DYSLEXIA DETECTION FROM EYE MOVEMENTS USING ARTIFICIAL NEURAL NETWORKS

M. Macaš*, L. Lhotská*, D. Novák*, M. Vyhnálek** and R. Brzezny**

* Czech Technical University in Prague/Gerstner Laboratory, Technicka 2, Prague, Czech Republic
** 2nd Medical Faculty of the Charles University/Department of Neurology, V úvalu 84, Prague, Czech Republic

mmacas@seznam.cz

Abstract: The main goal of the study was to propose and implement a neural network based classifier for dyslexia detection from eye movement signal. Eye movements of 76 school children were measured using a videooculographic (VOG) technique, during one reading and four non-reading tasks. Time and frequency domain features were extracted and various feature selection methods were applied to select subsets of significant features. Finally a feed-forward neural network using back-propagation algorithm was used for a supervised learning. A suitable topology was chosen and learning parameters were set experimentally. A number of experiments were performed with different subsets of features. The final classifier reached about 90% correct identification of the presence of dyslexia and about 90.5% correct identification of the absence of dyslexia. The method described here may also serve as an impulse for new research focusing on diagnostic significance of eye movements.

Introduction

Dyslexia (specific reading disability) is a common, cognitively and behaviourally heterogeneous developmental condition, characterized primarily by severe difficulty in the mastery of reading despite average intelligence and adequate education [1]. Dyslexia is one of the specific learning disabilities. Its prevalence in Czech school population is estimated to 2-3% [2], which is considerably less than in English speaking countries (5-17,5%) [3]. The children with dyslexia need special educational methods for acquiring reading skills, that is why the early diagnostic is very important [4]. An early dyslexia detection will give those people the chance of being treated by means of the most accurate and specialized intensive therapy.

This paper focuses on the diagnostic problem – searching for a relationship between eye movements and dyslexia. The problem has been described in several studies. However, these studies do not provide any clear unified conclusions. The approach proposed here uses artificial intelligence methods in order to contribute to better problem understanding. An artificial neural network computation model has been used for an eye movement signal classification.

The human eye is never in entire calm. Its movement is a consequence of an ophthalmogyric muscle's work. There are two main kinds of eye movements - the big and the small. Saccades are rapid eye movements which allow binocular turning or version of the eyes from one fixation point to another. During these conjugated and volitional movements, the eye is browsing a visual field. The direction and the magnitude of the saccade are voluntary. Each saccade has its direction. People usually read from left to right and most of saccadic eye movements are oriented accordingly. These normal reading movements are called forward saccades. Reading movements going from right to left are called regressions.

The saccade alternates with the period of fixation made when the eyes are directed to a particular target. The fixation consists of three kinds of small eye movements - a drift, microsaccades and a tremor. Sequences of fixations and saccades (rapid eye movements between fixations) define scanpaths, providing a record of visual attention on a subject of interest.

Although the connection between dyslexia and eye movements has been examined in several studies, it is still not clear, if this connection exists. The functional significance of eye movements is far from being understood. Few studies carried out comparison between eye movements of dyslexics and matched control subjects.

Many properties of dyslexic's eye movements were described, adverting to diagnostic potential of eye movements. Subjects with reading difficulties make a higher percentage of regressive eye movements than normal readers [5].

In another study [6] authors found, that children with learning disabilities made significantly more fixations and regressions than normal readers. But there were no significant differences in their fixation's duration. The similar results were obtained in [7], where analysis showed, that the number of forward and regressive eye movements was significantly higher for dyslexic than for retarded and normal readers. Differences were not only in absolute number of regressions, but also in the percentage of regressions of the total number of eye movements. Another study [8] aimed for using techniques of nonlinear dynamical systems to assess

ISSN: 1727-1983

visual dysfunction. Authors used power spectral density or fractal dimension.

Materials and Methods

The eye movements of 76 female subjects were recorded using iView 3.0 videooculography system at Department of Neurology, 2nd Medical Faculty, Charles University, Czech Republic. Only 52 of the 76 measured subjects were used for subsequent experiments because of difference in age or poor quality of signal record. There were three types of subjects – normal readers (N=22, mean age was 11.5 years), retarded readers without dyslexia (N=20, mean age was 11.46 years) and dyslexics (N=10, mean age was 11.72 years). The measurements have been executed in dark. Subject with fixated head looked at a stimulus on the screen. The screen has been placed 1 meter from the subject. Subjects were stimulated by four non-verbal and one verbal stimuli consisting of bitmap pictures (resolution 1024x768). Each stimulus corresponded to one task. Nonverbal tasks included e.g. browsing through dots, taking a view of a picture or looking through numbers placed randomly on monitor. The only verbal task was reading a text.

Before feature extraction itself, the signals were preprocessed and analyzed using special software for analysis of eye movement signal [9]. First, the calibration of mapping on visual stimuli was carried out using projective geometric transformation. After calibration, the noise was filtered out of the signals by means of convolution filter with Gaussian kernel. Finally, the blinking artefacts were removed. The pre-processed signals were then analyzed.

The analysis included automatic detection and description of fixations and saccades. Next, the data set of 300 features was created from the pre-processed and analyzed horizontal and vertical VOG signals over all five stimuli. The implementation of feature extraction was based on extraction scripts written in Matlab 6.1. The extraction scripts were proposed for automatic extraction of features from eye movement analysis files produced by the eye movement analysis software. Several feature types were extracted. Data were then individually and manually checked for usability. Abnormal, suspicious (big or too different) and missed values were replaced by the mean of the corresponding variable over those patterns for which its value was available. After the extraction, 300 feature variables from time and frequency domains were gathered (100 features from each domain). Only the most important features are shortly described in this paper.

The first group of features can be called *Fixation Types Ratios* and contains the ratios of directions of eye movement from one fixation to the next. The proportion of each categorised direction and the transition matrix of each two successive categorised directions are then calculated. The categorised direction West-East, for example, means that the eye entered from the previous fixation area to the actual one from the East direction and then the eye exited from the actual fixation area to the next one in the West direction. These measures were proposed as objective indicators of the strategies employed by an observer engaged in a visual search task [10].

Different, but very similar group of features is called *Relative Occurrence of Directions of Saccades.* Eight directions were used - east, south-east, south, etc. The usability of this vector lies in short and apposite representation of directionality of saccades.

The frequency domain features are inspired by a power pattern concept described in [11]. The power spectrum represents the distribution of power in the space of frequencies. To reduce the size of feature set, first, the entire frequency space is divided into F sections, called "frames”. Next the power for each frame is calculated as the sum of squared Fourier coefficients for frequencies within the corresponding frame. The power pattern is the vector of numbers expressing the powers in particular frames. The width of the frame was 1Hz in and the frames for frequencies between 1Hz and 20Hz have been used in our experiments.

Many other features have been extracted, however results presented here have not used them (the selection procedure eliminated them). For example, there have been also extracted characteristics as the mean and the standard deviation of fixation duration, time of completing the task, number of fixations and saccades, etc.

Next, data have been normalized and the dyslexic subjects were replicated four times and a random vector was added onto each copy. It balances class representation and can reduce over-fitting. The resulting data-set has 82 patterns with 300 feature variables and was prepared for feature selection procedure.

The feature selection has been performed to find subsets of most significant features usable for classification. Suitable choice of inputs can significantly influence the classification performance of the neural network. The quantity of training data needed to specify the mapping grows exponentially with the dimensionality of the input space. Reducing the number of input variables can sometimes lead to improved performance for a given data set, even though some information is being discarded. Selection criterion enables to compare different feature subsets and determine the best one. In our case, we use very simple criterion based on class separability - Euclidean distance between class means.

The second part of feature selection problem is searching strategy. We use a systematic searching procedure, which searches through candidate subsets of features. We use backward elimination [12] as the searching algorithm to find subsets of 5 - 8 input variables. The backward elimination method starts with the full set of features and at each stage, one feature is removed from the set, which gives rise to smallest decrease in the criterion value.

As described above, there are three groups of subjects in examined population. However, the goal is to classify only dyslexic subjects. Therefore we understand the problem as dichotomia (two classes - dyslexic and non-dyslexic subjects). The feed-forward network is used for classification experiments considering only two classes.

The back-propagation network is one of the most frequently used networks in many applications of artificial neural networks [13]. It is a net with feed-forward topology using error back-propagation to learn. Error back-propagation learning rule is central to much current works on learning in artificial neural network.

The algorithm of back-propagation provides a computationally efficient method for changing the weights in a feed-forward network to learn a training set of input-output examples. The application of this algorithm has two phases. In the first phase, input is propagated forward to the output units where the error of the network is measured (forward propagation). In the second phase, the error is propagated backward through the network and is used for adapting connections. In this study, one hidden layer network is used with 4-7 inputs and one output unit with saturating linear activation function. The number of hidden units (hyperbolic tangent sigmoid activation function) has been determined experimentally in particular cases. The back-propagation algorithm with momentum and adaptive learning rate has been used for learning. Momentum constant has been 0.9.

Generally, a network with an excess of free parameters (weights) tends to generate mapping, which has a lot of curvatures and performs over-fitting [12]. One approach to optimizing the generalization performance is to control its effective complexity. It can be done by adding a complexity term Ω to the error function E that penalizes large weights. The new error function $\hat{E}$ is $\hat{E}=E+v\Omega$, where v is a regularization parameter that controls the extent to which the penalty term influences the form of solution. A simple form of the regularization term is:

$$\Omega = \frac{1}{W}\sum_{i=1}^{W} w_i^2 \tag{1}$$

The classification performance cannot be evaluated using a test set, because of small amount of data and splitting whole data set in training and testing subsets could waste the data. Therefore a more suitable cross-validation method [14] has been used. The cross-validation is markedly superior for small data sets. Cross-validation consists of dividing the data into m subsets. In each step, one subset is used as the testing set and the rest of the original data is used for network's training. The estimated error rate is the average error rate (expressed in percents of miss-classifications) from these subsets. In this work, leave-one-out cross-validation where m equals the sample size is used to estimate the classification performance error. Together with the cross-validation error, diagnostic accuracy measurements are used - sensitivity and specificity [15]. True positive is the situation when the test is positive for a subject with dyslexia. The true-positive fraction known as sensitivity represents the percentage of correctly classified dyslexic subjects. True negative is the situation when the test is negative for a subject without dyslexia. The true-negative fraction known as specificity indicates accuracy in identifying non-dyslexic subjects and represents the percentage of correctly classified non-dyslexic subjects.

Results

Many experiments have been performed, however only some of the results are described in this paper. The set of experiments presented here contains 6 basic experiments.

Table 1: Summary of results - cross-validation error, sensitivity and specificity values for the best results of 10 runs.

Input data	$\mathbf{E_{best}}$ **[%]**	$\mathbf{SE_{best}}$ **[%]**	$\mathbf{SP_{best}}$ **[%]**
Time domain	30.5	62.5	76.2
Horizontal frequency domain	26.8	75.0	71.4
Vertical frequency domain	15.9	77.5	90.5
All domains	18.3	72.5	90.4
Verbal task only	14.0	80.0	93.0
Verbal task only +regularization	9.7	90.0	90.5

For each experiment, the classification performance has been estimated using cross-validation error, sensitivity and specificity. For each experiment setting, 10 runs have been realised, the best and the average result have been used for classifier's performance evaluation. The best results are listed in the result summary in Table 1 and are used for training curves in Figure1 and Figure 2. The average results are presented in Figure 2 and Figure3.

The first three experiments have used only features from particular domains. The results for time domain features have been quite poor as well as for features extracted from frequency domain of horizontal signal. However, the results for features extracted from frequency domain of vertical signals have been promising. The performances for different training times are depicted in Figure 1. The minimum of error is 16% of misclassifications. The specificity of the best result is 90% and the sensitivity reaches 78%.

In the fourth experiment we have used the whole set of 300 features extracted from all domains, from which a subset has been selected and used for classification.

The result has been poor which can be explained by too big feature set used for selection.

The best results have been achieved for the fifth experiment, where purely features corresponding to verbal (reading) stimulus have been used.

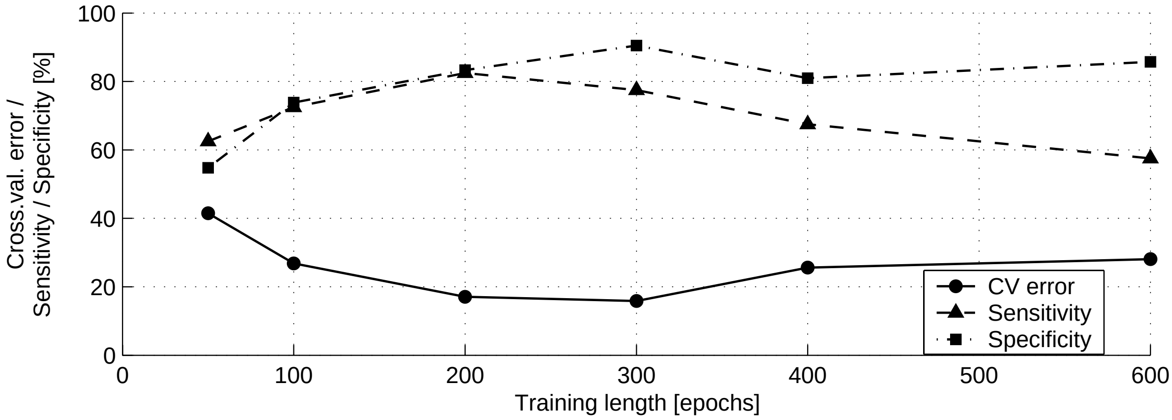

Figure 1: The results of classification experiments using features extracted from frequency domain of vertical signal. The back-propagation network with 5 hidden neurons is used (determined experimentally).

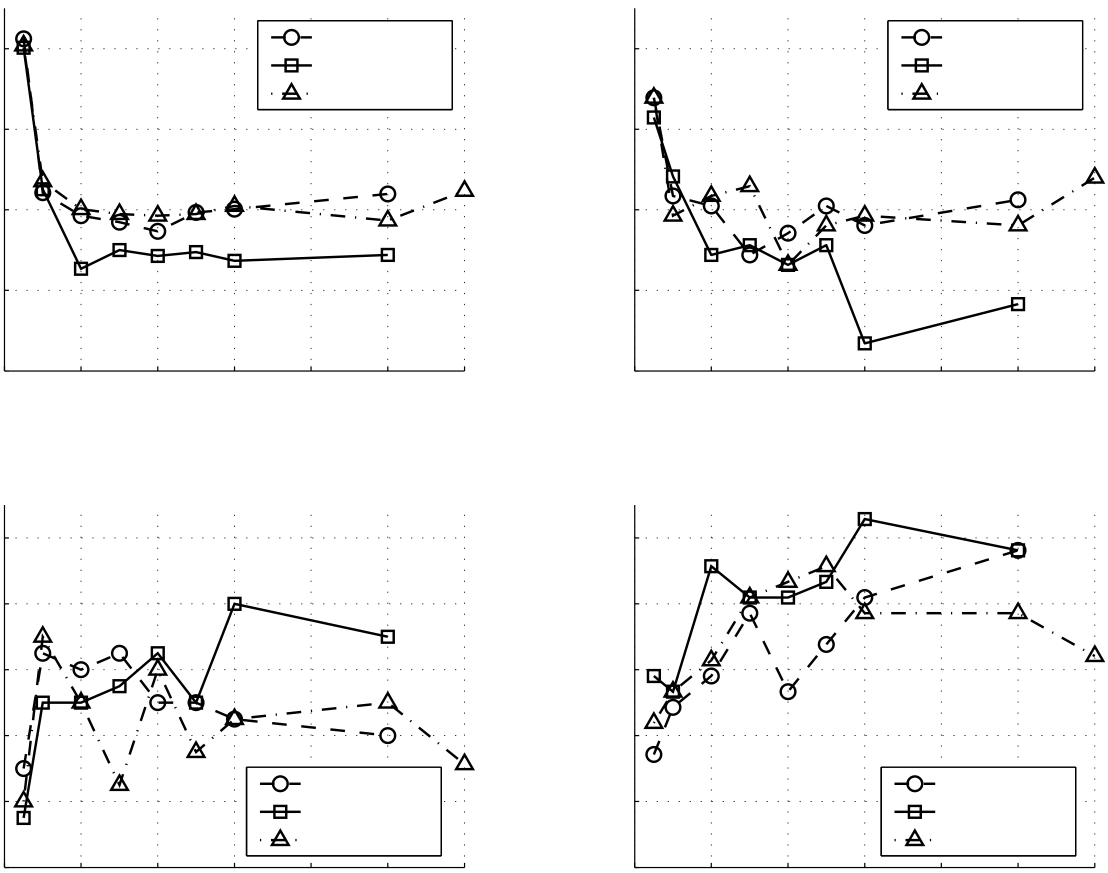

Figure 2: The results of classification experiments using features extracted from verbal task. The back-propagation network with 9 hidden neurons is used.

ISSN: 1727-1983 © 2005 IFMBE

The results are presented in Figure 2. Three subsets with different numbers of features (four, five and six) have been selected from 58 features corresponding to the verbal task.

We have used here the average (the mean of cross-validation errors from 10 runs of experiment) and the best result for comparison. It is obviously evident that the network with five inputs has much better classification performance (for our feature subsets) than the other networks. The minimal error (best result) reaches 14% of misclassifications and the average of 10 runs has its minimum at about 22%. The sensitivity reaches 80% and the specificity reaches even 93%. It points out that the verbal task is perhaps very important for our method of classification.

Finally, the regularization method described above has been used for result improving. Only the last result, features extracted from verbal task, has been improved. The results have been calculated for nine values of the regularization parameter *ν*. The dependence of the average minimal error (the mean of minimal errors from 10 runs of experiment) on the regularization parameter is shown in Figure 3.

A minimum can be seen there for *ν =0.09*, which is the optimal choice of the regularization parameter *ν*. The best error reached in this experiment is 9.8% of misclassifications (only the average result, not the best one is shown in Figure 3), which corresponds to 90.5% sensitivity and 90.0% specificity.

Discussion

A back-propagation neural network based classification system for dyslexia detection has been

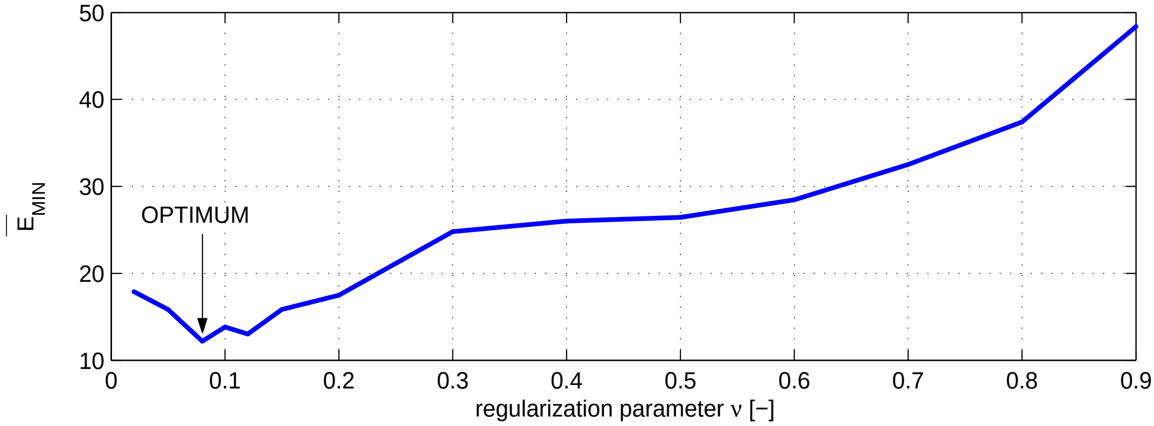

Figure 3: A dependence of classification performance on the regularization parameter. The performance $\bar{E}_{min}$ was measured as the mean of minimal errors obtained in 10 runs.

proposed in this work. A set of many features from time and frequency domain has been extracted and features have been selected. During classification experiments it has been found, that eye movement signal has shown itself as significant for dyslexia detection.

The best result has been obtained for feature subset selected from the set corresponding purely to verbal task. It corresponds to expectation that dyslexia is detectable mainly from reading tests.

However, in the case of the classifier using vertical frequency domain features, the four of the five most significant features have been extracted from the non-verbal tests.

The classifier with best classification capabilities (using features from verbal task only) has been further improved. The generalization has also markedly improved using regularization method. The final classifier reaches 90% selectivity and 90.5% specificity. Its topology is feed-forward neural network with five inputs, nine neurons in one hidden unit and one output neuron. It uses back-propagation learning strategy with momentum (0.9), adaptive learning rate and regularization term (of weight decay type) added to mean squared error.

All results are greatly affected by lack of measured data. The size of the training set, relative to the number of features determines the accuracy and reliability of classification system.

Conclusions

Various possibilities for obtaining better accomplishments can be proposed. The main aim now is to get larger testing population. Especially, more dyslexic patients should be measured to get more reliable results. For new measurements, additional tests can be proposed using not only static, but also dynamical stimuli (e.g. a moving light).

The feature extraction procedure can be improved as well. We propose to use other standard methods for feature extraction (wavelet transform, autoregressive methods, etc.) as a future direction.

The method described here may also serve as an impulse for new researches focusing on diagnostic significance of eye movements, e.g. in the diagnostics of other brain dysfunction, balance disorders, schizophrenia, paralytic strabism and autism.

Acknowledgement

The research is supported by the research programme No. MSM 6840770012 “Transdisciplinary Research in the Field of Biomedical Engineering II”.

References

[1] GRIGORENKO E. L. (2001): ‘Developmental Dyslexia: An Update on Genes, Brains, and environments’, *J Child Psychol Psychiatry*, **42**, pp. 91-125

[2] MATĚJČEK, Z. (1995): ‘Dyslexia – specific reading disabilities’, H&H (in Czech)

[3] HABIB, M. (2000): ‘The Neurological Basis of Developmental Dyslexia: An Overview and Working Hypothesis’, *Brain*, **123**, pp. 2373-2399

[4] SHAYWITZ, S. (1998): ‘Dyslexia’, *N Engl J Med*, **338**, pp. 307-312

[5] GILBERT, C. L. (1953): ‘Functional motor efficiency of the eyes and its relation to reading’, University of California Publications in Educations, pp. 159-231

[6] RUBINO A., MINDEN A. H. (1973): ‘An Analysis of Eye Movements in Children with Reading Disability’, *Cortex*, **9**, pp. 217-220

[7] PAVLIDIS G. (1985): ‘Eye Movements in Dyslexia – Their Diagnostic Significance’, *Journal of Learning Disabilities*, **18**, pp. 42-49

[8] HUTZLER F., WIMMER H. (2004): ‘Eye Movements of Dyslexic Children when Reading in Regular Orthography’, *Brain and Lang.*, **89**, pp. 235-242

[9] SNOPEK J. (2003): ‘Methods of Eye Movement Record Analysis at Reading and Sequential Tasks’, Master Thesis – Dept. of Cybernetics, Czech Technical University in Prague, Czech Republic (in Czech)

[10] PONSODA V., Scott D., Findlay J. (1995): ‘A Probability Vector and Transition Matrix Analysis of Eye Movements during Visual Searching’ *Psychologica,* **88**, pp. 167-185

[11] TERANISHI M., OMATU S., KOSAKA T. (2002): ‘Neuro-classification of Bill Fatigue Levels Based on Acoustic Wavelet Components’, Proc. of ICANN 2002, p. 1074-1079

[12] BISHOP, C. M. (1995): ‘Neural Networks for Pattern Recognition’, Oxford University Press Inc., New York

[13] HASSOUN, M. H. (1995): ‘Fundamentals of Artificial Neural Networks’, Massachusetts Institute of Technology

[14] WEISS S. M., KULIKOWSKI C. A. (1991):’Computer Systems That Learn’, Morgan Kaufmann

[15] RANGAYYAN R. M. (2002):’Biomedical Signal Analysis – A Case Study Approach’, IEEE Press