried out. In the first analysis the box constraint was varied from 0.001 to 1000 by increasing factor of 10 while kernel scale was set to 1. In the second analysis the kernel scale, γ was varied from 0.001 to 1000 by increasing factor of 10 while the box constraint was fixed to 1. Cross-validation with K-fold equal to ten folds was applied to predicts classification accuracy with the lowest error is performed with training data.

Confusion matrix for multiclass were then employed in order to verify the performance of classification model. The sensitivity, specificity and accuracy were determined using Equation (6), (7) and (8) respectively.

$$Sensitivity, S_e = T_{PR} = \frac{T_P}{T_P + F_N} \quad (6)$$

$$Specificity, S_p = T_{NR} = \frac{T_N}{T_N + F_P} \quad (7)$$

$$Accuracy, A_c = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \quad (8)$$

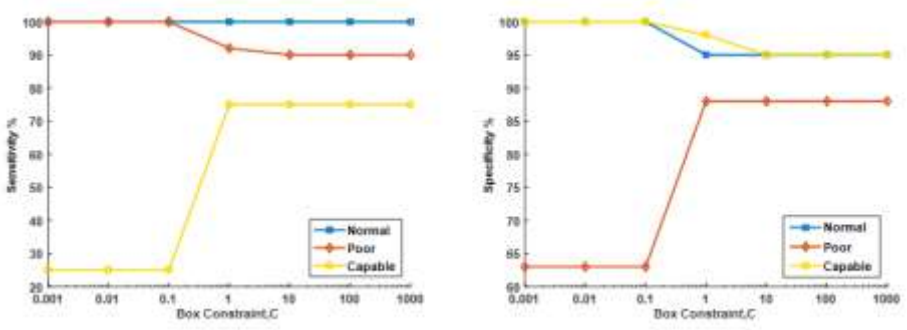
3. RESULTS AND ANALYSIS

Table 2 shows the result of k-fold cross-validation error for various C and kernel scales. It is obvious that scale 1 for both C and γ gives the lowest error, which is 23%.

Table 2. K-Fold Cross-Validation Error

	Scale						
Cross Validation	0.001	0.01	0.1	1	10	100	1000
Box Constraint, C	0.45	0.43	0.40	0.23	0.20	0.21	0.20
Kernel Scale, γ	0.71	0.71	0.71	0.23	0.38	0.55	0.52

The sensitivity versus C plot of the multiclass SVM classifier when C is varied from 0.001 to 1000 is shown in Figure 4(a). As can be seen, increasing C more than 0.1 decreases the classifier sensitivity from 100% to 92% for poor dyslexic, while for capable dyslexic the sensitivity rapidly increases from 25% to 75%. In contrast, the sensitivity for normal subject does not change and stays at 100%. Furthermore, increasing C above 1 give no changes to classifier sensitivity for all classes.



(a) (b)

Figure 4. Multiclass SVM Classification Performance When C is varied for Normal, Poor Dyslexic and Capable Dyslexic (a) Sensitivity (b) Specificity

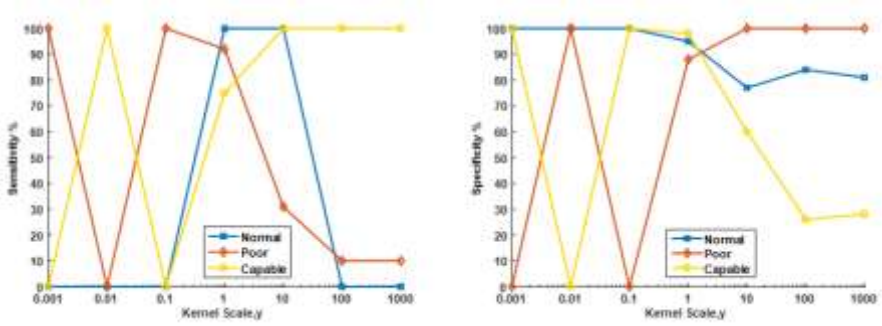
Figure 4(b) shows the specificity of multiclass SVM classification performance which was measured for various range of C (0.001 to 1000). It can be seen that the specificity for classifying capable

dyslexic and normal subject decreases from 100% to 98% and 95% respectively, while for poor dyslexic the specificity increases from 63% to 88% when C is set at 1. The result remains unchanged when C is above 1.

Although C in the range of 0.001 to 0.1 performs better in specificity for normal and capable dyslexic, it does not perform well for poor dyslexic. Thus, it can be concluded that C equals to 1 is the optimal setting that gives the best overall sensitivity and specificity for classifying normal, poor dyslexic and capable dyslexic.

Figure 5(a) and (b) shows the sensitivity and specificity for normal, poor dyslexic and capable dyslexic resulted from SVM classification when γ value is varied from 0.001 to 1000. When γ is set from 0.001 to 0.1, the SVM sensitivity for poor and capable dyslexic is fluctuated, between 0% and 100%. While for normal subject it is not sensitive at all. However, when γ is set to 1, the sensitivity increases to 100% for normal, 92% for poor dyslexic and 75% for capable dyslexic. Above scale of 10, the sensitivity drops tremendously when classifying normal and poor dyslexic.

The same trend is observed in the specificity for γ in the range of 0.001 to 0.1. At scale equal to 1, specificity for classifying normal subject is 95%, while for poor dyslexic and capable dyslexic, it is 88% and 98% respectively. The best sensitivity and specificity are obtained for all groups when γ is set to 1.



(a) (b)

Figure 5. Multiclass SVM Classification Performance When Kernel Scale is Varied for Normal, Poor Dyslexic and Capable Dyslexic (a) Sensitivity (b) Specificity

It is observed that in Figure 6, classifier accuracy for C is high, which is in the range of 94% to 89%. However, classifier accuracy is not stable for γ, which increases and decreases between 91% to 9%. When both γ and C are 1, the SVM accuracy is 91%. The accuracy decreases when both parameters is set above 1. Thus, the optimal value for C and γ is 1 since these values give good accuracy.

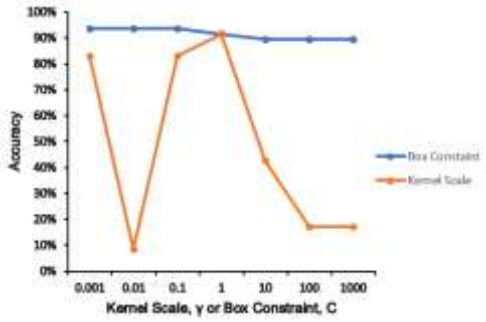


Figure 6. Multiclass SVM Classification Overall Accuracy for RBF Kernel

4. CONCLUSION

This work was carried out to examine the classification performance of multiclass SVM in distinguishing EEG signal of normal, poor and capable dyslexic children. The extraction of features which

are Beta and Theta-Beta ratio was carried out using wavelet db2 and these features were used as the input to the classifier. The box constraint of SVM and the RBF kernel parameter were varied to find the optimum results. Cross-validation also was carried out. The results obtained in this study shows that RBF kernel parameter γ affects the classification performance. Setting γ to 1 in the RBF kernel and C to the same value in the SVM yielded the highest accuracy, which is at 91%. The SVM with RBF kernel could classify the normal, poor dyslexic and capable dyslexic children accurately with high sensitivity and specificity using the optimum parameters.

ACKNOWLEDGEMENT

This work was supported by Fundamental Research Grant Scheme (FRGS), Malaysia (600-RMI/FRGS 5/3(137/2015)). The authors would like to thank Ministry of Higher Education, Malaysia, Research Management Institute and Faculty of Electrical Engineering, Universiti Teknologi MARA, Shah Alam, for financial support, facilities and various contributions, and to Dyslexia Association Malaysia for their assistance.

REFERENCES

[1] B. a. Shaywitz *et al.*, "The neurobiology of dyslexia," *Clin. Neurosci. Res.*, vol. 1, pp. 291–299, 2001.
[2] B. Sklar, J. Hanley, and W. W. Simmons, "A computer analysis of EEG Spectral Signatures from Normal and Dyslexic Children," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 1, pp. 20–26, 1973.
[3] H. M. Al-Barhamtoshy and D. M. Motaweh, "Diagnosis of Dyslexia using computation analysis," *2017 Int. Conf. Informatics, Heal. Technol. ICIHT 2017*, vol. 6, no. 2, pp. 462–482, 2017.
[4] M. Vandermosten, F. Hoeft, and E. S. Norton, "Integrating MRI brain imaging studies of pre-reading children with current theories of developmental dyslexia: a review and quantitative meta-analysis," *Curr. Opin. Behav. Sci.*, vol. 10, no. June, pp. 155–161, Aug. 2016.
[5] Y. Sun, J. Lee, and R. Kirby, "Brain Imaging Findings in Dyslexia," *Pediatr. Neonatol.*, vol. 51, no. 2, pp. 89–96, Apr. 2010.
[6] S. I. Dimitriadis *et al.*, "Altered temporal correlations in resting-state connectivity fluctuations in children with reading difficulties detected via MEG," *Neuroimage*, vol. 83, pp. 307–317, Dec. 2013.
[7] A. H. Jahidin, M. N. Taib, N. M. Tahir, and M. S. A. M. Ali, "IQ Classification via Brainwave Features: Review on Artificial Intelligence Techniques," *Int. J. Electr. Comput. Eng.*, vol. 5, no. 1, pp. 84–91, 2015.
[8] S. Motamedi-Fakhr, M. Moshrefi-Torbati, M. Hill, C. M. Hill, and P. R. White, "Signal processing techniques applied to human sleep EEG signals—A review," *Biomed. Signal Process. Control*, vol. 10, pp. 21–33, Mar. 2014.
[9] A. T. Tzallas, M. G. Tsipouras, and D. I. Fotiadis, "Epileptic seizure detection in EEGs using time-frequency analysis.," *IEEE Trans. Inf. Technol. Biomed.*, vol. 13, no. 5, pp. 703–710, 2009.
[10] H. Ullah, S. Mahmud, and R. H. Chowdhury, "Identification of Brain disorders by Sub-band Decomposition of EEG signals and Measurement of Signal to Noise Ratio," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 4, no. 3, p. 568, Dec. 2016.
[11] I. Güler and E. D. Ubeyli, "Multiclass support vector machines for EEG-signals classification.," *IEEE Trans. Inf. Technol. Biomed.*, vol. 11, no. 2, pp. 117–126, 2007.
[12] M. S. Bascil, A. Y. Tesneli, F. Temurtas, Á. Pca, and Á. L. V. Q. Á. Mlnn, "Multi-channel EEG signal feature extraction and pattern recognition on horizontal mental imagination task of 1-D cursor movement for brain computer interface," *Australas. Phys. Eng. Sci. Med.*, vol. 38, no. 2, pp. 229–239, 2015.
[13] S. K. Bashar, A. B. Das, and M. I. H. Bhuiyan, "Motor imagery movements detection of EEG signals using statistical features in the Dual Tree Complex Wavelet Transform domain," *2015 Int. Conf. Electr. Eng. Inf. Commun. Technol.*, pp. 1–6, 2015.
[14] X. Li, X. Chen, Y. Yan, W. Wei, and Z. J. Wang, "Classification of EEG signals using a multiple kernel learning support vector machine," *Sensors (Basel).*, vol. 14, no. 7, pp. 12784–12802, 2014.
[15] A. B. M. A. Hossain, W. Rahman, and M. A. Riheen, "Left and Right Hand Movements EEG Signals Classification Using Wavelet Transform and Probabilistic Neural Network," *Int. J. Electr. Comput. Eng.*, vol. 5, no. 1, pp. 92–101, 2015.
[16] A. Z. A. Zainuddin, K. Y. Lee, W. Mansor, and Z. Mahmoodin, "Optimized KNN classify rule for EEG based differentiation between capable dyslexic and normal children," in *2016 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, 2016, pp. 685–688.
[17] T. Gandhi, B. K. Panigrahi, and S. Anand, "A comparative study of wavelet families for EEG signal classification," *Neurocomputing*, vol. 74, no. 17, pp. 3051–3057, 2011.
[18] J. E. Walker and C. A. Norman, "The Neurophysiology of Dyslexia: A Selective Review with Implications for Neurofeedback Remediation and Results of Treatment in Twelve Consecutive Patients," *J. Neurother.*, vol. 10, no. 1, pp. 45–55, Dec. 2006.
[19] C. Spironelli, B. Penolazzi, and A. Angrilli, "Dysfunctional hemispheric asymmetry of theta and beta EEG activity during linguistic tasks in developmental dyslexia," *Biol. Psychol.*, vol. 77, no. 2, pp. 123–131, 2008.