

Optimal determination of wavelet for football player EEG using SVM classifier.

Ashish Aggarwal*, Ravinder Agarwal

Thapar University, Punjab, India

Abstract

This study investigates EEG for evaluation of the cerebrum and cognitive processes for sports' applications. An infamous issue with EEG signals for the cerebral information is frequently contaminated by artifacts of the non-cerebral cause. The nearness of curios makes the examination of EEG troublesome for sports' applications. And, to manage these artifacts, various strategies and procedures have been developed by numerous specialists applying conventional filters, artificial intelligence and time-frequency based techniques. Wavelet Transform outperforms for denoising non-stationary EEG signals, but the performance of this technique is highly dependent on wavelet selection. This paper explores an appropriate selection of particular wavelet which presents a modular statistical framework providing a pathway. A comparative analysis of 6 different wavelet families for effective filtering was done. The results show that biorthogonal wavelet family (bior3.1) was more reasonable and a classification accuracy of 91.68% was achieved for denoising the above-stated signals.

Keywords: EEG, Wavelet, ANOVA, SVM, SNR.

Accepted on November 14, 2016

Introduction

Electroencephalography (EEG) signal refers to the recording of the unconstrained electrical movement of the cerebrum over a little time frame [1-3]. The neurons of the human mind process data, by changing the stream of electrical ebbs and flows over the membrane. These changing streams produce electric fields that can be recorded by putting electrodes on the surface of the scalp. The potentials between various electrodes are then amplified and recorded as the EEG; which implies the composition out of the electrical movement of the cerebrum [4]. In this way EEG recordings finish learning about the general movement of a large number of neurons in the mind. The mind is a standout amongst the most imperative organs of people, for controlling the coordination of human muscles and nerves. The EEG is utilized as a part of the assessment of mind issue [5-8]. It is additionally used to assess individuals who are having issues connected with the cerebrum. Though, the investigation of ceaseless EEG signal is mind boggling.

Diverse sorts of EEG waves are classified by the recurrence in particular Delta waves (0.5-4 Hz), Theta waves (4-7.5 Hz), Alpha waves (7.5-14 Hz), Beta waves (14-30 Hz), Gamma waves (above 30 Hz). All these mentioned waves speak to various mental conditions of an individual [9,10]. EEG signals consist of little amplitudes and can be effortlessly defiled by artifacts. Such artifacts can be generated by an electrode or by the subject body itself [11,12]. The different sorts of artifacts that can happen in the signals amid recordings are the electrode noise, baseline movement, EMG disturbance, eye blink, neck

movement, hand movement etc [13-15]. These artifacts should be expelled primarily for appropriate handling before studying the mind of the subject.

Many techniques have been proposed like auto-regression, artificial intelligence techniques, blind source separation techniques (BSS), empirical mode decomposition (EMD), wavelet analysis to extract artifacts from EEG signals [12]. Wavelet transform is an important tool for examination of EEG signals; it is localized in both time and frequency and gives a successful answer for denoising non-stationary EEG. Also, it offers great time resolution for low-frequency parts and great frequency resolution for high-frequency segments of the signal being broken down. To expel artifacts, the wavelet transform is used by numerous authors as a part of their work. Hundreds of wavelets are present but, information obtain always vary with different types of wavelets. So, it is important to find the best wavelet for artifact reduction. The existing literature occupied with number of studies for particular application by applying various kind of wavelets which proves the importance of this technique. But, at the same time these studies fails to provide reasoning for wavelet selection [11,15-17]. That's why; this study is carried out to find the best wavelet, so that maximum information could be obtained.

Materials and Methods

Twenty five young male football players were chosen and their EEG signals were recorded (with an age scope of 18-23 years, mean 20.63 years, SD \pm 1.69 years). All members were

informed about the experiment and their written consents were taken. EEG estimation involves the connection of electrodes to institutionalized areas on the scalp using Emotive Epoc. It is a high determination, 14 channels (and 2 reference channels CMS and DRL) versatile system appeared as a part of Figures 1 and 2. EEG movement was recorded constantly from 14 channels organized by broadened worldwide 10-20 system [17]. The real time recording of the EEG signal from the subject's scalp is demonstrated as in Figure 3. EEG action at the accompanying electrodes locales: (AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8 and AF4) were recorded [5,10], as represented in Figure 4. Electrodes are appended to the skin using conductive glue with impedances, for the most part, kept underneath 5 kΩ. Preceding appending the electrodes, the skin is normally arranged with an abrasive paste, e.g., Nu-Prep to decrease skin impedance. Signals were examined at sample rate 2048 Hz.



Figure 1. Emotive.

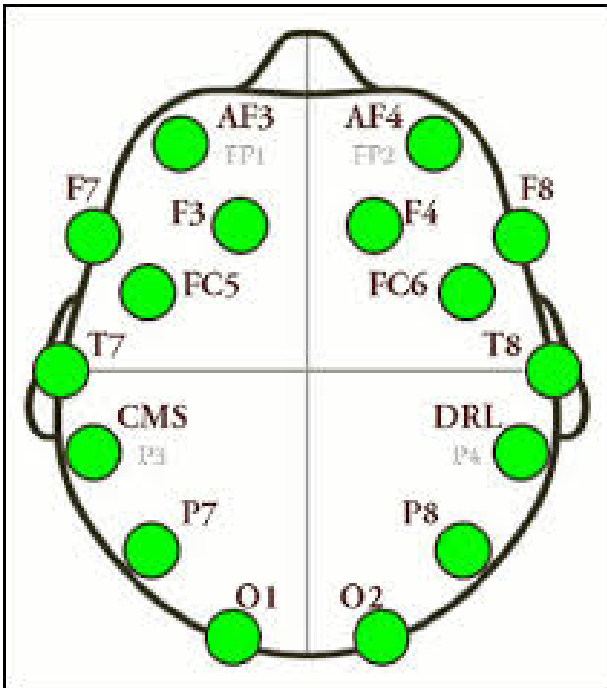


Figure 2. 10-20 Map of EEG electrodes.



Figure 3. Real time data acquisition using Emotive Epoc.

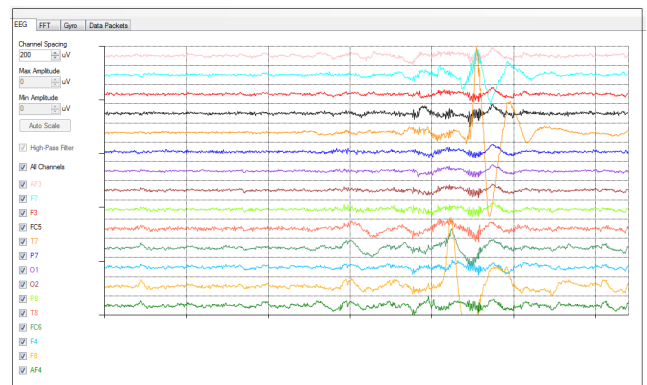


Figure 4. Emotive test bench window showing EEG signal and artifacts.

EEG signals were procured by Emotive test seat v1.5.0.3 while playing the MOT (multiple object tracking) game. 2D MOT game was developed in unity which had 14 levels and difficulty increased with the expansion of the levels. All the participants were seated individually in front of computer screen. Each subject had to play the game for 20 min. At the primary level, 8 still circles were showed on the screen, 3 were green and the remaining was blue. These 3 circles are taken as targets and the remaining ones as obstacles. At that moment, circles moved on screen for 10 sec. After this, the circles stopped moving and got blue. Here the subjects demonstrated the three targets by selecting them with the aid of a PC mouse and needs to recognize the green circles. After participants had picked right circles, they were permitted to continue; otherwise, the same level was repeated. Figure 5 demonstrates the MOT game. With the expansion of the levels, a number of target balls and obstruction were also increased [1]. In an MOT, where the dependent measures used was the maximum number of elements that could be tracked. It was proposed that speed thresholds may be a more relevant and controllable measure of MOT that demands rather than using the total number of items as a dependent measure (Figure 6) [9].

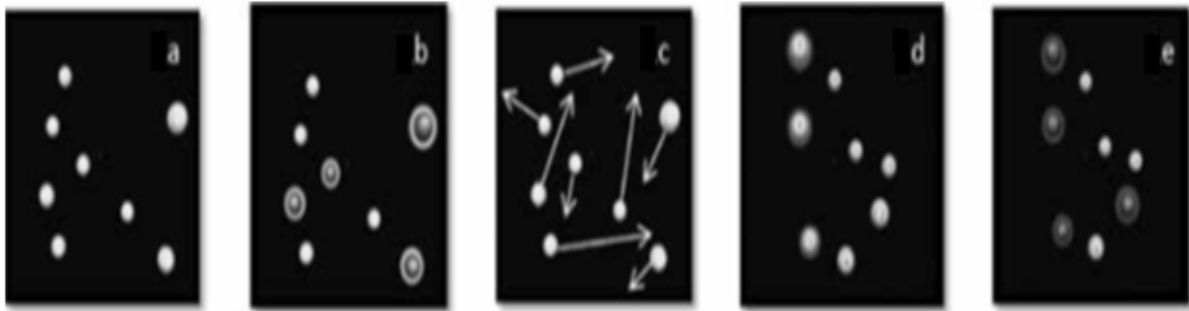


Figure 5. (a) Presentation of randomly positioned spheres in a virtual volumetric space, (b) identification of the spheres to track during the trial, (c) removal of identification and movement of all spheres with dynamic interactions, (d) observer's response by identifying the spheres, (e) Feedback is given to the observer.

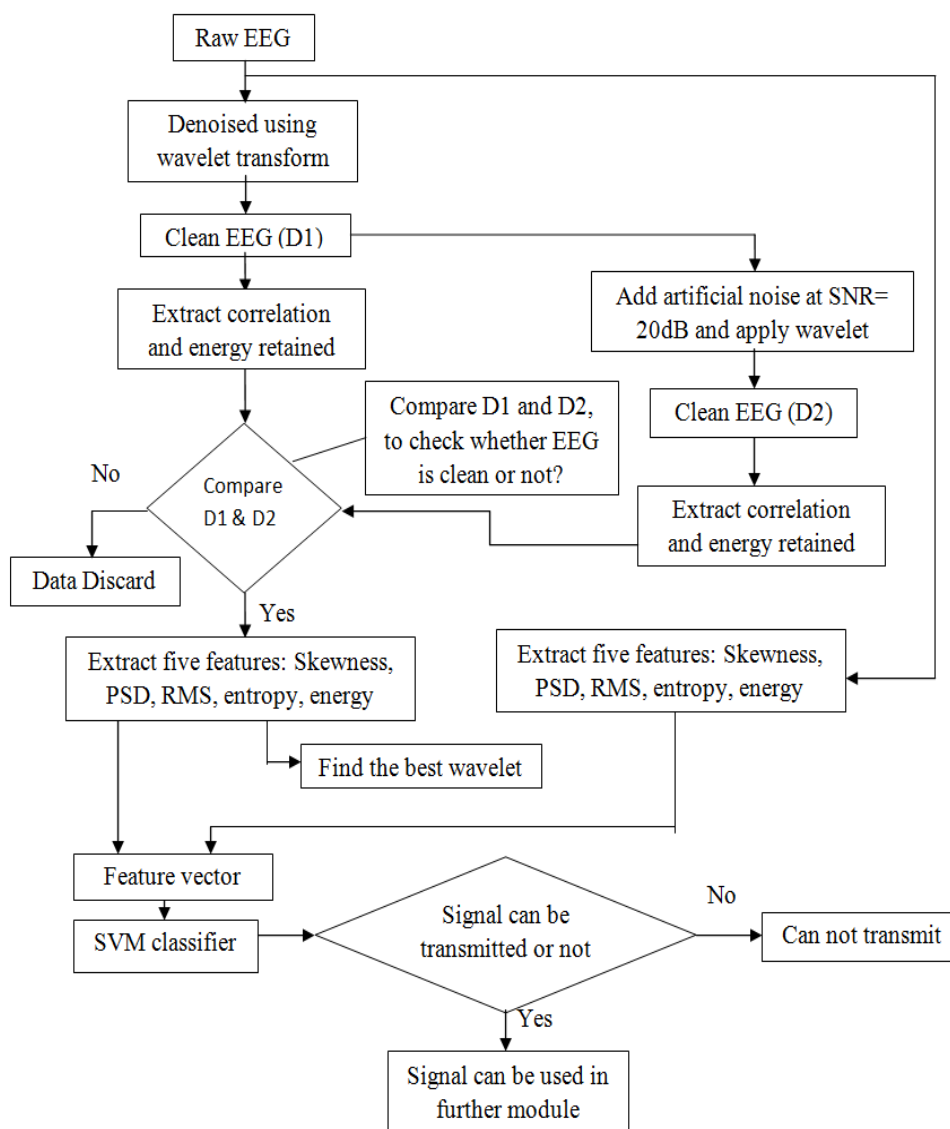


Figure 6. Flow chart representing the algorithm.

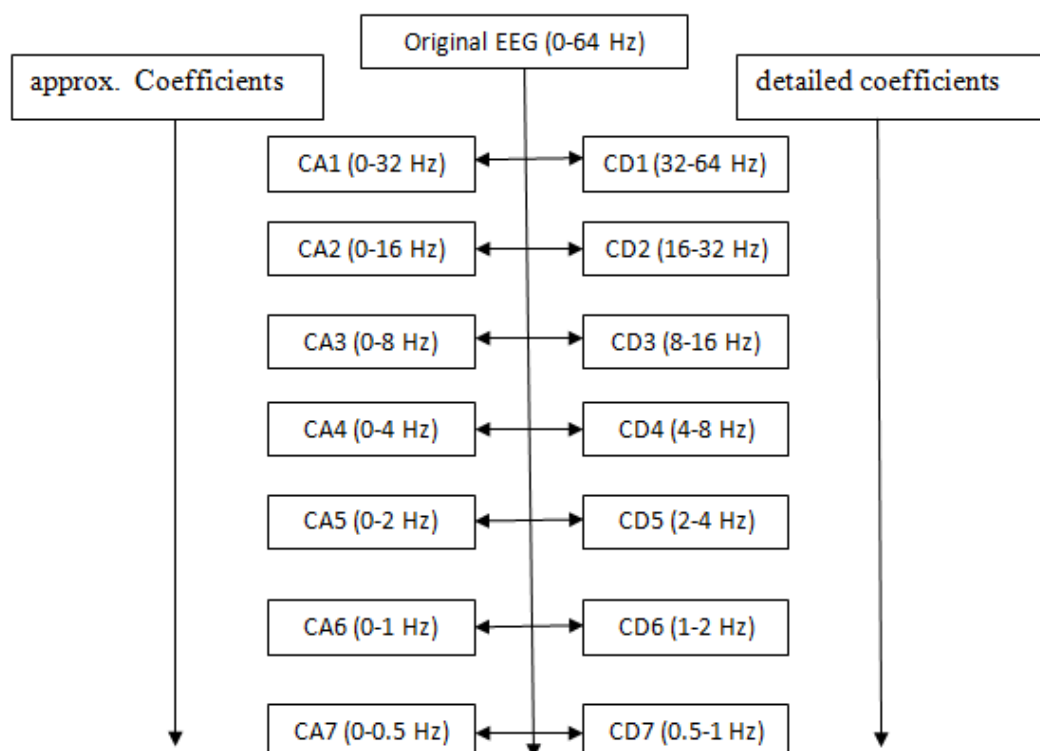


Figure 7. Wavelet transforms decomposition structure of EEG.

EEG signals acquired were contaminated by the artifacts and wavelet transform were used to obtain clean EEG. For this work, 109 of unique wavelets were used and signals were denoised by every wavelet. The guideline of wavelet denoising comprises of breaking down (7 level) of the signals by performing wavelet, applying appropriate global thresholds to the detail coefficients. Lastly, to expel the baseline wanders in the signal, the approximation coefficients (cA7) were disposed by remaking the denoised signal and taking the altered coefficients into account. The decomposition of the signal by wavelet transform is shown in Figure 7. This denoised signal is named as D1. After achieving clean EEG, it is important to check, whether is it artifact free or not? If yes, which wavelet should be used to obtain more information with minimum losses. To solve this purpose, an artificial noise (random noise and white noise at 20 db SNR) was added to the denoised (D1)

EEG signal which got denoised (named as D2) with the same procedure. Also, to figure out the results, components (correlation coefficients (r) amongst D1 and D2 and energy retained) were compared.

EEG is a non-stationary signal and its amplitude varies randomly above and below zero values, the further analysis becomes essential in a way to define characteristic properties of the signal. A wide variety of features have been considered individually and in the group, representing both the EEG amplitude and spectral content. So, feature extraction was done for interpretation of the recorded signal. Five features, skewness, RMS, energy, entropy, PSD [3,13,16] were used to figured out the best wavelet for denoised signals (Table 1). All these features have certain specific advantages.

Table 1. List of features.

	Features
Denoised signal features	Skewness
	Energy
	Root mean square (RMS)
	Entropy
	Power spectral density (PSD)
Artificial added noise denoised features	Correlation
	Energy retained

We were interested in refining the experiment to increase its sensitivity for detecting differences in the dependent variables. To promote the elucidation, ANOVA was used for all features for various specimens with the earlier comparison to test the results if significant or not. The essential technique for ANOVA is to determine two different estimates of the variance from the information and by this; proportion (F ratio) of these two assessments is calculated. One of these estimates is the sum of squares between the groups (SSB) and second is the sum of squares within the groups (SSW). At that point, a significant F ratio between two assessments is ascertained for $p < 0.05$. The noteworthy F ratio shows that the population means are not all equivalent to each other. It is the measurement that is used to test the hypothesis and the genuine impacts [7]. Analysis of variance (ANOVA) tells us whether the results are accepted or not.

These results can be transmitted to further module and further information can be extracted. Human brain is a very complex system and it is very difficult to understand it. Further, for complex systems, segregation between raw EEG and clean EEG is also a matter of concern. For the classification of EEG signals, conventional classifiers like neural network, linear Discrimination Analysis (LDA), Support Vector Machines (SVM) are used. The current study refines the experiment to build its affect ability for distinguishing contrasts between two

classes and accuracy was obtained for every wavelet. The strategy proposed in this paper shown in Figure 6 includes the accompanying strides:

1. Apply wavelet transform to tainted EEG and break down it up to the seventh level (select any wavelet from Table 2) appeared in Figure 7.
2. Global thresholding is applied after this; it includes the disposing of approximation coefficients to zero at the seventh deterioration level to expel baseline drift (0-0.5Hz).
3. Apply wavelet reconstruction to produce denoised EEG signal.
4. Add artificial noise (random noise and white Gaussian noise at 20 dB SNR)
5. $Y = [\text{awgn}(D1, 20, 'measured') + A * \text{rand}(1, \text{length}(D1))]$;
6. Where, A is arbitrary constant, and $A = 2$ in this study.
7. Again denoised the signal.
8. Correlation coefficients and energy retained figure out.
9. Calculate the features: skewness, energy, RMS, entropy, PSD
10. One way ANOVA was executed to determine the significance of results
11. SVM classifier applied between two classes (raw EEG and clean EEG).

Table 2. Type of wavelets.

Wavelet family	Wavelets	No. of wavelets
Haar	Haar	01
Daubencies	db1 to db45	45
Symlets	Sym2 to sym28	27
Coiflet	Coif1 to coif5	05
Biorsplines	bior1.1, bior1.3, bior1.5, bior2.2, bior2.4, bior2.6, bior2.8, bior3.1, bior3.3, bior3.5, bior3.7, bior3.9, bior4.4, bior5.5, bior6.8	15
Reversebior	rbio1.1, rbio1.3, rbio1.5, rbio2.2, rbio2.4, rbio2.6, rbio2.8, rbio3.1, rbio3.3, rbio3.5, rbio3.7, rbio3.9, rbio4.4, rbio5.5, rbio6.8	15
Dmeyer	Dmey	01
	Total	109

Results and Discussion

The investigations of the components said above were done to break down the relationship amongst noisy and denoised signal using MATLAB 2012 for feature extraction and SVM classifier. The outcomes are displayed as the mean of all emphases. Counterfeit commotion (white Gaussian noise and random noise) were added to the denoised EEG which was denoised by using the same technique (named as D2). Two more elements were made sense to discover the connection amongst D1 and D2. Firstly, numerical 'correlation coefficient'

was found between two signals. Correlation is a measurable system that demonstrates the emphatically combination of variables. The primary consequence of a relationship is known as the correlation coefficient (or "r"). It ranges from -1.0 to +1.0. Figure 8 demonstrates the correlation coefficients versus wavelets. "bior3.1" showed the best correlation coefficients (0.9975) as it is closer to +1. Energy retains between D1 and D2 was also calculated. Energy retain is the ratio of the energy of D1 and D2 signals. This feature shows that much energy of the signal is retained after removing the artificial noise. Figure 9 shows that maximum energy is retained by "bior3.1". So, it is

concluded that artifacts are effectively eliminated from the EEG signals and it can performed effectively by "bior3.1", but

which wavelet performed best to remove artifacts is further calculated.

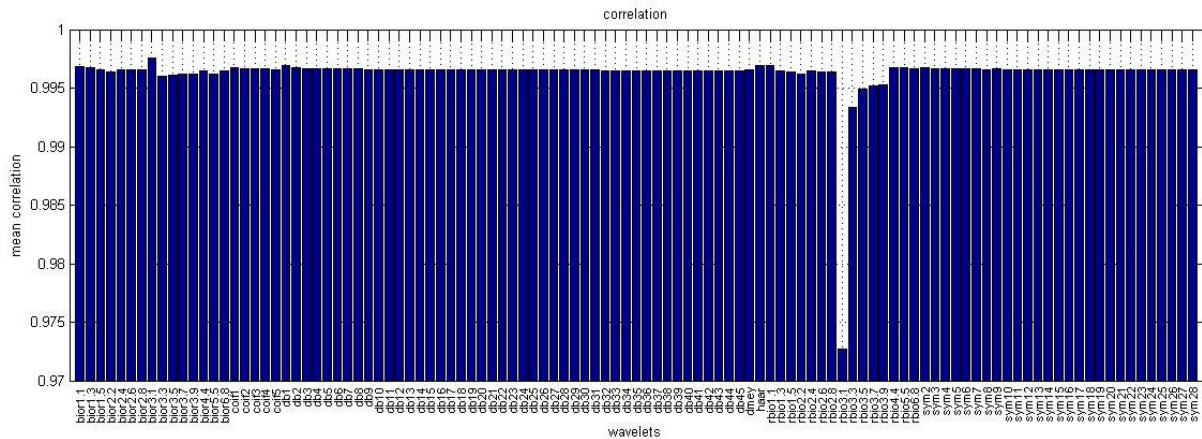


Figure 8. Correlation coefficients vs. wavelets.

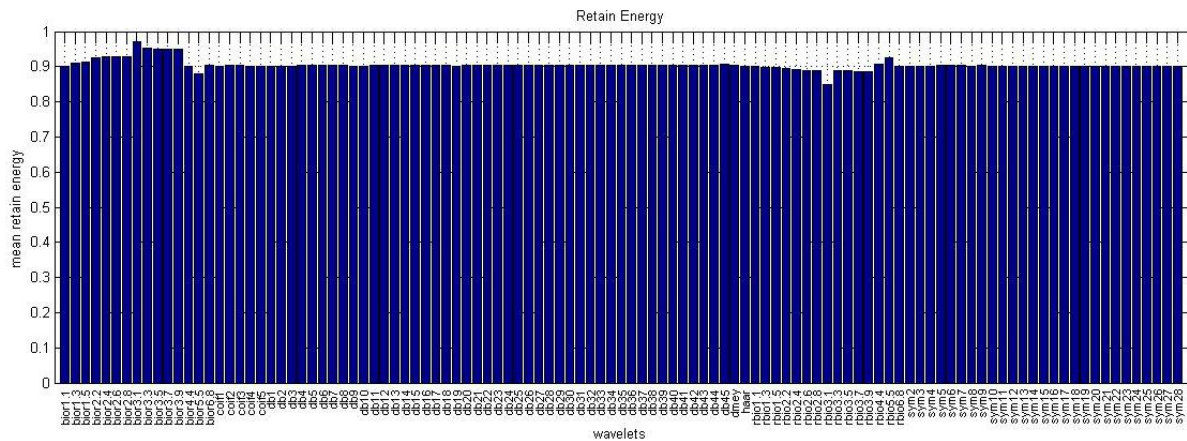


Figure 9. Energy retains vs. wavelets.

The RMS is generally called the quadratic mean, are a factual measure of the extent of a fluctuating quantity. It is particularly helpful when variations are positive and negative. It is used for characterizing the power of the signal. This parameter gives the physical significance of the signal for a specific genuine force. EEG is denoised by wavelets for 25 subjects and the average RMS value is taken. Figure 10 demonstrates the plot of average RMS versus wavelets. As indicated in Figure 10, "bior3.1" demonstrated superior execution esteem while "rbior3.5" performed the most exceedingly terrible. This implies wavelet function bior3.1 is equipped for denoising EEG superior to another wavelet. To examine the significance of the RMS esteem, F ratio was figured and it gave $F(1,216)=11495$ for $p=2.82153e-189$ which is quite smaller than 0.05, i.e., null hypothesis of equal means is rejected, and at last the test statistic is significant.

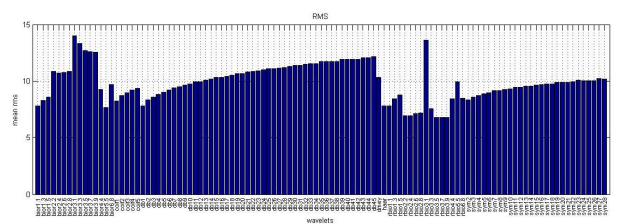


Figure 10. Average RMS values vs. wavelets.

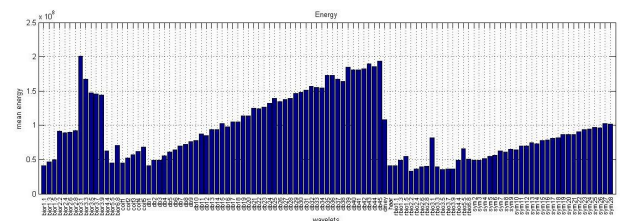


Figure 11. Average Energy vs. wavelets.

Another parameter, energy of the denoised EEG by various wavelets additionally was figured out [13,16,17]. EEG signal is denoised by all wavelets one by one and the average energy computed. Figure 11 demonstrates the plot of average energy versus wavelets which indicate "bior3.1" demonstrated a superior execution esteem and "rbior2.2" performed the worst. To research the importance of this element, F ratio was computed and it gives $F(1,216)=13711.23$ for $p=2.09498e-197$ which is much smaller than 0.05, i.e., null hypothesis of equivalent means is rejected and at long last, the test statistic is significant.

Furthermore, attributes like PSD, entropy and skewness were examined to support the above said result. Entropy is a measure of regularity to quantify levels of complexity within a time series. These components versus wavelets were plotted as appeared in Figure 12 for PSD, Figure 13 for entropy and Figure 14 for skewness. "bior3.1" again demonstrates the better execution as appeared in Figures 10 and 11. Skewness is the measure of the normalization of the signals. EEG got denoised by "rbior3.1" demonstrates the minimum skewness and "bior3.1" indicates second best. Further, F ratio was calculated and $F(1,216)=6056.67$ ($p=5.4686e-160$), $F(1,216)=132.12$ ($p=3.59017e-24$) and $F(1,216)=1.29081e+06$ ($p=0.000$) was observed for elements PSD, entropy and skewness, respectively. In every case, the estimation of p is much smaller than 0.05 means the all test statistic are significant [18].

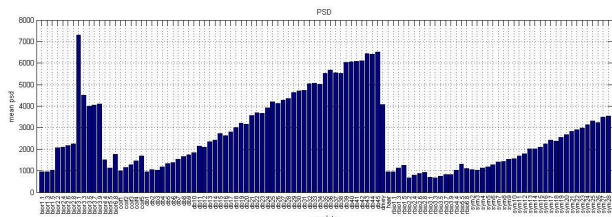


Figure 12. Average PSD vs. wavelets.

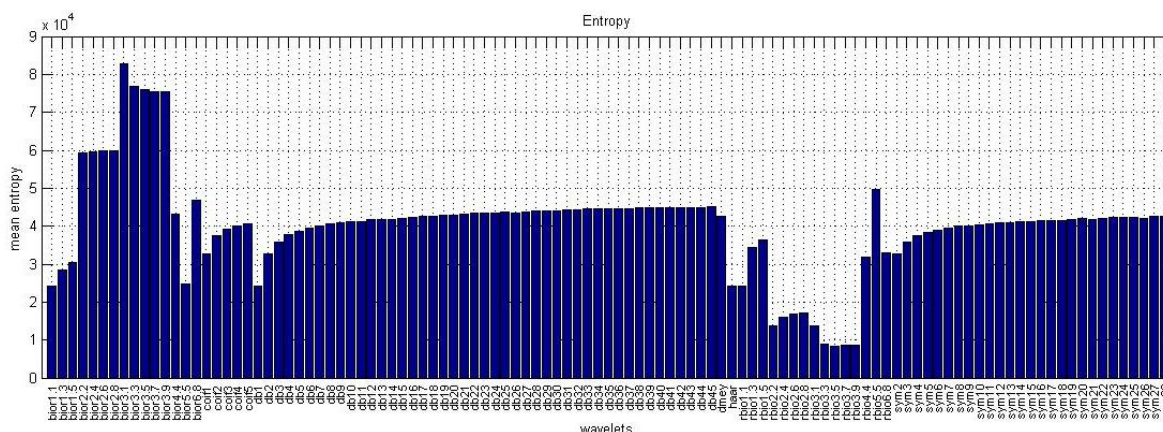


Figure 13. Average entropy vs. wavelets.

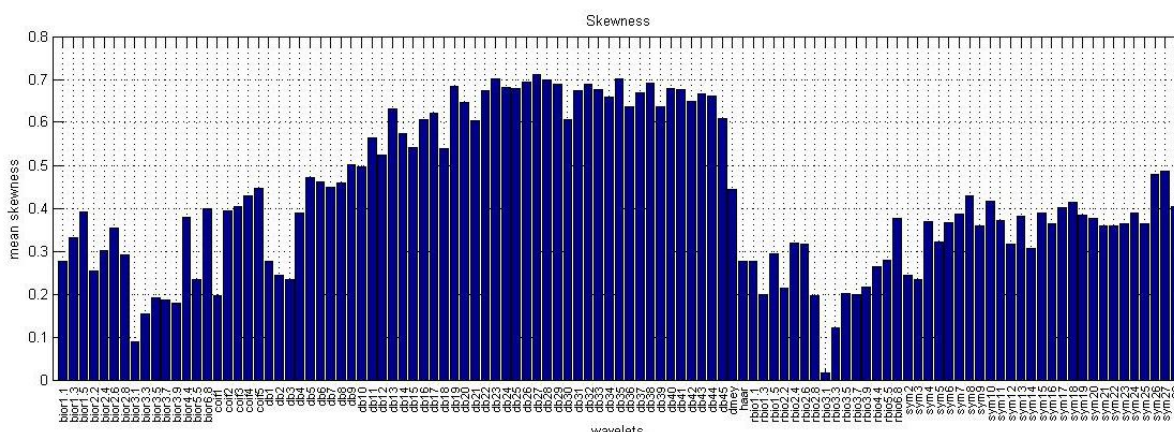


Figure 14. Average skewness vs. wavelets.

Here, five distinct elements were executed for each wavelet and each element has diverse significance for EEG. Artifacts were removed effectively from the EEG signals. EEG signals

are used for several cognitive processes in complex systems. Thus, classification between raw and clean EEG signals is also important. Usually, feature selection relies on two primary

errands: a pertinence model and a seeking technique. The mutual information could be used as significance foundation since it has been generally acknowledged as a decent pointer to gauge the significance of the component [3]. Therefore, SVM was used to characterize the information. The complete information set was partitioning into two sections; training data and testing data. 80% of the complete information set were chosen as training data randomly and SVM trained by this information, remaining 20% were selected as testing

information. MATLAB based SVM testing was done for testing information and the classification accuracy was calculated. The testing model was keep running for 25 emphases for the better outcomes and the average accuracy was calculated. Figure 15 exhibits the plot of average accuracy versus wavelets. This infers wavelet 'bior3.1' is prepared for denoising EEG better than anything other wavelet and classification accuracy 91.68% observed while 'rbio3.7' performs the worst, giving 68.13% accuracy.

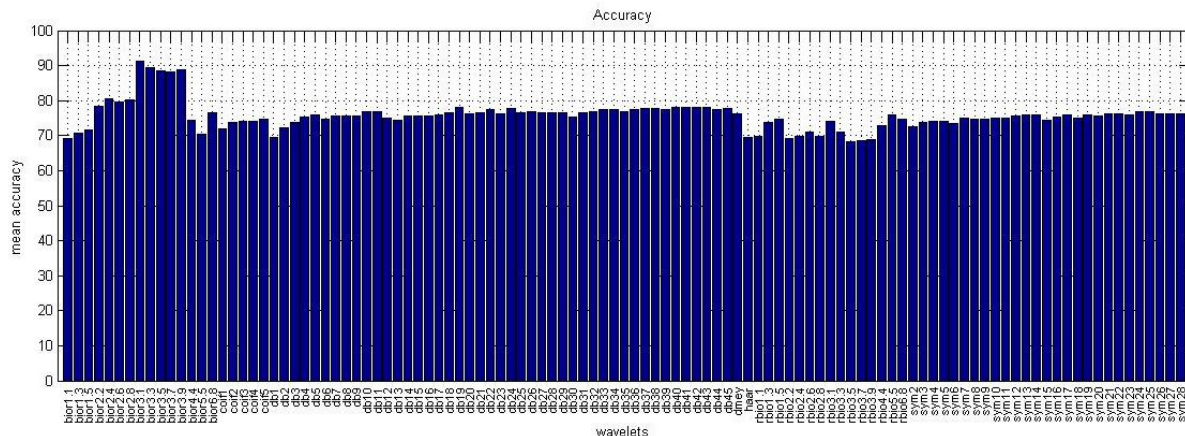


Figure 15. Average accuracy vs. wavelets.

Conclusion

Finally, it was concluded that for any EEG signal occurring under the aforesaid conditions and spectrum, the designed algorithm is capable of removing the artifacts signal up to an enhanced level, so that it could be analysed easily, and wavelet bior3.1 performs denoising the best and is suitable for accurate classification of signals. One-way ANOVA helps ensure class separability relationship for proving data is significant. Future work will further evaluate the techniques used in this study by incorporating the more wavelet comparative studies to improve the accuracy in representing the nature of EEG by considering brain waves (delta, theta, alpha, beta and gamma waves) with different noise levels.

References

1. Arend A, Zimmer H. Successful training of filtering mechanisms in multiple object tracking does not transfer to filtering mechanisms in a visual working memory task: Behavioral and Electrophysiological evidence. *Neuropsychologia*, 2012; 50, 2379-2388.
2. Kumar A, Mohanty M, Routray A. Design of support vector machines with time frequency kernels for classification of EEG signals. *Proceedings of the 2010 IEEE Students' Technology Symposium, IIT Khargpur, IEEE 2010; 330-333.*
3. Rabbi A, Ivanca K, Putnam A, Musa A, Thaden C, Rezai R. Human performance evaluation based on EEG signal analysis: a prospective review. *31st Annual International Conference of the IEEE EMBS, Minneapolis, Minnesota, USA, 2009; 1879-1882.*

4. Stewart A, Nuthmann A, Sanguinetti G. Single-trial classification of EEG in a visual object task using ICA and machine learning. *J Neurosci Methods* 2014; 228, 1-14.
5. Zaammouri A, Aitmousa A, Chevaller S, Monchelli E. Intelligent ocular artifacts removal in a non-invasive single channel EEG recording. *IEEE 2015; 1-5.*
6. Yang B, He L, Lin L, Wang Q. Fast removal of ocular artifacts from electroencephalogram signals using spatial constraint independent component analysis based recursive least squares in brain-computer interface. *Springer, 2015; 16: 486-496.*
7. Mosquera C, Verleysen M, Vazquez A. EEG feature selection using mutual information and support vector machine: A comparative analysis. *EMBS, IEEE, 2010; 4946-4949.*
8. Garrett D, Peterson D, Anderson C, Thaut M. Comparison of Linear, Non-linear, and Feature Selection Methods for EEG Signal Classification”, *IEEE 2013; 11: 141-144.*
9. Legault I, Allard R, Faubert J. Healthy older observers show equivalent perceptual-cognitive training benefits to young adults for multiple object tracking. *Front Psychol* 2013; 323: 1-7.
10. Blinowska K, Durka P. *Electroencephalography (EEG). Wiley Encyclopedia of Biomedical Engineering 2006; 1-15.*
11. Zhang L, Wu D, Zhi L. Method of removing noise from EEG signals based on HHT method. *ICISE, IEEE 2009; 596-599.*

12. Sheoran M, Kumar S, Chawla S. Methods of denoising of electroencephalogram signal: a review. *IJBET, Inderscience* 2015; 18: 385-395.
13. Babu P, Prasad K. Removal of ocular artifacts from EEG signals using adaptive threshold PCA and wavelet transforms. *ICCSNT, IEEE* 2011; 572-575.
14. Kumar P, Kareemullah H. EEG signal with feature extraction using SVM And ICA classifiers. *ICICES, IEEE* 2014; 1-7.
15. Patil S, Pawar M. Quality advancement of EEG by wavelet denoising for biomedical analysis. *ICCICT, IEEE*, 2012; 1-6.
16. Olund T, Duun-Henriksen J, Kjaer TW, Sorensen HB. Automatic detection and classification of artifacts in single-channel EEG. *Conf Proc IEEE Eng Med Biol Soc* 2014; 2014: 922-925.
17. Ting W, Guo-zheng Y, Bang Y, Hong S. EEG feature extraction based on wavelet packet decomposition for brain computer interface. *Measurement* 2007; 41: 618-625.
18. Zhang Y, Ji X, Liu B, Huang D, Xie F, Zhang Y. Combined feature extraction method for classification of EEG signals. *NCA, Springer*, 2016; 1-9.

***Correspondence to:**

Ashish Aggarwal
Thapar University
Patiala, Punjab 147004
India