ABSTRACT
Blind Source Separation (BSS) is an approach to extract the meaningful data from the non-Gaussian independent element of the combined sources. The count and the mixing of pattern from the different sources are not known and hence the name ‘blind’. Joint blind source separation (JBSS) algorithm is beneficial to get common sources at a time exist across multiple datasets like ElectroEncephaloGram (EEG). In this research work, extract of the signal from expected independent elements using an effective algorithm is presented. It is compared with several other BSS algorithms like STFT, ICA, EEMD and IVA. This analysis also helps to early diagnosis of neurological diseases such as brain hypoxia, epilepsy, sleep disorders, and Parkinson’s disease etc. The observational results have higher SNR and Average Correlation Coefficient (ACC) values for the proposed algorithms compared to other BSS techniques.

Keywords: Blind Source Separation (BSS), ElectroEncephaloGram (EEG), ICA, SNR, ACC.

INTRODUCTION
Blind Source Separation techniques are the most beneficial and common method in signal processing. In the area of multichannel recording many techniques of BSS are introduced which work accurately in contrast to the multichannel recording in the single channel measurement.\(^2\) In the area of biomedical signals, independent sources are frequently blended together with the measured signal. Our job is then to separate contribution sources in order to have a nearer look at the signal of interest. In multichannel recordings such as EEG this problem is efficiently managed by using blind source separation techniques to sort the given mixed signal into an original non-mixed sources.\(^3\) On EEG techniques, sensors are pointed at the head surface and large number sources are active during each human action. There have no estimate about the origins or the mixing process of EEG signals in the cognitive system. Therefore brain signal analysis to be regaled as a BSS problem.\(^1\) Independent Component Analysis (ICA) is a statistical technique for recovering statistically independent sources of mixed data.\(^4\) In order to improve the accuracy and the stability of BSS, the family of ICA algorithms were implemented to extract the Independent Components. There are also many criteria to extract independent component and the FASTICA algorithm is one of the most well-known and popular method.\(^5\) The algorithm is established on a fixed point iteration scheme and maximizing Non-Gaussianity of the component. To recognize the source signal, many methods are proposed by the researchers which are classified based on the field of application. The time domain BSS method prominently encounters signal attenuation and permutation problem. Using Fourier transforms to the time domain convolutive mixture, commutes to an instantaneous mixture problem in the frequency domain.\(^6\) In the frequency domain BSS method seldom happens signal ambiguity (i.e., Scaling and Permutation), to overcome this problem use time-frequency domain methods, such as STFT and Wavelet FASTICA. Wavelet FASTICA breaks down the signal into different frequency bins.\(^7\) Very recently, independent vector analysis (IVA), as an extension of ICA from one of multiple data sets, has drawn increasing attention. IVA was originally designed to address the permutation problem in the frequency domain for the separation of acoustic sources.\(^15\) IVA was formulated as a general JBSS framework to ensure that the extracted sources are independent within each dataset and correlated well across multiple information sets.\(^10\) Empirical mode decomposition (EMD) is a worthy alternative for this determination. EMD is a single-channel technique that decomposes a non-stationary and nonlinear
time series into a finite act of intrinsic mode functions (IMFs).\textsuperscript{17} Compared with other decomposition methods (e.g., wavelet transform), EMD is completely data driven, meaning that it breaks down a signal in a natural manner without needing a prior knowledge.\textsuperscript{18} It has been proven to be effective in many biomedical applications, e.g., denoising electrohysterogram (EHG) signals and EEG signals.\textsuperscript{19,20} Nevertheless, the original EMD algorithm is extremely sensitive to noise and it causes mode mixing. Lately, a noise-assisted version of the EMD, called ensemble EMD (EEMD), was proposed and has been shown to be more rich in real-life applications.\textsuperscript{21} In EEG signal acquisition there is a probability of getting both Uni and multidimensional data, to separate these different types of dimensional data concurrently EEMD-IVA technique was mainly used.\textsuperscript{14} But it has a limitation of poor converging rate which is rectified by the proposed method. We have examined the operation of Proposed Algorithm algorithms and several other BSS methods on both synthetic data and real EEG data. The convergence rate of the proposed method is faster than the normal EEMD-IVA. We first validate it on simulated data, then employ it to the real EEG data collected from the patients. Lastly, we compare the operations of each method with the help of ACC and SNR.

**EEG SIGNAL AND ELECTRODE SYSTEM**

EEG signals are usually acquired using scalp electrodes; it is placed according to the 10-20 international electrode system depicted in Figure 1 (C). The “10” and “20” refer to the percentage of the distance between the landmark points namely, the Nasion, the Inion, and the Preaurical points, as shown in Figure 1(A) and (B), used to draw the lines at which intersections, the electrodes are placed. In other words, given the landmark points, the electrode positioning is established by looking at the intersections between communication channels which are surgically and carnally drawn, spaced at 10 or 20% of the space between the landmark points. Since the early research on EEG analysis, it has been observed that the areas of a healthy human cortex have their own intrinsic rhythms in the range of 0.5 – 40Hz. In general, five main rhythms can be found from an EEG recording: Delta (δ) 0.5–4Hz, Theta (θ) 4–8Hz, Alpha (α) 8 – 14Hz, Beta (β) 14 – 30Hz and Gamma (γ) over 30Hz. The amount of action in different EEG frequency bands is quantified employing spectral analysis techniques.\textsuperscript{8}

**EXISTING METHODS**

During the seventies, EEG analysis implied interpreting the EEG waveform using descriptive and heuristic methods.\textsuperscript{9} In time, various methods have been employed to analyze several subtle changes in the EEG signal.

**Short Term Fourier Transform- Independent Component Analysis (STFT-ICA)**

The time - frequency response of EEG signal is efficiently extracted by STFT-ICA. In this method windowing techniques are applied to the input signal. The STFT ICA algorithm steps are described below.\textsuperscript{10}

- Generate mixed signal $x_{i=1,2,3}(t)$ from dataset $s_{i=1,2,3}(t)$
- Obtain $X(f, \tau)$ by taking STFT of $x_i(t)$.
- Combine time and frequency information $X_{ft}$.
- Take FASTICA on $X_{ft}$ , to find mixing matrix $A$.
- Calculate demixing matrix $W= A^{-1}$.
- Recover the time domain response by computing the FT and HT (Hilbert transform) or FT-LT (logarithm-decrement technique) using demixing matrix

**Wavelet ICA**

Wavelet decomposition is a time invariant, the
phase relationship does not change after the decomposition of the input signal. As a result, there is no time delay introduced in this operation, thus it simplifies the algorithm significantly. So after the mixed signal separation wavelet ICA keeps the time and frequency information as it is.

- **Wavelet decomposition**: Apply DWT to every channel of the recordings to obtain non-overlapping spectra.
- **Identification and selection**: To identify and select only the details that contains frequency components in a specified range.
- **Preprocessing**: Preprocessing step whitening is performed in order to reduce the data dimensionality and to lighten the computational charge.
- **ICA**: Apply the ICA algorithm to the selected details.
- **Demixing**: Compute demixing matrix of the selected frequency components.
- **Wavelet reconstruction**: Perform the wavelet reconstruction using the non selected details and the cleaned details after ICA step and obtain the artifacts removed separated signal.

### Ensemble Empirical Mode Decomposition (EEMD)

Empirical mode decomposition (EMD) is a worthy alternative for BSS determination. EMD is a single-channel technique that decomposes a non-stationary and nonlinear time series into a finite act of intrinsic mode functions (IMFs). Compared with other decomposition methods (e.g., wavelet transform), EMD is completely data driven, meaning that it breaks down a signal in a natural manner without needing prior knowledge. The IVA steps are described below.

- Identify all extrema of the mixed input signal $x(t)$.
- Find maximum and minimum of $x(t)$.
- Calculate the mean $m(t) = \frac{\text{max}[x(t)] + \text{min}[x(t)]}{2}$.
- Extract the details $d(t) = x(t) - m(t)$.
- Iterate on the residual $m(t)$
- Fix the standard deviation (0.3-0.4).
- Extract IMFs and calculate the ensemble average.

### Independent Vector Analysis (IVA)

Independent vector analysis is an extension of ICA, it deals with multivariate sources. In the frequency domain the separated signals are swapped which leads to permutation ambiguity to overcome this IVA is mainly used. The IVA model consists of a set of standard ICA models. The univariate sources across different layers are dependent such that can be aligned and grouped as a multivariate variable. The IVA steps are described below.

- **Preprocessing**: Whiten the input signal $X_f$, to make it as an un-correlated signal (mean=0, variance=1).
- **Mutual information**: Find $I(y) = D(f_j \| \prod_{i} f_{y_i})$, using the contrast function and determine the source distribution using information geometry.
- **Complex variable**: Assume $Z = u + jv$, where $j = \sqrt{-1}$.
- **Contrast function**: Spherically symmetric exponential norm distribution (SEND), Gaussian or Laplacian contrast function is used.
- **Contrast optimization**: Newton’s rule/ Gradient descent rule is used for optimization.
- **Find the demixing matrix**: For each frequency bin, calculate $W$ to separate the signal.

### PROPOSED ALGORITHM

If the input signal having both Uni and multidimensional dataset means current existing algorithms are failing to handle this situation. Hence, joint Enhanced EEMD and IVA algorithms proposed to solve this problem. This proposed method gives better ACC and SNR value compare to normal EEMD-IVA algorithm by obtaining an easy separation of the different dimensional EEG signals. The algorithm steps are same as mentioned earlier only small corrections in the contrast function and ensemble average. Instead of the Gaussian contrast function use LoG (Laplacian of Gaussian) contrast function, which gives very fine and accurate separated signals. On normal EEMD algorithm, ensemble average is taken at the end of IMFs decomposion, but this enhanced algorithm ensemble average is taken at each step to get accurate IMFs. So combine these two enhanced algorithms to get high ACC and SNR value. The Proposed Algorithm flow is shown in figure 2.
The EEG signals are non-Gaussian in nature, it is verified with the help of statistical measures. If the mean value is zero and variance is constant, then the signal is Gaussian. Skewness returns the symmetrical measurement. A negative value of skewness indicates that the left side of the probability density function is longer than the right side. The positive value of skewness indicates that the right side of the probability density function is longer than the left side. In kurtosis, normal distribution has a value of three. A kurtosis value of less than 3 indicates a flatter distribution than normal. The kurtosis value of greater than 3 indicates a sharper distribution than normal.13 Eight regions of EEG signals were investigated and tested for non-gaussianity, and their measures are given in table 1.

Existing BSS methods like STFT, Wavelet ICA, EEMD and IVA algorithm results are compared by the means of waveforms, SNR and average correlation coefficients (ACC). First, the algorithm is applied for generating test signals, then followed by EEG signals.

As an example, consider three generated signals and these signals are mixed with a mixing matrix $A$.

The following three source signals were seen:

$$s_1 = 1.5\cos(0.01t)\sin(0.5t)$$

(1)

$$s_2 = 1.5\sin(0.025t)$$

(2)

$$s_3 = 1.5\sin(0.025t)\sin(0.2t)$$

(3)

Where $s_i$ is a simulated source and it varied from $i=1, 2, 3$ and $t$ is a number of samples ($T=1000$).

These three simulated sources, as shown in Figure 3. Note that here $s_i$’s are row vectors. The mixed datasets, were brought forth as follows, with each column denoting one observation in their respective data space.

$$X = A \cdot S$$

(4)
Where $s = [s_1; s_2; s_3]$ With 
\[
A = \begin{bmatrix}
0.2590 & 0.3264 & 0.6512 \\
0.4522 & 0.1219 & 0.7194 \\
0.0855 & 0.9133 & 0.9203
\end{bmatrix}
\]
(5)

Where $A$ is a mixing matrix and $X$ is a mixed signals. The Figure 3 shows the separated signals in all the types of BSS techniques. In STFT one particular size of the time window is selected for all the frequencies, which restricts the flexibilities of the input signal, but wavelet ICA is flexible in all the signals so Wavelet ICA decomposes the signal it may be any form weather it is a one or multi dimensional signal. But it needs prior information of the input signal for decomposition. IVA is applied for each frequency bin of the mixed sources, it will give the demixing matrix of all the frequency bin and also it remove the permutation ambiguity. Compare to all BSS techniques IVA will give better SNR and correlation coefficient value. Likewise the same BSS algorithm is put on for EEG signals and the results recorded in Figure 3.

All these methods deal with multidimensional signals except wavelet ICA, if both Uni and multidimensional signals are getting concurrently the EEMD – IVA method is employed to sort out the signals normally this pattern came in brain signal acquisition. But this EEMD-IVA method has some disadvantages the speed of the convergence rate is very slow and it will give an Average correlation coefficient value around 0.8 this is insufficient value to diagnose some neurological disorders, the separated signals using EEMD-IVA is shown in Figure 6. So Proposed Algorithm method is proposed to achieve the better ACC and SNR value. EEMD method is to sort out the one-dimensional signal into a finite number of IMFs then apply an EIVA algorithm to tell apart the mixed signals.

As an example, consider five generated signals and these signals are mixed with a mixing matrix $A$

The following five source signals were seen:
\[
s_1 = 1.5 \cos(0.01t)\sin(0.5t)
\]
(6)
\[
s_2 = 1.5 \sin(0.025t)
\]
(7)
\[
s_3 = 2\cos(0.08t) \sin(0.006t)
\]
(8)
\[
s_4 = 1.5\sin(0.025t)\sin(0.2t)
\]
(9)
\[
s_5 = 1.5\sin(0.2t))
\]
(10)

Where $s_i$ is a simulated source and it varied from $i=1, 2, 3, 4, 5$ and $t$ is a number of samples ($T=1000$). It is shown in Figure 4. Note that here $s_i$’s are row vectors. The mixed datasets, were brought forth as follows, with each column denoting one observation in their respective data space.

$X[n] = A[n] \cdot S[n], n = 1, 2, 3 \quad (11)$

Where $S[1] = [s_1; s_3; s_2]$ , $S[2] = [s_2; s_4; s_5]$ , $S[3] = [s_1; s_3; s_4]$ With

\[
A[1] = \begin{bmatrix}
0.2590 & 0.3264 & 0.6512 \\
0.4522 & 0.1219 & 0.7194 \\
0.0855 & 0.9133 & 0.9203
\end{bmatrix}
\]
(12)

\[
A[2] = \begin{bmatrix}
0.3598 & 0.6248 & 0.7426 \\
0.3821 & 0.5749 & 0.8063 \\
0.5320 & 0.9358 & 0.2793
\end{bmatrix}
\]
(13)

\[
A[3] = [0.8945 & 0.7831 & 0.2763]
\]
(14)

![Figure 4. Mixed uni and multidimensional EEG signal.](image-url)
In this segment, the EEEMD - IVA method is applied to the synthetic data. This simulation is used to demonstrate the specific procedure and the source separation effect of the EEEMD-IVA. In Figure 4, present the mixed datasets generated according to equ (14). Since X3 was one-dimensional, to apply EEEMD to X3 for decomposition and obtain a lot of averages IMFs, as indicated in Figure 5. Then the each data set was multivariate, and to remove irrelevant redundant information across the multivariate mixed generated signal datasets.

Hence the multi-LV method to extract subLVs from each dataset. It is noted in the beginning, these subLVs could carry as much varied information as possible within each dataset and meanwhile be correlated as highly as possible across data sets. Some other significant attribute of these subLV’s is that the subLV’s within each dataset were interrelated with each other.
Nevertheless, multi latent variable may not be capable to totally recover the underlying sources. So, finally, IVA was performed to these extracted sub-LVs and helped to attain the end of the JBSS technique. The recovered sources are shown in Figure 7, from this to understand that the roots of each dataset have been accurately recovered and the corresponding sources are highly correlated across data sets. Mention that the EEEMD-IVA method recovered all underlying sources of the unidimensional dataset X3. Likewise, this method is used in real time mixed both Uni and multi dimensional EEG signals. Table 2 shows the comparison results of existing and proposed algorithm for various metrics.

Table 2. Comparison results

<table>
<thead>
<tr>
<th>Metric</th>
<th>Separated signal</th>
<th>STFT ICA</th>
<th>Wavelet ICA</th>
<th>IVA</th>
<th>EEEMD-IVA</th>
<th>Proposed Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACC(Generated Signals)</td>
<td>X1</td>
<td>0.5516</td>
<td>0.5902</td>
<td>0.6829</td>
<td>0.8239</td>
<td>0.9147</td>
</tr>
<tr>
<td></td>
<td>X2</td>
<td>0.7204</td>
<td>0.7236</td>
<td>0.7928</td>
<td>0.8569</td>
<td>0.8949</td>
</tr>
<tr>
<td></td>
<td>X3</td>
<td>0.5824</td>
<td>0.7623</td>
<td>0.7922</td>
<td>0.8421</td>
<td>0.9021</td>
</tr>
<tr>
<td>ACC(EEG Signals)</td>
<td>X1</td>
<td>0.4892</td>
<td>0.5896</td>
<td>0.7346</td>
<td>0.8123</td>
<td>0.8721</td>
</tr>
<tr>
<td></td>
<td>X2</td>
<td>0.3756</td>
<td>0.4832</td>
<td>0.5928</td>
<td>0.8062</td>
<td>0.8635</td>
</tr>
<tr>
<td></td>
<td>X3</td>
<td>0.2267</td>
<td>0.4548</td>
<td>0.6321</td>
<td>0.8259</td>
<td>0.8894</td>
</tr>
<tr>
<td>SNR Value for Generated Signals</td>
<td>X1</td>
<td>28.2016</td>
<td>28.9509</td>
<td>32.3696</td>
<td>39.6702</td>
<td>45.5391</td>
</tr>
<tr>
<td></td>
<td>X2</td>
<td>34.4792</td>
<td>34.6353</td>
<td>35.4191</td>
<td>41.0856</td>
<td>44.9852</td>
</tr>
<tr>
<td></td>
<td>X3</td>
<td>28.4946</td>
<td>30.0236</td>
<td>35.4054</td>
<td>40.2754</td>
<td>45.2116</td>
</tr>
<tr>
<td>SNR Value of EEG Signals</td>
<td>X1</td>
<td>27.9894</td>
<td>28.8772</td>
<td>34.8421</td>
<td>39.6702</td>
<td>45.1427</td>
</tr>
<tr>
<td></td>
<td>X2</td>
<td>23.7251</td>
<td>27.8352</td>
<td>29.5642</td>
<td>37.9638</td>
<td>42.8631</td>
</tr>
<tr>
<td></td>
<td>X3</td>
<td>20.2414</td>
<td>27.1214</td>
<td>31.9894</td>
<td>39.0381</td>
<td>43.5491</td>
</tr>
</tbody>
</table>

CONCLUSION
EEG signals can be applied effectively to examine the mental states and ailments related to the mind. The inherent issues with the EEG signal are that it is highly nonlinear and non Gaussian in nature and its visual interpretations are tedious and subjective prone to inter observer variations. To help researchers better analyze EEG signals, in this research presented various signal analysis techniques such as linear, frequency, time–frequency domain methods. Compared to all other methods EEEMD-IVA is working better it is handled both multi and unidimensional brain signal (EEG) concurrently, the average correlation coefficient and SNR values are higher compare to other methods.

REFERENCES
8. Patrizio Campisi, Daria La Rocca, Brain Waves for Automatic Biometric-Based
User Recognition, IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL. 9, NO. 5, MAY 2014.


