Detection of Ocular Artifacts using Bagged Tree Ensemble Model

Sandhyalati Behera
Department of EEE
ITER, Siksha 'O' Anusandhan (Deemed to be University)
Bhubaneswra, India
sandhyalatibehera@gmail.com

Mihir Narayan Mohanty
Department of Electronics and Communication Engineering
ITER, Siksha 'O' Anusandhan (Deemed to be University)
Bhubaneswra, India
mihirmohanty@soa.ac.in

Abstract-Removal of artifacts from the biomedical signal is a cumbersome task. As the desire of physicians is a clean signal for diagnosis authors have tried to detect the artifacts within the signal. Detection of artifacts is a primary job to remove them with further techniques. In this paper we have collected data from Mendeley database and focused on the ocular artifacts. The ensemble method is chosen by the method of bagging and boosting to enhance the detection accuracy. As the technique is one of the statistical techniques, it is found better accuracy in this work. From the dataset 19 channel Ocular artifactual signals are considered along with the healthy signal. The features have been extracted using time domain and frequency domain techniques. Finally, the combination with ensemble classifier shows better accuracy as explained in the result section.

Keywords- EEG, Artifact, Feature Extraction, Classification, Ensemble Classifier.

I. INTRODUCTION

Generally, EEG signal amplitude is small and in the range of microvolt. The EEG signal undergoes for signal distortions called artifacts at the time of recording. The artifacts usually are of higher amplitude and different shape in comparison to the original signal. EEG artifacts are maybe either human-related or technical. Human related artifacts are due to minor body movements. Due to artifact the important information in EEG may lost. So, artifact detection and removal play a vital role in EEG analysis.

Different methods such as manual method, filter method and automatic methods are used for the detection and removal of artifacts from EEG signals[1-3]. ICA, RLS (recursive least square) with the help of adaptive filtering used to separate artifacts from the brain waves. The interference is estimated by RLS algorithm after projecting the brain wave into ICA domain. The estimated interference is subtracted from the EEG signal to get clean EEG[4-7]. It requires EOG reference channel to estimate interference. The artifacts create the difficulty in analysing the EEG so, it is necessary to decrease the artifacts from EEG records. Wavelet transforms used to remove the ocular artifacts without using a reference EOG channel. The statistical techniques such as kurtosis, Multiscale entropy are used to identify artifacts and wavelet transform is used for removal of artifacts[6, 8-10]. The modified multiscale entropy is used to identify the artifacts effectively.[11] The ICA, PCA, Wavelet ICA are used to reduce the dimension of EEG data. After reducing the dimension, the kurtosis and modified multiscale entropy are used to identify the artifactual components and the Wavelet-ICA removes the artifacts. The artifact removal process involves the identification of artifacts and cancelling or correcting the artifacts without distorting the original signal of interest. The artifact may overlap with EEG signal in time and spectral domains so the use of simple filtering or straight forward signal processing techniques cannot remove artifact[12].

The regression method requires a reference signal and it fails if a reference signal is not available. The adaptive filtering method also requires a reference input to compare the desired output with observed output.[13] The time-domain filter implemented by Empirical mode decomposition technique uses fractional Gaussian noise (fGn) as reference signal to detect the distinguishing feature of EOG signal. To cancel Ocular, muscular and Cardiac artifacts from the EEG signal adaptive filter based on FLN-RBFN can be used. The spatial decomposition methods in [14] can be used to remove ocular artifacts.

The classification methods such as Linear Discriminant Analysis (LDA), Support vector machine (SVM), Artificial Neural Network (ANN), Nearest neighbour classifier are commonly used classifiers to detect the artifacts from EEG[15]. Due to high dimension and noisy nature of EEG the nonlinear classifiers such as neural networks and support vector machines produce slightly better results. There are many applications of

978-1-7281-3908-1/19/$31.00 ©2019 IEEE
DOI 10.1109/ICAML48257.2019.00016

EEG classification like artifact detection, sleep analysis, to understand depth of anaesthesia, detecting drowsiness, etc. The feature extraction plays an important role in classification. This is first step in classification process. To obtain satisfactory performances appropriate features are necessary[12]. To classify the general artifactual source components the automatic classification method like ICA, which takes temporal correlation into account. Wavelet transform represents a non-stationary signal into both time and frequency domain. The analysis of EEG signal using wavelet transform and classification using artificial neural network (ANN) and logistic regression (LR) gives effective results. To increase the computational speed and processing speed lifting based wavelet transform (LBDWT) is used.[12, 16] The weighted distance nearest neighbour (WDNN) classification algorithm which is an adaptive classification method, assigns a weight to each training sample. These weights are used to find nearest neighbour of an input pattern. The Bayesian classifier used to distinguish between clean EEG and artifactual EEG segments. In Bayesian classifier the probability of artifacts presents in the features calculated from the Independent components of EEG segments[17]. The simple Bayesian classifier calculates the posterior probability from the given prior probability. The Bayesian classifier which calculates a weighted-log posterior function and supervised learning gives effective results than the simple Bayesian classifier.

In this paper, the brief introduction about artifact detection, removal, and classification given in section 1. The methods for obtaining the features are given in section 2. The classifier taken for classification of the features is given in section 3. Results and discussion given in section 4 and section 5 are about the conclusion and future work.

II. METHOD

A. Feature extraction

The signal is observed in time domain for which the time domain methods been chosen for the feature set. Amplitude peaks and correlation peaks are obtained as the time domain features. Further spectral peaks are obtained and combined with each set of peaks.

The amplitude peak is obtained by taking maximum amplitude of an EEG signal. The correlation peak is obtained by taking maximum value after correlation of two EEG signal. Mathematically correlation between two EEG signal is given as:

$$Y_{x_1x_2}(n) = \sum_{m=0}^{N-n-1} x_{1_{m-n}} x_{2_m} \qquad n \geq 0$$
$$= Y_{x_2x_1}(-n) \qquad n < 0$$

Where x_1 and x_2 are two EEG signal. While taking Cross correlation of two clean EEG signal One signal is taken as reference with that reference EEG signal other clean EEG signal is cross correlated. Similarly, for obtaining cross correlation between two contaminated EEG signal one signal is taken as reference with that other contaminated EEG signal is cross correlated.

The spectral peak is obtained by making peak frequency spectrum of EEG signal.

Frequency spectrum is obtained by taking DFT of EEG signal. Mathematically DFT of an EEG signal is obtained as,

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{\frac{j2\pi kn}{k}}$$
$$k = 0,1,2,3.....N-1 \tag{2}$$

B. Classification using ensemble method

Ensemble method is the combination of decision tree to produce better prediction as compared to single decision tree. The measure principle behind it is that a group of weak learners come together to form a strong learner. These techniques include bagging and boosting. Bagging is meant to reduce the variance of decision tree by creating several subsets of data from the training sample chosen randomly. Further the new tree is replaced the earlier trees as effective one. Finally, with an ensemble of different models, the consecutive tree is fit at every step towards the goal to reduce the error from prior tree. This method can even suitable for datamining techniques also. Every tree in ensemble is grown on an independently drawn root strap replica of input data.

III. RESULTS AND DISCUSSION

The EEG signal considered for the proposed work is shown in fig 1 and fig 2. The clean EEG signal shown in fig 1 is of 6001 samples.

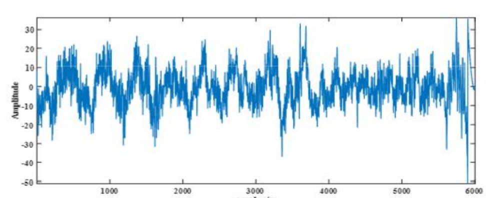

Fig.1. Clean EEG signal

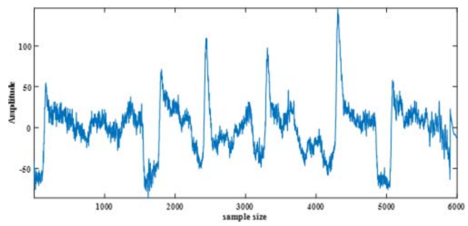

Fig.2. Contaminated EEG signal

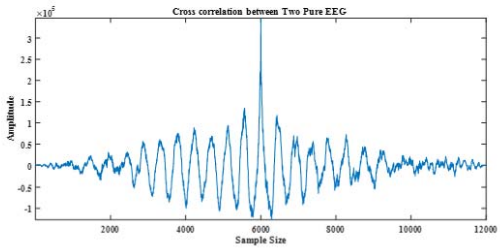

Fig.3. Correlation of two clean EEG signal

The classification of features is performed by using MATLAB2015a. The time domain and frequency domain features are classified by using the SVM classifier. The time-domain features are obtained from the cross-correlation are given in table 1 and the frequency domain features are given in table 2.

Features	Ensemble Classifier performance
Amplitude Peak	81.69%
Correlation Peak	91.23%
Spectral Peak	89.69%
Combined Feature	**92.55%**

For the proposed classification work 5 time-domain features and five frequency domain features are considered. Time-domain features are represented as Amplitude peak features and correlation peak. Spectral peaks are frequency domain features.

IV. CONCLUSION

In this paper, the Ensemble classifier is successfully implemented for detection of artifact. The combined feature of time and frequency domain gives better results. In future the artifacts are to be removed to find the clean signal for diagnosis.

REFERENCES

[1] M. Teplan, "Fundamentals of EEG measurement," *Measurement science review,* vol. 2, pp. 1-11, 2002.

[2] O. Aydemir, "Classifying Various EMG and EOG Artifacts in EEG Signals," *Przeglad Elektrotechniczny,* vol. 88, pp. 218-222, 11/01 2012.

[3] E. Parvinnia, M. Sabeti, M. Z. Jahromi, and R. Boostani, "Classification of EEG Signals using adaptive weighted distance nearest neighbor algorithm," *Journal of King Saud University-Computer and Information Sciences,* vol. 26, pp. 1-6, 2014.

[4] R. J. Croft and R. J. Barry, "Removal of ocular artifact from the EEG: a review," *Neurophysiologie Clinique/Clinical Neurophysiology,* vol. 30, pp. 5-19, 2000.

[5] A. Delorme, T. Sejnowski, and S. Makeig, "Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis," *Neuroimage,* vol. 34, pp. 1443-1449, 2007.

[6] R. Dhiman, J. Saini, and A. Priyanka, "ARTIFACT REMOVAL FROM EEG RECORDINGS–AN OVERVIEW," *Proc. NCCI,* pp. 1-6, 2010.

[7] H. Ghandeharion and A. Erfanian, "A fully automatic ocular artifact suppression from EEG data using higher order statistics: Improved performance by wavelet analysis," *Medical engineering & physics,* vol. 32, pp. 720-729, 2010.

[8] P. S. Kumar, R. Arumuganathan, K. Sivakumar, and C. Vimal, "A wavelet based statistical method for de-noising of ocular artifacts in EEG signals," *IJCSNS International Journal of Computer Science and Network Security,* vol. 8, pp. 87-92, 2008.

[9] N. Mammone, F. La Foresta, and F. C. Morabito, "Automatic artifact rejection from multichannel scalp EEG by wavelet ICA," *IEEE Sensors Journal,* vol. 12, pp. 533-542, 2011.

[10] A. Schlögl, C. Keinrath, D. Zimmermann, R. Scherer, R. Leeb, and G. Pfurtscheller, "A fully automated correction method of EOG artifacts in EEG recordings," *Clinical neurophysiology,* vol. 118, pp. 98-104, 2007.

[11] S. Behera and M. N. Mohanty, "A Statistical Approach for Ocular Artifact Removal in Brain Signals," in *2018 2nd International Conference on Data Science and Business Analytics (ICDSBA)*, 2018, pp. 500-503.

[12] P. LeVan, E. Urrestarazu, and J. Gotman, "A system for automatic artifact removal in ictal scalp EEG based on independent component analysis and Bayesian classification," *Clinical Neurophysiology,* vol. 117, pp. 912-927, 2006.

[13] R. Mahajan and B. I. Morshed, "Unsupervised Eye Blink Artifact Denoising of EEG Data with Modified Multiscale Sample Entropy, Kurtosis, and Wavelet-ICA," *IEEE Journal of Biomedical and Health Informatics,* vol. 19, pp. 158-165, 2015.

[14] S.-D. Wu, C.-W. Wu, K.-Y. Lee, and S.-G. Lin, "Modified multiscale entropy for short-term time series analysis," *Physica A: Statistical Mechanics and its Applications,* vol. 392, pp. 5865-5873, 2013.

[15] I. Winkler, S. Haufe, and M. Tangermann, "Automatic classification of artifactual ICA-components for artifact removal in EEG signals," *Behavioral and Brain Functions,* vol. 7, p. 30, 2011.

[16] M. K. I. Molla, M. R. Islam, T. Tanaka, and T. M. Rutkowski, "Artifact suppression from EEG signals using data adaptive time domain filtering," *Neurocomputing,* vol. 97, pp. 297-308, 2012.

[17] R. Sameni and C. Gouy-Pailler, "An iterative subspace denoising algorithm for removing electroencephalogram ocular artifacts," *Journal of neuroscience methods,* vol. 225, pp. 97-105, 2014.

[18] M. R. Lakshmi, T. Prasad, and D. V. C. Prakash, "Survey on EEG signal processing methods," *International Journal of Advanced Research in Computer Science and Software Engineering,* vol. 4, 2014.

[19] M. Nandish, M. Stafford, P. H. Kumar, and F. Ahmed, "Feature extraction and classification of EEG signal using neural network based techniques," *International Journal of Engineering and Innovative Technology (IJEIT),* vol. 2, pp. 1-5, 2012.