Comparison Between Support Vector Machine with Polynomial and RBF Kernels Performance in Recognizing EEG Signals of Dyslexic Children

A. Z. Ahmad Zainuddin, W. Mansor, Khuan Y. Lee, and Z. Mahmoodin

Abstract

Dyslexia is seen as learning disorder that causes learners having difficulties to recognize the word, be fluent in reading and to write accurately. This is characterized by a deficit in the region associated with learning pathways in the brain. Activities in this region can be investigated using electroencephalogram (EEG). In this work, Discrete Wavelet Transform (DWT) with Daubechies order of 2 (db2) based features extraction was applied to the EEG signal and the power is calculated. The differences between beta and theta band with responding to learning activities were explored. Multiclass Support Vector Machine (SVM) was used to classify the EEG signal. Performance comparison of Polynomial and Radial Basis Function (RBF) kernel recognizing EEG signal during writing word and non-word is presented in this paper. It was found that SVM with RBF kernel performance was generally higher than that of the polynomial kernel in recognizing normal, poor and capable dyslexic children. The SVM with RBF kernel produced 91% accuracy compared to the polynomial kernel.

Keywords

Dyslexia • Electroencephalogram • Support vector machine • RBF kernel • Polynomial kernel

A. Z. A. Zainuddin · W. Mansor (✉) · K. Y. Lee · Z. Mahmoodin
Faculty of Electrical Engineering, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia
e-mail: wahidah231@salam.uitm.edu.my

A. Z. A. Zainuddin · W. Mansor · K. Y. Lee · Z. Mahmoodin
Computational Intelligent Detection RIG, Pharmaceutical and Life Sciences CORE UiTM, 40450 Shah Alam, Selangor, Malaysia

A. Z. A. Zainuddin · Z. Mahmoodin
Medical Engineering Technology Section, Universiti Kuala Lumpur, 53100 Gombak, Selangor, Malaysia

1 Introduction

Dyslexia is a neurological disorder that causes learners having difficulties to decode a word, read and write despite receiving the adequate level of academic education [1]. Generally, the dyslexic children intelligent quotient (IQ) is normal or above average even though they have the problem to acquire smooth skill in reading and writing [2]. Schools in Malaysia screen children with dyslexia through an assessment that consists of measuring capability in spelling, reading, writing and as well as children strength and weakness in learning [3]. According to the report from Malaysia Ministry of Education, approximately 53,613 children enrolled the special program for learning disability in 2016 in which 8.35% expected to have dyslexia [4]. Another report shows that dyslexic children that enrol the intervention program have increased from 1,679 in 2014 to 10,329 in 2017 in which 5,806 is at primary level (age 7–12 years old) [5]. This number is increasing every year.

Looking into brain function, the cerebral cortex is the part of the brain that consists of four lobes which associated with a different function known as a frontal, temporal, parietal and occipital lobe. When an activity is carried out, the bio-electrical signal is generated in the area that related to its function which can be recorded using EEG. Compared to other imaging technique to identify dyslexia such as fMRI, PET and MEG [6], EEG has advantages as it can record higher temporal resolution of the signal where time and frequency domains of the signal are kept, is portable, easy to use, low cost, noninvasive and practical to be applied during learning activities [7].

A few studies have been conducted using EEG to determine area associates with brain functions such as sleep studies [8], epileptic [9], mental task, mental imaginary, motor imaginary, brain-computer interface [10] and learning disabilities [11]. This EEG signal is extracted to find good features for classification. Some of the features that can be

L. Lhotska et al. (eds.), *World Congress on Medical Physics and Biomedical Engineering 2018*,
IFMBE Proceedings 68/3, https://doi.org/10.1007/978-981-10-9023-3_17

extracted from EEG signal are power, skewness, variance, energy, entropy and standard deviation [12].

Dyslexia information in EEG signal can be obtained by extracting the features of the signal and then classified the signal using a suitable classifier. SVM is one of the well-known classifiers that can produce accurate results [13]. It is based on statistical learning theory and can work in small sample size, nonlinear and multiple classifications [14]. Choosing different kernel function of SVM may produce different performance [15]. Polynomial and RBF were widely used nonlinear kernel that projected data into infinite dimensional feature space [16]. The SVM performance using both kernels in classifying EEG of dyslexic children has not been reported.

This paper describes the classification of EEG signals of normal, poor and capable dyslexic children using SVM with Polynomial and RBF kernels. In this work, the performance of Polynomial and RBF kernel through writing known word and non-word is examined for suitability in identifying dyslexia.

2 Research Methodology

In this work, the classification of EEG signal of dyslexic children was carried out in several stages which include signal acquisition, subject identification and processing, features extraction and SVM classification using Polynomial and RBF Kernels.

2.1 Signal Acquisition, Subject Identification, and Processing

EEG signals were acquired using wireless bio-signal acquisition system called g.Nautilus with 8 active electrodes placed on the subject scalp in accordance with the International 10/20 System. These electrodes act as a sensor to pick up brain waves. Eight (8) electrode locations were chosen with reference to the areas associated with reading and writing pathways. At the left hemisphere of the brain, the electrodes are positioned at C3, P3, T7, and FC5 while at the right hemisphere of the brain, the electrodes are located at C4, P4, T8, and FC6.

There were four tasks carried out by each subject while EEG signal was recorded. The subject was asked to sit comfortably on a chair with a piece of paper and a pencil. A screen monitor was placed on a table in front of the subject. In the first task, the subject has to write 3 simple words and in the second task the subjects are required to write 3 complex words, these words are the words that have a specific meaning and can be understood. While in the third task, the subject has to write 3 simple non-words and in the fourth task, the subject must write 3 complex non-words. These non-words are the words that have no specific meaning. Each word and non-word was shown on the monitor screen one by one.

These sets of words were prepared according to age-appropriate academic level. Set A is for the subjects aged 7–8 which comprises 3 alphabets, set B is for the subjects aged 9–10 which contains 4–5 alphabets and set C is for the subjects aged 11–12 and have 5–8 alphabets. The choice of words and non-words were based on the assessment used by Dyslexia Association of Malaysia.

In this study, EEG data were recorded from 8 normal control subjects, 17 poor dyslexic children, and 8 capable dyslexic children. Normal control subjects are children from public school that can read and write smoothly. Poor dyslexic is referred to children that could not read and write correctly compared with normal control subject with the same age group level while capable dyslexic children refer to children that are able to read and write after they went through a dyslexia intervention program. The subject age was in the range of 7–12 years old since at this stage they start to receive formal learning activity at school where the symptom of dyslexia can be clearly seen from reading and writing. These subjects were first screened to identify the level of learning disorder which is poor or capable dyslexic with the assistance from Dyslexia Association of Malaysia and Rakan Dyslexia Malaysia group. During the assessment, physiological background, medical history, right and left hand dominant and IQ were recorded to ensure conformity of data.

EEG signals were recorded using g.Nautilus wireless biosignal acquisition system that has a built-in amplification and provides 24bit resolution with 500 Hz sampling rate. Noise embedded in the signal was removed using 2 types of filter. A notch filter was used to eliminate artifacts from power lines frequency at 50 Hz and a high pass filter with cut off frequency at 0.5 Hz was employed to remove noise from dc source. Once the artifacts were removed, features extraction was carried out.

2.2 Features Extraction

EEG signals are divided into five frequency bands known as delta δ (up to 4 Hz), theta θ (4–8 Hz), α alpha (8–13 Hz), beta β (13–30 Hz) and gamma γ (above 31 Hz). The delta is associated with deep sleep, theta is related to drowsiness, alpha indicates relaxed awareness, beta refers to the concentration or active attention and finally, gamma is simultaneous processing of information from different brain areas. Learning activities such as reading and writing, are mental activities which associated with the beta band frequency. While in theta band, the brain focusing is withdrawn.

Since EEG signal has non-stationary properties, time-frequency domain approaches using DWT was used for extracting the signal features. Daubechies of order 2 (db2) was employed to provide time-frequency scale representation due to its ability to localize features and provide smooth EEG signals [12]. Hence, db2 decomposes EEG signal into 5, however, in this work, only beta (13–30 Hz) and theta (4–8 Hz) bands were considered.

The power features were computed from reconstructed signal detail coefficient and the power was calculated from the sum of squared reconstructed signal values (x) divided by the signal length (L) as shown in using Eq. (1).

$$Power = \sum x^2/L(x) \tag{1}$$

The beta band power and the ratio of theta/beta band power are the two statistical feature vectors used as input to the classifier.

2.3 Classification

As mentioned previously, SVM with polynomial and RBF kernels were used to classify the three categories of EEG signals; normal, poor and capable dyslexic. SVM performs classification by finding maximum separation boundary by optimizing the spaces between two classes. In the linear case, a straightforward separation can be done using linear kernel but in nonlinear condition, the data need to be placed in features space where the separation is carried out in hyperspace. Nonlinear separation is accomplished by employing Radial Basis Function (RBF) and Polynomial kernel. Multiclass SVM with one versus one was employed in this work to classify normal, poor and capable dyslexic children. One versus one mechanism was carried out by separating each pair of classes against each other and using majority voting scheme to determine the output.

The SVM classifier equation used in the work is shown in (2).

$$f(x) = \sum_i^N \alpha_i y_i k(x_i, x) + b \tag{2}$$

where b is the bias, $k(x_i, x)$ is the kernel used in SVM, α_i is the weight vector, y_i is the target vector and N is the size of training data. While maximizing the margin of the data separation, the SVM minimizes the misclassification to zero. The trade-off between the misclassification and the margin is controlled by a parameter called box constraint. For the polynomial kernel, the order of polynomial kernel is determined by d as shown in Eq. (3). Here, the parameter d was set to 3. The RBF kernel projects vectors into an infinite dimensional space to compute the inner product between two projected vectors. The RBF equation used in the work is shown in Eq. (4) where the tuned parameter, σ that specifies the kernel width was set to 1. Both parameters were selected since it gives the lowest error from ten-fold cross-validation.

$$k(x_i, x) = (x_i.x + 1)^d \tag{3}$$

$$k(x_i, x) = exp\left(-\frac{||x_i - x||^2}{2\sigma^2}\right) \tag{4}$$

To select the optimum kernel, the box constraint was varied from 0.001 to 1000. The performance of each kernel was then evaluated and the accuracy, sensitivity, and specificity were determined using Eqs. (5), (6) and (7) respectively. Confusion matrix for multiclass was then employed to verify the performance of the classification models.

$$Accuracy, A_c = \frac{T_N + T_P}{T_P + T_N + F_P + F_N} \tag{5}$$

where T_N is the true negative, T_P is the true positive, F_P is the false positive and F_N is the false negative.

$$Sensitivity, S_e = T_{PR} = \frac{T_P}{T_P + F_N} \tag{6}$$

$$Specificity, S_p = T_{NR} = \frac{T_N}{T_N + F_P} \tag{7}$$

3 Results and Discussion

In this study, one dataset refers to total features obtained from a recording of EEG signals from 8 channels (C3, C4, P3, P4, FC5, FC6, T7 and T8) during performing a task. Since two features which are beta band power and theta/beta band ratio were computed for a task, one dataset gives 16 features. As each subject completes a total of 4 tasks, the accumulative dataset is 132 for 33 subjects. Therefore, the total data used is 2112. The datasets later were divided into 64% for training and 36% for testing. As mentioned

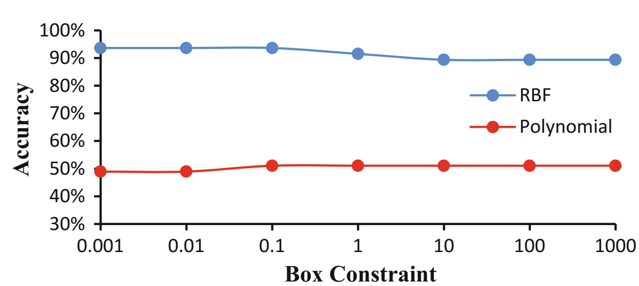

Fig. 1 Accuracy for SVM with Polynomial and RBF kernels when box constraint is varied

Table 1 Classification Performance of SVM with both kernels using box constraint = 1

Kernel	Type	Sensitivity	Specificity	Accuracy
Polynomial	Normal	1.00	0.67	
	Poor Dyslexic	0.44	1.00	0.51
	Capable Dyslexic	0.75	0.79	
RBF	Normal	1.00	0.95	
	Poor Dyslexic	0.92	0.88	0.91
	Capable Dyslexic	0.75	0.98	

previously, the optimum parameter for RBF and polynomial kernel of SVM were selected using K-Fold cross-validation.

Figure 1 shows the accuracy of SVM in identifying normal, poor and capable dyslexic when box constraint is varied. The results show that RBF kernel provides high accuracy (94%) when the box constraint is between 0.001 and 0.1 whereas the polynomial kernel maintains high accuracy (51%) when the box constraint is in the range of 0.1–1000.

Table 1 shows the classification performance of SVM with polynomial and RBF kernels. The SVM with polynomial kernel provides the highest sensitivity when classifying the normal subjects and have the highest specificity when recognizing poor dyslexic children. It is also found that using the polynomial kernel, the SVM provides an accuracy of 51% in classifying the normal, poor and capable dyslexic children.

It can be seen that the SVM with RBF kernel gives good performance when classifying EEG signals of normal, poor and capable dyslexic children. It provides 91% accuracy in classifying all subjects. The highest sensitivity which is 100% is obtained when classifying the normal subjects and the highest specificity (98%) is achieved when distinguishing the capable dyslexic. Comparing the performance of these two types of kernel at box constraint is 1, it is obvious that the RBF kernel is the most accurate kernel since it produces the highest classification accuracy which is 91% whereas the polynomial kernel only gives 51%. The RBF kernel performs better than the polynomial kernel since it uses Gaussian curve with infinite dimensionality in separating data points which offers more predictive efficiency.

4 Conclusion

The performance of SVM with polynomial and RBF kernels in recognizing EEG signals of dyslexic children has been described in this paper. The sensitivity, specificity, and accuracy of each kernel were determined to select the optimum kernel. It was found that the SVM with RBF kernel performance is much better than that of polynomial kernel since it produces an accuracy of 91% in classifying all subjects. The SVM with polynomial kernel was unable to identify poor dyslexic correctly compared to normal and capable dyslexic. Therefore, the SVM with RBF Kernel is proposed to be used in recognizing EEG signals of normal, poor and capable dyslexic.

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 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.

Conflict of Interest The authors declare that they have no conflict of interest.

References

1. E. S. Norton, S. D. Beach, and J. DE Gabrieli, "Neurobiology of dyslexia," *Curr. Opin. Neurobiol.*, vol. 30, pp. 73–78, Feb. 2015.
2. U. Goswami, "Dyslexia, Developmental," in *International Encyclopedia of the Social & Behavioral Sciences*, vol. 6, Elsevier, 2015, pp. 727–730.
3. Ministry of Education, "Instrumen Senarai Semak Disleksia," 2011.
4. B. P. K. K. P. Malaysia, "Data Pendidikan Khas 2016," 2016.
5. Z. Mahfuzah, "Statistik Murid Disleksia di Malaysia," 2017. [Online]. Available: https://www.mahfuzahzainol.com/single-post/2017/12/06/Statistik-Murid-Disleksia-di-Malaysia. [Accessed: 19-Jan-2018].
6. Y. Sun, J. Lee, and R. Kirby, "Brain Imaging Findings in Dyslexia," *Pediatr. Neonatol.*, vol. 51, no. 2, pp. 89–96, Apr. 2010.
7. S. Mohamad, W. Mansor, L. Y. Khuan, C. W. N. F. C. W. Fadzal, N. Mohammad, and S. Amirin, "Development of computer-based assessment for brain electrophysiology technique of dyslexic children," in *2016 IEEE Symposium on Computer Applications & Industrial Electronics (ISCAIE)*, 2016, pp. 79–83.
8. S. Qureshi and S. Vanichayobon, "Evaluate different machine learning techniques for classifying sleep stages on single-channel EEG," in *2017 14th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, 2017, pp. 1–6.
9. S. Siuly and Y. Li, "Designing a robust feature extraction method based on optimum allocation and principal component analysis for epileptic EEG signal classification," *Comput. Methods Programs Biomed.*, vol. 119, no. 1, pp. 29–42, 2015.

10. 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.
11. D. C. Hammond, “What is Neurofeedback: An Update,” *J. Neurother.*, vol. 15, no. 4, pp. 305–336, 2011.
12. 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.
13. X. Liu, C. Gao, and P. Li, “A comparative analysis of support vector machines and extreme learning machines,” *Neural Networks*, vol. 33, pp. 58–66, 2012.
14. Y. Ma, X. Ding, Q. She, Z. Luo, T. Potter, and Y. Zhang, “Classification of Motor Imagery EEG Signals with Support Vector Machines and Particle Swarm Optimization,” *Comput. Math. Methods Med.*, vol. 2016, no. 5, pp. 667–677, 2016.
15. E. A. Zanaty, “Support Vector Machines (SVMs) versus Multilayer Perception (MLP) in data classification,” *Egypt. Informatics J.*, vol. 13, no. 3, pp. 177–183, 2012.
16. C. K. I. Williams, “Learning With Kernels: Support Vector Machines, Regularization, Optimization, and Beyond,” *J. Am. Stat. Assoc.*, vol. 98, no. 462, pp. 489–489, Jun. 2003.