Machine learning and deep learning performance in classifying dyslexic children's electroencephalogram during writing

Ahmad Zuber Ahmad Zainuddin[1,2,3,4], Wahidah Mansor[1,2,3], Khuan Yoot Lee[2,3], Zulkifli Mahmoodin[4]
[1]Microwave Research Institute, Universiti Teknologi MARA, Shah Alam, Malaysia
[2]School of Electrical Engineering, College of Engineering, Universiti Teknologi MARA, Shah Alam, Malaysia
[3]Computational Intelligence Detection RG, Health and Wellness ReNEU, Universiti Teknologi MARA, Shah Alam, Malaysia
[4]Medical Engineering Technology Section, Universiti Kuala Lumpur British Malaysian Institute, Malaysia

Article Info

Article history:

Received Jul 17, 2021
Revised Jun 23, 2022
Accepted Jul 18, 2022

Keywords:

Deep learning
Dyslexia
Electroencephalogram
Extreme learning machine
K-nearest neighbors
Long short-term memory
Support vector machine

ABSTRACT

Dyslexia is a form of learning disability that causes a child to have difficulties in writing alphabets, reading words, and doing mathematics. Early identification of dyslexia is important to provide early intervention to improve learning disabilities. This study was carried out to differentiate EEG signals of poor dyslexic, capable dyslexic, and normal children during writing using machine learning and deep learning. three machine learning algorithms were studied: k-nearest neighbors (KNN), support vector machine (SVM), and extreme learning machine (ELM) with input features from coefficients of beta and theta band power extracted using discrete wavelet transform (DWT). As for the deep learning (DL) algorithm, long short-term memory (LSTM) architecture was employed. The kernel parameters of the classifiers were optimized to achieve high classification accuracy. Results showed that db8 achieved the greatest classification accuracy for all classifiers. Support vector machine with radial basis function kernel yields the highest accuracy which is 88% than other classifiers. The support vector machine with radial basis function kernel with db8 could be employed in determining the dyslexic children's levels objectively during writing.

Corresponding Author:

Wahidah Mansor
Microwave Research Institute, Universiti Teknologi MARA
40450 Shah Alam, Selangor, Malaysia.
Email: wahidah231@uitm.edu.my

1. INTRODUCTION

Fundamental learning skills that children need to acquire are reading, writing, and doing calculations [1], [2]. However, this is the most challenging skill faced by dyslexic children and becomes apparent when they start schooling compared with normal children. Signs of dyslexia become obvious when learning activities required them to read and write. Early identification of dyslexia is crucial to help them to master the basics before academic content becomes harder. The level of difficulties faced by children is varied, some dyslexic children have severe difficulties and some have mild difficulties. Classification of this level requires a skilled examiner who assesses the children learning disability. Intervention programs designed specifically for dyslexic children would help them to overcome their learning difficulties early and provide an opportunity for them to be on par with normal learners.

The study of dyslexia through brain-based techniques was implemented previously using structural and functional connectivity imaging such as magnetoencephalography (MET) [3], functional magnetic resonance imaging (fMRI) [4], and positron emission tomography (PET) [5]. However, these techniques incurred high costs, are bulky and require the subject to lie down which is not suitable for children to perform activities while their brain activities images are monitored. Electroencephalography (EEG) provides a better solution by analyzing the brain's electrical activity. EEG consists of five frequency bands related to information from different parts of the brain. Most previous studies only focused on the differences between EEG signals of dyslexic and normal learners [6]. They do not further categorize dyslexic children into several types. This work analyses the dyslexic children's EEG signals and categorizes them into poor dyslexic and capable dyslexic. This is important to monitor dyslexic children's progress of learning in the intervention program objectively.

Studies on dyslexia mainly focus on reading [7] with limited work concentrating on writing, even if it is one of the main indicators and part of the consequence of having dyslexia [8]. Writing can be seen as a complex process that requires coordination between cognitive related processes and motor skills [9]. In differentiating brain activity of normal and dyslexic children, it is paramount to seek tasks that could stimulate the required changes that differ between the groups, in which the writing process is seen as an option as it requires active attention from learners. During writing, the brain activates two main areas known as Broca and Wernicke. The language area of Broca is mainly responsible for expressive writing and speaking while Wernicke is associated with understanding the language of both writing and speaking. Other areas that are found to be related to machines are temporal and parietal where a study found these areas to be responsible for written word comprehension and converting visual images to written symbols that involve the programming of motor areas [8].

Previous studies investigating differences in EEG signals involved in determining intelligent quotient (IQ) [10] as well as EEG related problems such as sleep studies [11], epileptic, mental task [12], mental imaginary [13], motor imaginary [14], brain-computer interface [15], epilepsy [16], autism spectrum disorder [17] and learning disabilities [18]. This is done by applying several machine learning such as multilayer perceptron (MLP), linear discrimination analysis (LDA) and artificial neural network (ANN) to identify this symptom by examining the EEG signals. However, k-nearest neighbors (KNN) [19], support vector machine (SVM) [14] and extreme learning machine (ELM) [17] gradually gain popularity due to their algorithm simplicity. One of the main drivers behind the application of deep learning (DL) in EEG is due to the requirement for the application of features that indicate a specific mental state in machine learning (ML). Whereas DL would identify the features automatically. Most of the applications of DL in examining EEG use a convolution neural network (CNN) that process the images [12], [20]. A few researchers use a recurrent neural network (RNN) to process the signal through long short-term memory (LSTM) [21].

This paper describes the performance of ML and DL in identifying types of dyslexia from EEG signals during writing. Three types of ML classifiers which are KNN, SVM and ELM together with DL utilized LSTM were studied and optimized to attain the highest classification accuracy in distinguishing between poor dyslexic, capable dyslexic and normal children based on EEG signals during writing-related tasks. These are important steps to make use of EEG more practical and to become less dependent on trained professionals [22]. The classifiers were selected due to their classification capability and performance.

2. RESEARCH METHOD

The process of classifying EEG signals of normal, poor and capable dyslexic children using KNN, SVM, ELM classifier and DL is shown in Figure 1. The work was implemented in stages that start from the subject selection, data acquisition procedure which includes EEG signal recording, removal of artifacts, extracting features, optimization of kernel selection to achieve the highest classification accuracy and finally, the selection of the optimum classifier.

2.1. Subject identification

In this work, subjects with an age range from 7 to 12 years old were chosen as signs of their dyslexic and learning disability can be seen and readily diagnosed. They have also already started receiving formal learning activities at schools such as spelling, reading, and writing. A total of 54 subjects participated in this study. Among them are 18 poor dyslexic subjects, 18 capable dyslexic subjects and 18 normal subjects as the control group. The categorization of capable and poor dyslexics was made based on the assessment that was carried out by the Dyslexia Association of Malaysia.

Subject denoted as poor dyslexic is based on those that exhibit difficulties in both writings and reading when compared with children in their age group. While capable dyslexic is a subject that has undergone intervention program and has shown improvement in their overall reading and writing capabilities. Before the assessments were carried out, each subject's medical history, and background, along with the use of their dominant hand during writing, either right or left, were duly recorded. This is to ensure consistency and data conformity among all subjects. In addition, subjects were also selected among those without any record of prior neurological disorder or having any history of suffering from a double deficit. This work has obtained ethical approval for conducting the research from Research Ethics Committee UiTM. An explanation of the research and the consent form was given to the subject's caretaker upon receiving approval.

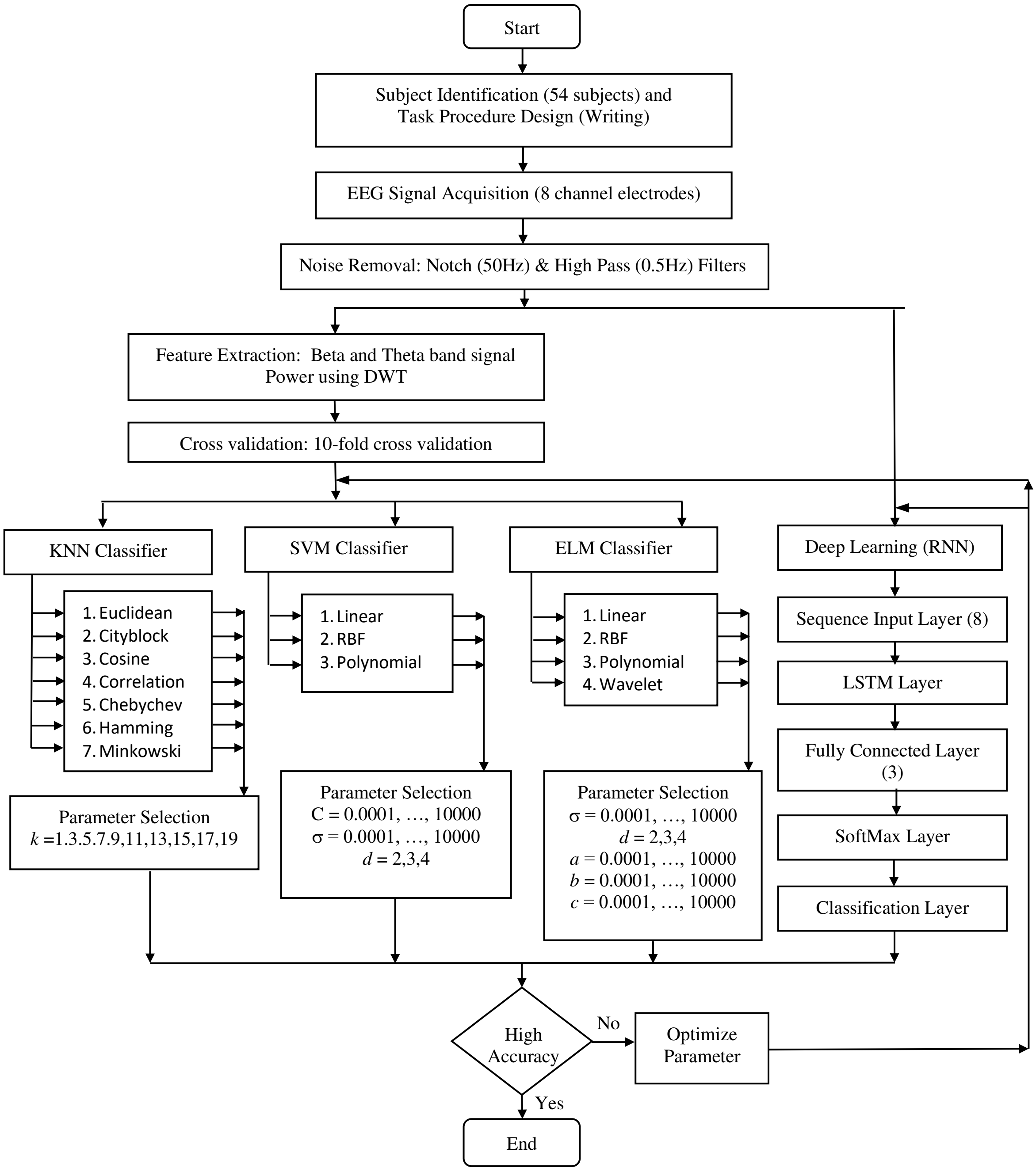

Figure 1. Flowchart showing the process of distinguishing EEG signals of subjects

2.2. Writing task

Five tasks were designed for writing assessment as shown in Table 1. The first task is relaxing with the eye closed to get EEG baseline recording and to eliminate physiological conditions that may influence the signal. The second task is writing known words (tasks B1 and B2) that will activate brain areas associated with memory recall and matching of the written word within the visual word form area (VWFA) in the occipital temporal regions. Task C1 and C2 are writing of non-words which are words that are made up and has no meaning. This will invoke the brain areas of Wernicke and Broca that will try to decode the non-word rather than recalling it from memory.

Table 1. Tasks designed for the assessment

Categories	Task
A	Relax state with close eye
B1	Writing of simple words
B2	Writing of complex words
C1	Writing of simple non-words
C2	Writing of complex non-words

2.3. EEG signal acquisition and pre-processing

In this study, EEG signals were recorded from normal and dyslexic subjects by our team and the database is not available to the public. During EEG signal acquisition, the subject was seated in a controlled environment holding a pencil with a paper on the table as shown in Figure 2. A computer screen was placed in front of the subject displaying a set of words one by one in turn and the subject was required to duplicate the word seen by writing it on the paper for tasks B1, B2, C1 and C2. Each set of words contains related alphabets that pose a difficulty for a dyslexic to write.

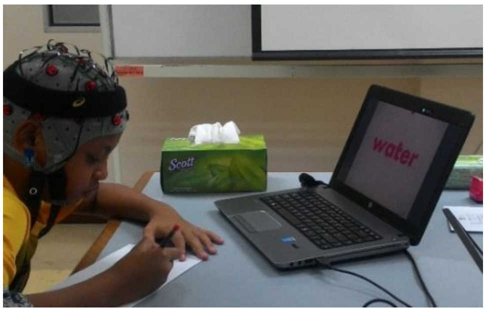

Figure 2. Recording of EEG signal from a subject during the assessment

EEG signals were recorded using g.Nautilus wireless biosignal acquisition system as shown in Figure 2 while subject completing tasks. A total of eight electrodes were positioned on the scalp following the International 10-20 electrode placement system. The placement of the electrodes is based on known areas that are associated with writing and reading found in our previous work [23]. Within the learning pathway of the left side of the brain, EEG signals were recorded from electrode locations C3, P3, T7 and FC5. To monitor any alternative pathway that may exist on the right side of the brain, signals were recorded mirrored to the left side of locations C4, P4, T8 and FC6. All signals were sampled at 250 Hz with a 24-bit resolution. Pre-processing of all acquired signals include the removal of power-line interference with a Notch filter and the removal of DC offset was achieved through a high pass filter having a cut-off frequency of 0.5 Hz. The EEG signals were then saved as *.mat* files for feature extraction and classification using a program written in MATLAB. Table 2 shows related electrode positions and their corresponding function. A total of 216 datasets containing 1,728 EEG signals recording were attained from the 8 electrodes. Out of these, 70% was used for the training dataset and the remaining 30% was for the testing dataset.

Table 2. Electrode positions and functions

Area	Brain Hemisphere Left/Right	Functions
Parietal Lobe	C3/C4	Analyze word – sensorimotor integration
Wernicke's Area	P3/P4	Recognition of word – cognitive processing
Temporal Lobe	T7/T8	VWFA - memory representation of the letter
Broca's Area	FC5/FC6	Connect sounds to letters - motor plan for vocalization

2.4. Feature extraction

Raw EEG signal recorded from the subject's scalp during the writing task contains all the signal frequencies that were combined in one single EEG signal. Since the EEG signal is non-stationary, the separation of a frequency band to localize features according to brain activities was achieved using the time-frequency scale representation discrete wavelet transform (DWT) which decomposes the signal into detail and approximation coefficients at different sub-bands [24]. The wavelet decomposition involves two digital filters which are low-pass and high pass filters denoted as g and h as shown in (1) and (2):

$$A[k] = \sum_n x[n] - g[2k - n] \quad (1)$$

$$D[k] = \sum_n x[n] - h[2k - n] \quad (2)$$

where *x[n]* represents the signals, *D[k]* is the detail, *A[k]* is the approximation.

Five frequency sub-bands as shown in Table 3 were obtained from the decomposition of the signal using Daubechies mother wavelet with orders 2, 4, 6 and 8 to provide smooth EEG signals. Each sub-band provides information related to brain activities. Detail coefficients at level 3 (D3) which represents the beta band and level 5 (D5) which represents the theta band were frequency bands of interest in this study. The beta band corresponds to learning-related tasks while theta band corresponds to relax conditions.

Table 3. Decomposition level of EEG signal frequency sub-band

Decomposition Level	Frequency Range	EEG Band
D1	64-128	Noise
D2	32-64	Gamma
D3	16-32	Beta
D4	8-16	Alpha
D5	4-8	Theta
A5	0-4	Delta

Band power calculates the magnitude of the signal in response to a given stimulus or task under measurement. It indicates activation and allows the determination of area localization to take place. Power features for EEG signal from the reconstruction of the beta band and theta band were computed using (3) to get the coefficient. Signal values are shown by symbol *x* and the signal length by symbol *L*.

$$Band\ Power = \frac{\sum x^2}{L(x)} \quad (3)$$

The ratio of theta/beta band power was calculated using (4). The coefficients of the beta band power served as the first feature vector and the coefficient of the ratio of theta/beta band power served as the second feature vector of the classifier. Having more than one feature vector would make classifier prediction more accurate and reliable.

$$Ratio = \frac{Theta\ Band\ Power}{Beta\ Band\ Power} \quad (4)$$

2.5. Cross-validation and classifier

Model evaluation was carried out using cross-validation to check how well a model generalizes to training data before the model is used in new data. It was accomplished by splitting the training data into several folds. K-fold cross-validation that splits the dataset into 10 folds was applied to determine the error rate in the model parameter selection. For each fold, it contains a dataset with equal distribution of classes to avoid bias in the results. After setting up the cross-validation, the classification was performed.

2.5.1. k-nearest neighbors

KNN performs classification by computing the distance of new data to training data that is the nearest [25]. The selected ‘k’ value is the number of neighbors that will be referred to and the new data will be classified according to the majority rule. This work sets the k value starting from 1 to a maximum of 19 with an incremental of 2. New data distances to the training sample were calculated using functions as shown in (5) to (10). *D(x,y)* is the distance of sample *x* and *y,* while x_i and y_i are the i^{th} features of the sample and the number of features represented by *k*. Note that Hamming distance is equivalent to Cityblock distance if it is looked at in the form of a real coordinate for binary string. Changing the *q* value in Minkowski distance would form Cityblock, Euclidean and Chebychev distance.

$$\text{Euclidean } D(x, y) = \sqrt{\sum_{i=1}^{k} (x_i - y_i)^2} \quad (5)$$

$$\text{Cityblock/Hamming} D(x, y) = \sum_{i=1}^{k} |x_i - y_i| \quad (6)$$

Correlation $D(x,y) = \frac{\sum_{i=1}^{k}(x_i-\overline{x_i})(y_i-\overline{y_i})}{\sqrt{\sum_{i=1}^{k}(x_i-\overline{x_i})^2 \sum_{i=1}^{k}(y_i-\overline{y_i})^2}}$ (7)

Cosine $D(x,y) = \frac{\sum_{i=1}^{k}(x_i \times y_i)}{\sqrt{\sum_{i=1}^{k}(x_i)^2} \times \sqrt{\sum_{i=1}^{k}(y_i)^2}}$ (8)

Chebychev $D(x,y) = max|x_i - y_i|$ (9)

Minkowski $D(x,y) = \left(\sum_{i=1}^{k}(|x_i - y_i|)^q\right)^{\frac{1}{q}}$ (10)

where $q \geq 1$.

2.5.2. Support vector machine

SVM works by computing the maximum boundary separation based on space optimization between two sets of classes [26]. For linear based categorization, a simple linear kernel can be utilized to achieve the separation, but in non-linear cases, data are to be placed in features space where the isolation of data can only be implemented in hyperspace. This nonlinear separation can be achieved by employing radial basis function (RBF) and the Polynomial kernel. This work uses multiclass SVM with a one versus one mechanism, separating each pair of classes against each other and using a majority voting scheme to determine the output. The SVM classifier can be written as in (11):

$f(x) = \sum_{i}^{N} \alpha_i y_i k(x_i, x) + b$ (11)

where α_i is the weight vector, y_i is the target vector, *N* is the size of training data, *b* is the bias and $k(x_i,x)$ is the SVM kernel as shown in (12) to (14). In this work, the parameters of all kernels were varied to obtain the optimal solution;

Linear Kernel $k(x_i, x) = x_i, x$ (12)

Polynomial Kernel $k(x_i, x) = (x_i, x + 1)^d$ (13)

where *d* is the degree of the polynomial

RBF Kernel $k(x_i, x) = exp\left(-\frac{\|x_i - x\|^2}{2\sigma^2}\right)$ (14)

where σ is the kernel width.

2.5.3. Extreme learning machine

In achieving fast learning speed and improved generalization, this work investigates ELM which is a feedforward neural network with a single hidden layer that is based on the risk minimization principle [27]. The only parameter that needed to be learned is the weights between hidden and output neurons, which are determined analytically. For *N* arbitrary distinct samples *(xi, ti)* $R^n \times R^m$, standard ELM with L hidden nodes and activation function *g(x)* were modelled as in (15):

$\sum_{i=1}^{L} \beta_i g(a_i b_i x_j) = t_j, j = 1, \dots, N$ (15)

where β_i is the weight vector connecting the i^{th} hidden node and output node; a_i is the weight factor connecting the i^{th} hidden node and the input node; b_i is the impact factor of the i^{th} hidden node.

As shown in (15) can be written efficiently as (16):

$H\beta = T$ (16)

The main goal of ELM is to achieve the smallest error in training and the smallest norm of output weights. The solution using least-square that computes the minimum norm using (17):

$\hat{\beta} = H \dagger T$ (17)

where H† is the Moore Penrose generalized inverse of hidden layer output matrix H.

Kernels are employed as output weight and integrated into ELM to obtain better generalization with less user intervention. In (18) to (21) shows the kernel for ELM that were used in this work:

Linear kernel $K(x, y) = x \times y$ (18)

Polynomial kernel $K(x, y) = exp(x \times y + n)^d$ (19)

where d is the degree of the polynomial

RBF kernel $K(x, y) = exp\left(\frac{-\|x-y\|^2}{\sigma}\right)$ (20)

where σ is the kernel width

Wavelet Kernel $K(x, y) = cos\left(\frac{d\|x-y\|}{c}\right) exp\left(\frac{-\|x-y\|^2}{b}\right)$ (21)

where b, c and d are the parameters for the wavelet kernel

2.5.4. Deep learning

Deep learning is a multilayer network structure that can extract significant features by learning from a lower layer to a higher layer. One of the architectures used in deep learning is RNN. In this work, LSTM which is the extension of RNN architecture was used since it can deal with sequential data [28]. It works with sequence and time-series data such as EEG signals. RNN take input in the shape of a sequence $x = (x_1, ..., x_T)$ and compute hidden vector sequence $h=(h_1, ..., h_T)$ and output vector $y=(y_i, ..., y_T)$ by iterating the following equations from $t = 1$ to T:

$$h_t = H(W_{xh}X_t + W_{hh}h_{t-1} + b_h) \quad (22)$$

$$y_t = W_{hy}h_t + b_y \quad (23)$$

where W terms denote weight matrices, b terms denote bias vectors and H is the hidden layer function. The hidden layer function for LSTM is computed in the following set of (24) to (28):

$$i_t = \sigma(W_{xi}X_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i) \quad (24)$$

$$f_t = \sigma\left(W_{xf}X_t + W_{hf}h_{t-1} + W_{ci}c_{t-1} + b_f\right) \quad (25)$$

$$c_t = f_t c_{t-1} + i_t\, tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \quad (26)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o) \quad (27)$$

$$h_t = o_t\, tanh(c_t) \quad (28)$$

Here σ is the logistic sigmoid function, and the components of the LSTM model, referred to as input gate, forget gate, output gate and cell activation vectors are denoted as i, f, o and c respectively. In this work, to achieve the optimal solution, the mini-batch size and max epoch were varied.

3. RESULTS AND DISCUSSION

Table 4 shows the accuracy of each distance function in the KNN classifier using features from DWT with the nearest neighbor parameter varied from 3 to 17 and the rule was set to calculate the nearest neighbor. The best result for each distance function and parameter is presented in Table 4. Note that the Correlation distance function with features db2 gives the highest accuracy which is 81.67% while keeping the nearest neighbor at 3 and using the smallest rule. It also found that db2 and db8 produce better results compared to other Daubechies wavelet orders. In terms of nearest neighbor, the highest accuracy can be achieved when the number is 3 or 5 for all distance functions except for Hamming.

Table 4. Accuracy of KNN classifier

Distance Function	Parameter			Accuracy
	DWT	N-neighbor	Rule	
Euclidean	db8	5	Smallest	78.33
Cityblock	db2	3	Nearest	73.33
Cosine	db8	3	Smallest	76.67
Correlation	db2	3	Smallest	81.67
Chebychev	db8	5	Smallest	78.33
Hamming	db8	11	Nearest	43.33
Minkowski	db8	5	Smallest	78.33

Table 5 shows the accuracy of each kernel for the SVM classifier. Parameters of each kernel were tested from 0.0001 to 1,000 to find better accuracy but only the parameter with the highest accuracy for each wavelet is shown in Table 5. It was found that RBF kernel with db8, the scale of 10 and box constraint of 1 gives the highest accuracy which is 88.33%. Overall, the RBF kernel outperforms other kernels for all wavelets. It also found that db2 and db8 give better performance in the SVM classifier. For the linear kernel, db8 give an accuracy of 85%. For the Polynomial kernel, db2 provides an accuracy of 71.67%.

Table 5. Accuracy of SVM classifier

Kernel	Parameter				Accuracy
	DWT	Box C	Order	Scale	
Linear	db8	-	-	-	85
RBF	db8	1	-	10	88.33
Polynomial	db2	0.1	4	-	71.67

Table 6 shows the accuracy of each kernel in the ELM classifier. It was found that the Wavelet kernel with db8 gives the highest accuracy which is 83.33% compared to other kernels in the ELM classifier. The performance of wavelet kernel is slightly better compared to RBF kernel. Overall, db8 gives the best features for Linear, RBF and wavelet kernels. In the Polynomial kernel, db2 provides the highest accuracy.

Table 6. Accuracy of ELM classifier

Kernel	Parameter				Accuracy
	DWT	Para	Order	Para	
Linear	db8	-	-	-	68.3
RBF	db8	1	-	-	81.67
Polynomial	db2	0.1	2	-	73.33
Wavelet	db8	1	1	0.1	83.33

The performance of DL using the LSTM is shown in Table 7. It was found that LSTM using epochs 30 with minibatch size 20 and Adams solver produces a higher accuracy which is 83.78% in detecting capable dyslexic children. In recognizing poor dyslexic children and normal children, the accuracy is 78.79% and 81.82% respectively. This shows that LSTM can differentiate the dyslexic subject between normal, poor and capable subjects during writing.

The sensitivity and specificity of all classifiers that produces the highest accuracy using their optimal parameters and functions are shown in Table 7. It is found that using the wavelet db8 as the feature extraction method, the SVM with RBF kernel provides the most consistent high accuracy, sensitivity and specificity in distinguishing all subjects. All machine classifiers perform very well and provide the highest accuracy in recognizing capable dyslexic children compared to when differentiating normal children and poor dyslexic children. In the case of detecting poor dyslexic children, the most accurate is SVM (91.67%) and the lowest is the LSTM which can only provide an accuracy of 78.79%. LSTM produces low accuracy than other classifiers since it has to find features itself, unlike other classifiers where features selection is given, however, LSTM still produce encouraging performance. For LSTM, to increase the accuracy, the number of subjects needs also to be increased. In recognizing capable dyslexic children, the highest accuracy which is 93.55% is achieved from SVM and the lowest accuracy is given by LSTM which is 83.78%.

This work has shown that the SVM with RBF kernel is the optimum classifier for distinguishing normal, poor dyslexic and capable dyslexic children. The recognition of poor and capable dyslexic children using machine learning and deep learning carried out in this work is unique and has not been reported by previous researchers. Even though there is a report revealed the performance of two types of machine learning (k-means and artificial neural network) and fuzzy logic in detecting normal and dyslexic children [6] using EEG, they do not further separate the dyslexic children into poor and capable dyslexic categories.

Table 7. Performance of classifiers with the optimized parameter (for machine learning)

Group	Performance	Classifier			
		KNN db2 Correlation	SVM db8 RBF	ELM db8 Wavelet	Deep Learning LSTM 30 epochs
Normal	Sensitivity	80.00	90.00	85.00	63.94
	Specificity	92.50	92.50	85.00	90.91
	Accuracy	88.33	91.67	85.00	81.82
Poor	Sensitivity	80.00	85.00	70.00	68.18
	Specificity	85.00	95.00	100.00	84.09
	Accuracy	83.33	91.67	90.00	78.79
Capable	Sensitivity	85.00	90.00	95.00	81.82
	Specificity	95.24	95.24	90.91	84.62
	Accuracy	91.94	93.55	92.19	83.78

4. CONCLUSION

In this work, KNN, SVM, ELM and DL were used to categorize EEG signals of normal children and differentiate between poor and capable dyslexic children during tasks that involved writing known words and non-words. The performance of each classifier was analyzed and assessed using a confusion matrix to ascertain its accuracy as well as its sensitivity and specificity. Results showed that the SVM classifier with RBF kernel and using features from db8 gives 88.33% overall accuracy outperformed KNN, ELM and DL. It was also found that features from db8 yielded the highest accuracy in determining normal, poor dyslexic and capable dyslexic subjects. In this study, the categories of dyslexic children can be distinguished through writing using ML and DL. In future work, more subjects have to be used to perform the classification using DL to get better accuracy. The optimum classifier found in this study can be used in the development of an automated dyslexia diagnosis system by other researchers which would help the school to design the appropriate intervention program for the dyslexic children and assist the trainer to monitor their progress during the training.

ACKNOWLEDGEMENTS

This work was supported by fundamental research grant scheme (FRGS), Malaysia (600-RMI/FRGS 5/3(137/2015)). The authors would like to thank the Ministry of Education, Malaysia for financial support and Research Management Institute and Faculty of Electrical Engineering, Universiti Teknologi MARA, Shah Alam, for their support and providing facilities as well as Dyslexia Association Malaysia for their assistance in the recruitment of dyslexic subjects and advice.

REFERENCES

[1] K. E. Waldie, A. J. Wilson, R. P. Roberts, and D. Moreau, “Reading network in dyslexia: Similar, yet different,” *Brain and Language*, vol. 174, pp. 29–41, Nov. 2017, doi: 10.1016/j.bandl.2017.07.004.
[2] E. S. Norton, S. D. Beach, and J. D. E. Gabrieli, “Neurobiology of dyslexia,” *Current Opinion in Neurobiology*, vol. 30, pp. 73–78, Feb. 2015, doi: 10.1016/j.conb.2014.09.007.
[3] 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, doi: 10.1016/j.neuroimage.2013.06.036.
[4] B. Tomasino *et al.*, “Double-letter processing in surface dyslexia and dysgraphia following a left temporal lesion: A multimodal neuroimaging study,” *Cortex*, vol. 73, pp. 112–130, Dec. 2015, doi: 10.1016/j.cortex.2015.08.010.
[5] Y. Sun, J. Lee, and R. Kirby, “Brain imaging findings in dyslexia,” *Pediatrics & Neonatology*, vol. 51, no. 2, pp. 89–96, Apr. 2010, doi: 10.1016/S1875-9572(10)60017-4.
[6] H. M. Al-Barhamtoshy and D. M. Motaweh, “Diagnosis of dyslexia using computation analysis,” in *2017 International Conference on Informatics, Health & Technology (ICIHT)*, Feb. 2017, vol. 6, no. 1, pp. 1–7, doi: 10.1109/ICIHT.2017.7899141.
[7] H. Perera, M. F. Shiratuddin, K. W. Wong, and K. Fullarton, “EEG signal analysis of real-word reading and Nonsense-word reading between adults with dyslexia and without dyslexia,” in *2017 IEEE 30th International Symposium on Computer-Based Medical Systems (CBMS)*, Jun. 2017, pp. 73–78. doi: 10.1109/CBMS.2017.108.
[8] P. Drotár and M. Dobeš, “Dysgraphia detection through machine learning,” *Scientific Reports*, vol. 10, no. 1, Art no. 21541, Dec. 2020, doi: 10.1038/s41598-020-78611-9.
[9] M. T. Borges, L. C. S. Aprígio, C. A. S. Azoni, and P. A. P. Crenitte, “Types of handwriting and signs of dysgraphia in children and adolescents with learning difficultie,” *Revista CEFAC*, vol. 22, no. 6, pp. 1–8, 2020, doi: 10.1590/1982-0216/202022617719.
[10] N. H. R. Azamin, A. H. Jahidin, M. S. A. Megat Ali, and M. N. Taib, “Intelligence quotient and perceptual ability: an inter-relationship based on brainwave power ratio features,” *Journal of Fundamental and Applied Sciences*, vol. 9, no. 6S, Feb. 2018, doi: 10.4314/jfas.v9i6s.70.
[11] N.-H. Lin *et al.*, “Detecting rapid eye movement sleep using a single EEG signal channel,” *Expert Systems with Applications*, vol. 87, pp. 220–227, Nov. 2017, doi: 10.1016/j.eswa.2017.06.017.

[12] Z. Jiao, X. Gao, Y. Wang, J. Li, and H. Xu, "Deep convolutional neural networks for mental load classification based on EEG data," *Pattern Recognition*, vol. 76, pp. 582–595, Apr. 2018, doi: 10.1016/j.patcog.2017.12.002.
[13] 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," *Australasian Physical & Engineering Sciences in Medicine*, vol. 38, no. 2, pp. 229–239, 2015, doi: 10.1007/s13246-015-0345-6.
[14] N. Hamzah, N. Zaini, M. Sani, and N. Ismail, "EEG analysis on actual and imaginary left and right hand lifting using support vector machine (SVM)," *International Journal of Electrical and Electronic Systems Research*, vol. 10, no. 1, 2017.
[15] A. B. M. A. Hossain, M. W. Rahman, and M. A. Riheen, "Left and right hand movements EEG signals classification using wavelet transform and probabilistic neural network," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 5, no. 1, pp. 92–101, Feb. 2015, doi: 10.11591/ijece.v5i1.pp92-101.
[16] A. S. M. Murugavel and S. Ramakrishnan, "An optimized extreme learning machine for epileptic seizure detection," *International Journal of Computer Science*, vol. 14, no. 4, pp. 212–221, 2014.
[17] S. Ding, L. Guo, and Y. Hou, "Extreme learning machine with kernel model based on deep learning," *Neural Computing and Applications*, vol. 28, no. 8, pp. 1975–1984, Aug. 2017, doi: 10.1007/s00521-015-2170-y.
[18] D. Meedeniya and I. Rubasinghe, "A review of supportive computational approaches for neurological disorder identification," in *Interdisciplinary Approaches to Altering Neurodevelopmental Disorders*, 2020, pp. 271–302, doi: 10.4018/978-1-7998-3069-6.ch016.
[19] S. A. M. Aris, S. Z. A. Jalil, N. A. Bani, H. M. Kaidi, and M. N. Muhtazaruddin, "EEG features extraction and k-NN classification during eyes closed," in *2016 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, Dec. 2016, pp. 679–684, doi: 10.1109/IECBES.2016.7843536.
[20] Y. R. Tabar and U. Halici, "A novel deep learning approach for classification of EEG motor imagery signals," *Journal of Neural Engineering*, vol. 14, no. 1, 2017, doi: 10.1088/1741-2560/14/1/016003.
[21] P. Bashivan, I. Rish, M. Yeasin, and N. Codella, "Learning representations from EEG with deep recurrent-convolutional neural networks," in *ICLR 2016*, Nov. 2015, pp. 1–15.
[22] A. Craik, Y. He, and J. L. Contreras-Vidal, "Deep learning for electroencephalogram (EEG) classification tasks: a review," *Journal of Neural Engineering*, vol. 16, no. 3, Art no. 31001, Jun. 2019, doi: 10.1088/1741-2552/ab0ab5.
[23] Z. Mahmoodin, W. Mansor, K. Y. Lee, and A. Z. A. Zainuddin, "Electroencephalogram electrode localization in the support vector machine classification of dyslexic children," in *2016 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, Dec. 2016, pp. 296–300. doi: 10.1109/IECBES.2016.7843461.
[24] G. Brihadiswaran, D. Haputhanthri, S. Gunathilaka, D. Meedeniya, and S. Jayarathna, "EEG-based processing and classification methodologies for autism spectrum disorder: A review," *Journal of Computer Science*, vol. 15, no. 8, pp. 1161–1183, 2019, doi: 10.3844/jcssp.2019.1161.1183.
[25] A. Z. A. Zainuddin, W. Mansor, L. Y. Khuan, and Z. Mahmoodin, "Classification of EEG signal from capable dyslexic and normal children using KNN," *Advanced Science Letters*, vol. 24, no. 2, pp. 1402–1405, Feb. 2018, doi: 10.1166/asl.2018.10758.
[26] A. Z. A. Zainuddin, W. Mansor, K. Y. Lee, and Z. Mahmoodin, "Performance of support vector machine in classifying EEG signal of dyslexic children using RBF kernel," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 9, no. 2, p. 403, Feb. 2018, doi: 10.11591/ijeecs.v9.i2.pp403-409.
[27] A. Z. A. Zainuddin, W. Mansor, K. Y. Lee, and Z. Mahmoodin, "Performance of extreme learning machine kernels in classifying EEG signal pattern of dyslexic children in writing," *International Journal of Integrated Engineering*, vol. 11, no. 3, pp. 129–138, Sep. 2019, doi: 10.30880/ijie.2019.11.03.014.
[28] T. I. Alshwaheen, Y. W. Hau, N. Ass'Ad, and M. M. Abualsamen, "A novel and reliable framework of patient deterioration prediction in intensive care unit based on long short-term memory-recurrent neural network," *IEEE Access*, vol. 9, pp. 3894–3918, 2021, doi: 10.1109/ACCESS.2020.3047186.

BIORAPHIES OF AUTHORS

Ahmad Zuber Ahmad Zainuddin obtained a B.Sc. in Medical Electronics from the University of Hertfordshire, United Kingdom, M.Sc. in Electronics System Design Engineering from Universiti Sains Malaysia. He is currently a senior lecturer at the Medical Engineering Technology Section, Universiti Kuala Lumpur, Malaysia. His research interests include EEG, signal processing, machine learning and medical instrumentation. He can be contacted at email: zuber@unikl.edu.my.

Wahidah Mansor is a Director at Microwave Research Institute, Universiti Teknologi MARA (UiTM), Shah Alam, Selangor, Malaysia and a Professor at the School of Electrical Engineering, College of Engineering, UiTM, Malaysia. She obtained her B.Eng (Hons) in Electronic Engineering from the City of Birmingham Polytechnic, United Kingdom (UK) in 1989, Master of Science in Electronic Engineering from the University of Nottingham, UK in 1992 and PhD in Electronic Engineering from the University of Nottingham, UK in 2007. During her working period in UiTM, she has been appointed as Head of Programme, Deputy Dean (Academic), Research Coordinator and Strategic Planning Coordinator. Her main research interests include biomedical signal processing, brain-computer interfaces, embedded system design, computer engineering and wireless communication. She can be contacted at email: wahidah231@uitm.edu.my.

Khuan Yoot Lee received her PhD in EE Eng (Biomedical) from Nottingham University, UK in 2007. She is now a Prof at the School of EE Eng, College of Engineering, UiTM. She has authored more than 200 Scopus/WOS indexed technical publications in conferences and journals. Her Research Experience includes analysis, compression and remote display of biomedical signals with infrequent short duration events; Classifying neonatal cry to determine the physiologic status of baby (asphyxia and hypothyroid) with SVM, PSO, BPSO and MLP, Neurofeedback for dyslexia children. She can be contacted at email: leeyootkhuan@uitm.edu.my.

Zulkifli Mahmoodin is currently with Universiti Kuala Lumpur with more than 21 years of experience in the field of Biomedical Engineering. He obtained his PhD in Electrical Engineering from Universiti Teknologi MARA, Malaysia where he worked on EEG-based dyslexia studies. An Open-Source hardware advocate, he has won numerous awards in international innovation competitions and acted as a consultant for the development of several engineering institutions for Biomedical Engineering, Oil and Gas and Aircraft MROs. He can be contacted at email: zulkiflim@unikl.edu.my.