

Comparison of Parametric and Non-parametric EEG Feature Extraction Methods in Detection of Pediatric Migraine without Aura

Kazemi S.¹, Katibeh P.^{2*}

ABSTRACT

Background: Migraine headache without aura is the most common type of migraine especially among pediatric patients. It has always been a great challenge of migraine diagnosis using quantitative electroencephalography measurements through feature classification. It has been proven that different feature extraction and classification methods vary in terms of performance regarding detection and diagnostic accuracy. Previous work on the subject was controversial, hence a comparison of these methods seems necessary.

Objective: The aim of this research is to compare two parametric and non-parametric feature extraction methods and also two classification methods in order to obtain optimal combinations of diagnostic accuracy.

Materials and Methods: Having recorded background EEG from 24 pediatric migraineurs and 19 control subjects, data was processed by Welch's and Yule-Walker's methods. Features were selected using genetic algorithm, and then given to a support vector machine and the linear discriminant analysis for the classification. Accuracy was calculated for all combinations having the dominant frequency and the correlated absolute power of each EEG wave band (theta, alpha, and beta) and for all wave bands combined.

Results: The highest migraine detection accuracy of 93% was obtained utilizing Welch's method for EEG feature extraction alongside support vector machine for a classifier. Besides, Yule-Walker autoregressive method showed better performance than Welch's, when only power bands (and not the dominant frequency) were used as classification input.

Conclusion: The superiority of Welch's method over Yule-Walker's and the support vector machine over linear discriminant analysis can be great help for further researches on computer aided EEG-based diagnosis of migraine.

Keywords

Migraine, Quantitative Electroencephalography, Welch, Yule-Walker Autoregressive Method, Support Vector Machine, Linear Discriminant Analysis

Introduction

Migraine headache is a significant subset of primary headaches, known as the most common type of pediatric recurrent headaches, and it is presented with intermittent episodes of moderate to severe headache characterized by the duration of the pain which is between 4 to 72 hours. Migraine headache without aura is more common than the one with aura; it is mostly diagnosed based on ICHD2 and ICHD3 criteria of headache disorders [1,2] and it is about 60% to 85%

¹School of Medicine, Shiraz University of Medical Sciences, Shiraz, Iran

²Pediatric Neurologist Assistant professor of Shiraz University of medical science, Shiraz, Iran

*Corresponding author:
P. Katibeh
Pediatric Neurologist
Assistant professor
of Shiraz University of
medical science, Shiraz,
Iran
E-mail: pegahkatibeh@
yahoo.com

Received: 22 July 2018
Accepted: 18 August 2018

of all migraineurs [3,4].

Electroencephalography (EEG) is an electrophysiological method of medical diagnosis and basic research recording discharges of superficial layers of cortex finding their way through the skull and radiating on the outside. Various EEG findings are correlated with the different types of headaches, most of which lack specificity, such as diffuse slowing, and increased slowing of high amplitude waves during photic stimulation [5].

The use of different EEG feature extraction methods such as various wavelet functions and Burg-AR method and classifiers like support vector machine, for migraine detection is not a new subject [6,7], although the comparison of different feature extraction methods and the efficacy of different classification methods is a novelty for migraine detection, which can absolutely aid migraine computer-aided diagnosis and follow up.

In this research, EEG signals were analyzed, using Welch's technique as a non-parametric feature extraction method and Yule Walker autoregressive technique as a parametric feature extraction method (knowing that, autoregressive techniques necessitate having signal dependent parameters such as 'order', one which is unnecessary while using non-parametric methods). The extracted features were classified using two classification methods 1. support vector machines and 2. the linear discriminant analysis. Lastly, classification results of all 4 combinations were compared to obtain best combination of migraine detection accuracy.

Materials and Methods

Data acquisition

EEG recording data is acquired from 24 children (8 males and 16 females) aged between 8 and 18 years old (12.7 ± 3.12 years, mean \pm standard deviation) diagnosed with migraine without aura, based on ICHD3 criteria of headache disorders, and 19 healthy subjects

(7 males and 12 females) in terms of migraine or any other headache disorder, also aged between 8 and 18 years old (12.6 ± 3.18 years, mean \pm standard deviation). Recording is done for 10 minutes of background EEG with eyes closed, in a quiet room having the minimum noise as possible, by an EEG device with sampling rate at 256 Hz and 19 electrodes according to standard 10-20 system. No particular segments are deleted or included by visual inspection. Pre-processing was performed using frequency bandpass filters for 0.5 to 30 Hz to delete the noise above 30 Hz including muscle contractions and electricity at 50 Hz. All analyses were performed in MATLAB 2014b.

Frequency Bands

After pre-processing, all signals from each electrode were filtered to obtain (4 to 7 Hz), alpha (9 to 13 Hz), and beta (14 to 30 Hz) frequency bands for further processing.

Feature Extraction

In this paper, Welch's method and Yule-Walker Autoregressive (AR) method of feature extraction are utilized to obtain power spectra.

Welch's method, as a non-parametric method of feature extraction, divides time into segments, calculating the power spectrum of each segment, and averages power spectra among time series segments. Length of segments (windows) were 1000, and number of FFT points were 1024 (commonly the next power of 2 from the window length).

Yule-Walker method of feature extraction, as a parametric method is an autoregressive model for which Akaike Information Criterion is utilized to calculate the 'order' (as a parameter for autoregressive models) using MATLAB function ARfit [8].

In our case, calculated orders for theta, alpha, and beta bands are 7, 10, and 12 respectively. Window length and number of FFT points are the same as above.

Using both feature extraction methods men-

tioned above, output variables include dominant frequency of each band (in Hz) and its absolute correlated power (in μV^2) which are used as inputs for classification.

Feature Selection

Since the number of features resulted by EEG processing is considerable, the feature selection is a necessary step to reduce the amount of information process and prepare the data for the accurate result. Therefore, in this article, a binary genetic algorithm is applied to obtain optimized features for classification and accuracy calculation. Genetic algorithm or GA (with binary genetic algorithm as its simplest form of use) is one of the most common feature selection tools used for EEG feature selection because it does not use one-by-one features and it is able to find different combinations of best features simultaneously. Therefore, in this article, a binary genetic algorithm is applied with the mutation parameter at 0.1, cross-over parameter at 0.8 and population size at 50 so as to obtain optimized features for classification and accuracy calculation [9].

Classification

Here, two methods of data classifications are used: support vector machines (SVM) and linear discriminant analysis (LDA) by MATLAB

functions `fitsvm` and `fitdiscr`, respectively.

Support vector machines (SVM) using hyperplane separation are useful tools to classify large data with a large number of predictors and have a high accuracy to split classes [10].

Linear discriminant analysis (LDA) is a statistical method commonly used for data classification which is based on linear combinations of features [11].

For both methods, accuracy was calculated using leave-one-out cross validation. Selected features and electrodes by GA were used as inputs for evaluation of dominant frequency and correlated absolute power (separately and combined) utilizing Welch's method and Yule-Walker AR method classified by both SVM and LDA tools. This was also done separately for theta, alpha, and beta frequency bands.

Results

The accuracy calculated for combined frequency and power selected features in each wave band using Welch's and Yule-Walker AR feature extraction methods is classified by SVM using the leave-one-out cross validation demonstrating higher accuracy using Welch's method especially for alpha band features (combined dominant frequency and correlated absolute power). The mentioned accuracy is 81% as shown in Figure 1.

Figure 2 demonstrates the same as Figure 1,

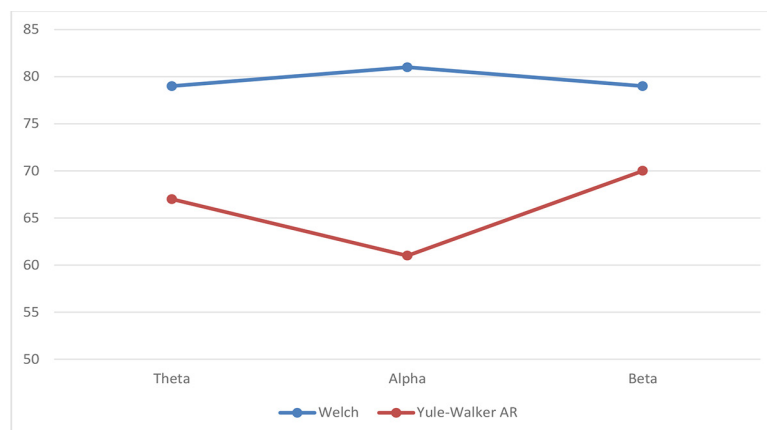


Figure 1: Diagnostic accuracy using combined frequency and power of wave bands using SVM. Alpha band extracted by Welch's method has the highest accuracy of 81% and Welch's results show better performance than Yule-Walker AR.

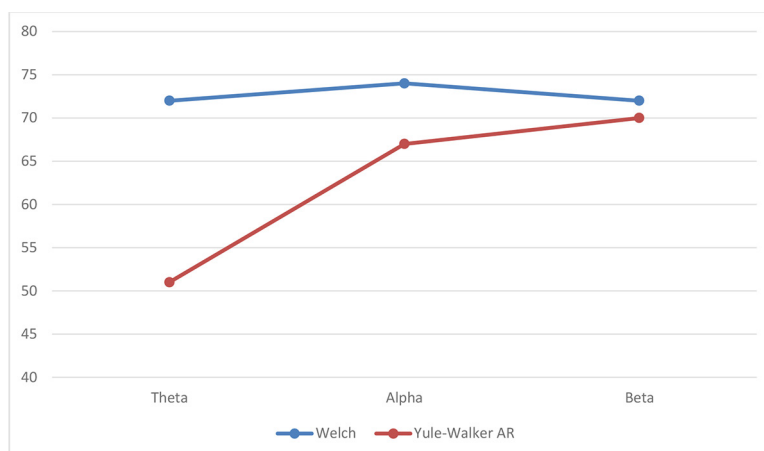


Figure 2: Diagnostic accuracy using combined frequency and power of wave bands using LDA. In total results are lower than SVM results although Welch’s results still remain the better one than Yule-Walker AR.

except for the classification using LDA method and the results show that accuracy values are larger using SVM compared to LDA. Overall, Maximum accuracy is 74% which is resulted by alpha band selected features using SVM.

As shown in bar charts, Figure 3 is dedicated exclusively to dominant frequency for all wave bands (combination selection by GA) as the only input of classification. This shows that Welch’s method results in higher values of detection accuracy especially when features are classified by SVM. It has 88% accuracy

as the highest detection accuracy for the mentioned combination.

Combined absolute power for all wave bands as the only inputs for classification reveals higher accuracy values when Yule-Walker AR is used as a feature extraction method especially with SVM as a classifier ensuing from a maximum detection accuracy of 86% that it is shown in Figure 4.

Combined dominant frequency and correlated absolute power selected features for all wave bands reveal the highest detection ac-

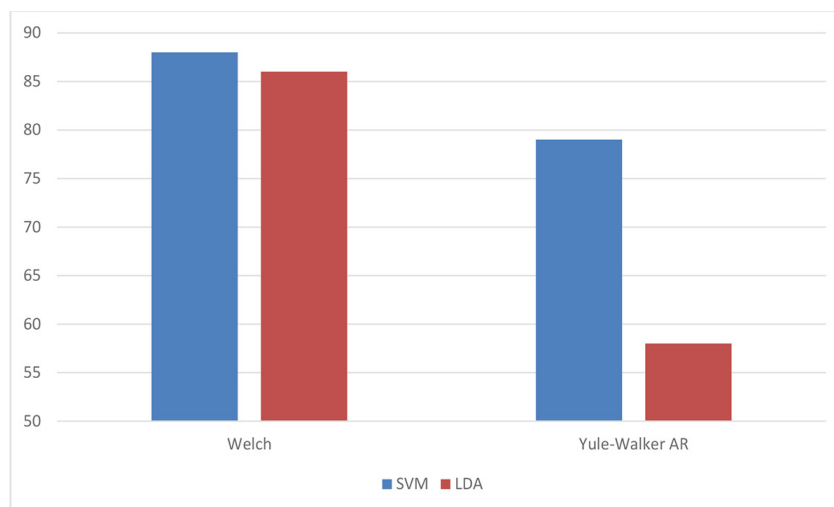


Figure 3: Diagnostic accuracy using exclusively dominant frequency of wave bands, combined. Welch’s method shows better performance with 88% accuracy using SVM.



Figure 4: Diagnostic accuracy using exclusively absolute power of wave bands, combined. Yule-Walker shows better performance especially using SVM with 86% accuracy.

curacy values. As shown in bar charts of Figure 5, data analysis adopting Welch’s method as the feature extraction method using SVM classification results in 93% accuracy and it is also shown clearly that SVM classifier, and Welch’s feature extraction method are the dominant ones between classifiers and powers spectral density methods of feature extraction, respectively.

Discussion and Conclusion

The highest accuracy of the current work is using both dominant frequency and power of

all wave bands with Welch’s method alongside SVM classification resulting in accuracy of 93%. Our results indicate that in almost all cases of inputs, SVM classification has a better performance than LDA. Besides, in all cases, except for power of all wave bands as the exclusive input, Welch’s method results in better values of diagnostic accuracy and this method has a much better performance than Yule-Walker AR model, when only dominant frequencies of all wave bands are used. The current work which includes a comparison of feature extraction and classification meth-

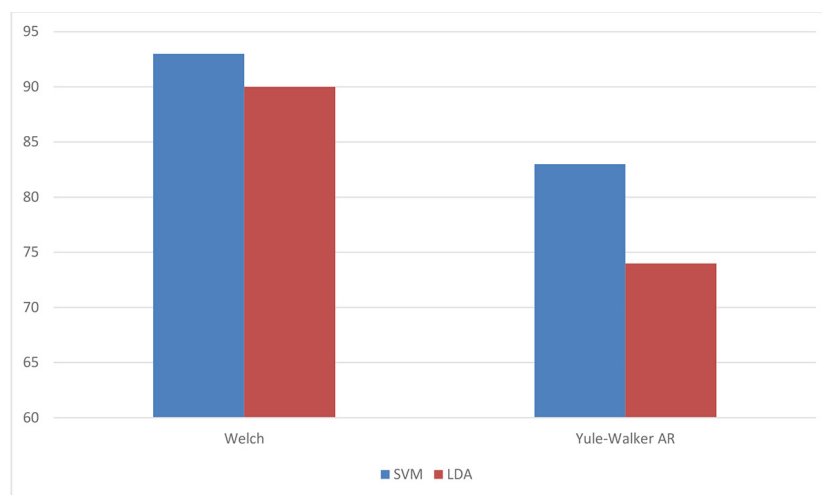


Figure 5: Diagnostic accuracy using exclusively absolute power of wave bands, combined. Yule-Walker shows better performance especially using SVM with 86% accuracy.

od in computer aided diagnosis of migraine without aura has very little amount of background of previous literature, which empowers the necessity, although this fact cannot deny their importance and delicacy. Using Burg AR method [6], as a parametric method has shown that EEG channels F1, T3, O1, and O2 are the most decisive ones in EEG-based diagnosis of migraine having a maximum of 88.4% diagnostic accuracy which is a close result to our findings of autoregressive feature extraction, although the work concentrated on adult migraine. No comparison of either feature methods or classification methods was performed, and only power bands were applied. Our results reveal that the use of power as a classification input, Yule-Walker AR has a better performance alongside SVM classification resulting in an accuracy of 86% and the most decisive channels selected by genetic algorithm were C3, P3, P4, O2, F8, T5, T6, and Fz. Application of genetic algorithm to discover the best combination of features and electrodes and achieve the maximal possible accuracy is another effort which makes this research distinct from previous efforts in migraine EEG classification. Besides, this work has some practical limitations. Various other feature extraction methods and classifiers exist and therefore, can lead to more complete researches in the future. Besides, although this study focuses on pediatric EEG migraine classification, further efforts in this field could include EEG from adult migraineurs.

In sum, analyzing pediatric migraine EEG using two feature extraction tools and two classification methods revealed that best combination of detection accuracy (93%) is when Welch's method and SVM are utilized together, in case dominant frequency is also applied to the classification method.

Acknowledgment

This research was financially supported as MD. Thesis, by Shiraz University of Medical

Science, code: 12359.

Conflict of Interest

None

References

1. Peng KP, Wang SJ. Migraine diagnosis: screening items, instruments, and scales. *Acta Anaesthesiol Taiwan*. 2012;**50**:69-73. doi: 10.1016/j.aat.2012.05.002. PubMed PMID: 22769861.
2. Olesen J, Steiner TJ. The International classification of headache disorders, 2nd edn (ICDH-II). *J Neurol Neurosurg Psychiatry*. 2004;**75**:808-11. PubMed PMID: 15145989; PubMed Central PMCID: PMC1739053.
3. Hedges TR, Jr. Pediatric migraine. *Am J Ophthalmol*. 1983;**95**:844-5. PubMed PMID: 6859200.
4. Teleanu RI, Vladacenco O, Teleanu DM, Epure DA. Treatment of Pediatric Migraine: a Review. *Maedica (Buchar)*. 2016;**11**:136-43. PubMed PMID: 28461833; PubMed Central PMCID: PMC5394581.
5. Kramer U, Nevo Y, Neufeld MY, Harel S. The value of EEG in children with chronic headaches. *Brain Dev*. 1994;**16**:304-8. PubMed PMID: 7818026.
6. Akben SB, Tuncel D, Alkan A. Classification of multi-channel EEG signals for migraine detection. *Biomedical Research*. 2016;**27**:743-8.
7. Bellotti R, De Carlo F, de Tommaso M, Lucente M. Migraine detection through spontaneous EEG analysis. *Conf Proc IEEE Eng Med Biol Soc*. 2007;**2007**:1834-7. doi: 10.1109/IEMBS.2007.4352671. PubMed PMID: 18002337.
8. Schneider T, Neumaier A. Algorithm 808: ARfit—A Matlab package for the estimation of parameters and eigenmodes of multivariate autoregressive models. *ACM Trans Math Softw*. 2001;**27**:58-65. doi: 10.1145/382043.382316.
9. Rejer I, Lorenz K. Genetic algorithm and forward method for feature selection in EEG feature space. *J Theor Appl Comput Sci*. 2013;**7**:72-82.
10. Subasi A, Gursoy MI. EEG signal classification using PCA, ICA, LDA and support vector machines. *Expert Syst Appl*. 2010;**37**:8659-66. doi: 10.1016/j.eswa.2010.06.065.
11. Hosseinifard B, Moradi MH, Rostami R. Classifying depression patients and normal subjects using machine learning techniques and nonlinear features from EEG signal. *Comput Methods Programs Biomed*. 2013;**109**:339-45. doi: 10.1016/j.cmpb.2012.10.008.