ICA With CWT and $k-$means for Eye-Blink Artifact Removal From Fewer Channel EEG

Ajay Kumar Maddirala and Kalyana C. Veluvolu, *Senior Member, IEEE*

Abstract—**In recent years, there has been an increase in the usage of consumer based EEG devices with fewer channel configuration. Although independent component analysis has been a popular approach for eye-blink artifact removal from multichannel EEG signals, several studies showed that there is a leak of neural information into the eye-blink artifact associated independent components (ICs). Furthermore, the leak increases as the number of input EEG channels decreases and leads to loss of valuable EEG information. To overcome this problem, we developed a new framework that combines ICA with continuous wavelet transform (CWT), $k-$means and singular spectrum analysis (SSA) methods. In contrast to the existing approaches, the artifact region in the identified eye-blink artifact IC is detected and suppressed rather than setting it to zero as in classical ICA. As most of the energy in the eye-blink artifact IC is concentrated in the artifact region, CWT and $k-$means algorithms exploits this feature to detect the eye-blink artifact region. Support vector machine (SVM) based classifier is finally designed for automatic detection of the eye blink artifact ICs. The performance of proposed method is evaluated on synthetic and two real EEG datasets for various EEG channels setting. Results highlight that for fewer channel EEG signals, the proposed method provides accurate separation without any neural information loss as compared to the existing methods.**

Index Terms—**Electroencephalogram (EEG), eye-blink artifact, independent component analysis (ICA), continuous wavelet transform (CWT), k−means clustering, singular spectrum analysis (SSA), support vector machine (SVM).**

I. INTRODUCTION

ELECTROENCEPHALOGRAM (EEG) signals manifest the electrical activity of the brain and are often used to diagnose the neurological disorders such as epilepsy, brain stroke and recently in applications like attention tracking in drivers [1]–[4]. In general, the measured EEG signals are often contaminated by several physiological artifacts, such as electrooculogram (EOG), eletromyogram (EMG) and other non-physiological artifacts. The influence of these artifacts in applications like brain computer interface (BCI) is also

studied in [5]. However, the recorded EEG signals are often contaminated by the EOG artifact than other artifacts and alters the signal characteristics [6]. As the eye-blink is an uncontrollable and involuntary activity and often occurs (once in every $5s$ [7]) than eye-movement, we hereafter refer to the EOG artifact as an eye-blink artifact. Removal of eye-blink artifact has been a challenging task, as they overlap with the frequency spectrum $0 - 12Hz$ of EEG signal [8]. The usage of traditional low-pass or band-pass filter for eye-blink artifact elimination may also remove the valuable components of the actual EEG signal. Therefore, accurate removal of eye-blink artifact from the EEG signals is necessary for better understanding of the brain function.

The adaptive filters have been applied to remove eye-blink artifact from EEG signals [9]. However, they require a reference signal that some how correlates with the eye-blink artifact in the EEG signal. An artifact subspace reconstruction (ASR) method has been proposed, which is similar to principle component analysis (PCA), to remove artifacts from multichannel EEG data [10]. Like in PCA, high variance components are rejected based on the pre-set cutoff parameter. To reject the artifact components, in this method, the clean regions of the EEG data are used as reference to determine the threshold. The optimal selection of ASR's cutoff parameter has been studied extensively in [11].

Independent component analysis (ICA) method is a statistical based filter that has been successfully employed to remove eye-blink artifacts from the multichannel EEG signals [12], [13]. Unlike ASR method, where the variance of the EEG data is considered to estimate the uncorrelated components, ICA method exploits the higher order statistics (HOS) of the mixed EEG data to estimate the independent components (ICs) [13].

In general, the performance of ICA in estimating ICs will be good with more number of EEG channels. To enhance the performance of eye-blink artifact removal, ensemble empirical mode decomposition (EEMD) technique [14] is jointly used with ICA [15], [16]. In EEMD-ICA method, each EEG channel data is mapped into k intrinsic mode functions using EEMD. Such transformation increases the input EEG data size to k times the number of EEG channels, which will be applied to ICA as input for the faithful separation of eye-blink artifact. In addition to this, the integration of wavelet transform with ICA (w-ICA) has been proposed in [17]–[22] to enhance the performance of artifact removal from the EEG data. In wavelet based ICA methods, the wavelet transform is applied either to decompose the EEG data into components before employing
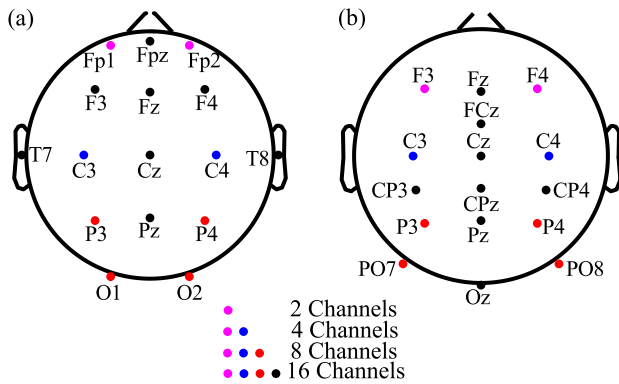
Fig. 1. Placement of electrodes for 2, 4, 8 and 16 channels configurations for (a) Synthetic and ERP-BCI data-1 and (b) covert and overt ERP-BCI data-2.

ICA [20], [21] or it is used to denoise the eye-blink artifact ICs obtained from ICA.

Recently, the consumer based EEG devices with fewer EEG channels have been popular for BCI related applications [23], [24] and also in the study of cognition (emotion, attention and fatigue) for vehicular applications [25], [26]. In general, these consumer based EEG devices comprises of single or at most sixteen EEG channels. The removal of eye-blink artifact in fewer EEG channel data is far more challenging compared to multichannel EEG. The use of ICA for fewer EEG channels and benefits has been first studied in [27] and developed a moving average based ICA method, called MAICA, for efficient separation of eye-blink artifact from the fewer channels EEG data [28].

Most of the existing ICA based methods rely on identifying the eye-blink artifact associated ICs. The identified ICs are then set to zeros to reconstruct the corrected EEG signal. As there exist neural cerebral information in estimated eye-blink artifact ICs, setting them to zero automatically in the reconstruction step removes the valuable EEG information [17]. The loss of neural components increases as number of input EEG channels to ICA decreases. Therefore, to preserve the valuable cerebral information, denoising of eye-blink artifact IC has to be performed rather than directly setting the component to zero. Only few methods have been proposed to denoise the estimated eye-blink artifact ICs [17], [19]. However, the existing methods fail to provide accurate separation of the eye-blink artifact without altering frequency spectrum of EEG signals under fewer channel configuration.

In this paper, we develop a new framework that combines ICA with continuous wavelet transform (CWT), $k-$means clustering and singular spectrum analysis (SSA) methods to denoise the eye-blink artifact IC without altering the original EEG information. In the proposed method, the identified eye-blink artifact IC is mapped into it's time-frequency representation. Next the columns of time-frequency matrix (each column is a feature vector) are clustered using $k-$means clustering algorithm. In general, the amplitude of samples in the eye-blink artifact region are high as compared to the samples in the non-artifact region. The proposed methodology exploits this feature to denoise the eye-blink artifact ICs. The novelty of the proposed method lies in combining CWT and

an unsupervised clustering algorithm ($k-$means) with ICA to denoise the eye-blink artifact without altering the EEG information. In addition, with simple two time-domain features of ICs, the support vector machine (SVM) [29] based classifier is developed to identify the eye-blink artifact ICs. The proposed methodology for eye-blink artifact removal is evaluated on synthetic and two real EEG datasets. Comparison with existing methods highlights the performance of the proposed approach.

The rest of the paper is organized as follows: the datasets used in this paper are discussed in in Section II. The frame work of the proposed method are discussed in Section III. The results and the discussion are presented in Section IV and V respectively. Finally, Section VI concludes the paper.

II. MATERIALS

In order to evaluate the performance of the proposed and the existing methods, we have considered one synthetic and two real EEG datasets: event related potential brain computer interface data (ERP-BCI) [30], [31] and covert and overt EEG [32], [33] datasets.

A. Synthetic EEG Data

For synthetic simulation study, we consider the resting state EEG data [34]. The EEG data is collected using 64 channel electrode cap with sampling frequency $256Hz$ from 10 subjects. For more details please refer [34]. For analysis in this paper, we only considered the EEG data corresponding to eyes-open task. To evaluate the performance of proposed method under fewer EEG channels setting, we have selected 16 channel EEG data from 64 channels raw EEG data. The 16 EEG channels considered for this study are: $Fp_1, Fp_z, Fp_2, F_3, F_z, F_4, T_7, C_3, C_z, C_4, T_8, P_3, P_z, P_4, O_1,$ and O_2, shown in Fig. 1(a). We have selected these electrodes to cover the total scalp regions such as pre-frontal, frontal, temporal, central, parietal and occipital. Before constructing the synthetic EEG data, we remove the dc drift and high frequency components from the data. Next, a band-pass filter is applied with cut-off frequencies 1 and $45Hz$. To assess the performance of the proposed technique over the existing methods, we have constructed 120 synthetically contaminated EEG signals. The synthetically contaminated EEG signals are constructed as follows: first, we have extracted a 10s artifact free EEG signals from a 16 channel lengthy EEG records and used as multichannel ground truth EEG data. A total of eighteen 10s artifact free multichannel EEG epochs were extracted (segmented) from six subjects EEG records.

Next, for constructing synthetic eye-blink artifact data, first, the eye-blink artifact region is identified manually and segmented from the EEG signal and zeros were appended to the segmented eye-blink artifact on both sides such that the signal length is 10s. The MATLAB smooth command has been used to remove the EEG remnants exist on the eye-blink traces and results ground truth eye-blink artifact. Based on the recommendations in [35], the amplitudes of the ground truth artifact is adjusted and stacked to construct the eye-blink artifact data matrix U (16 channels). We have
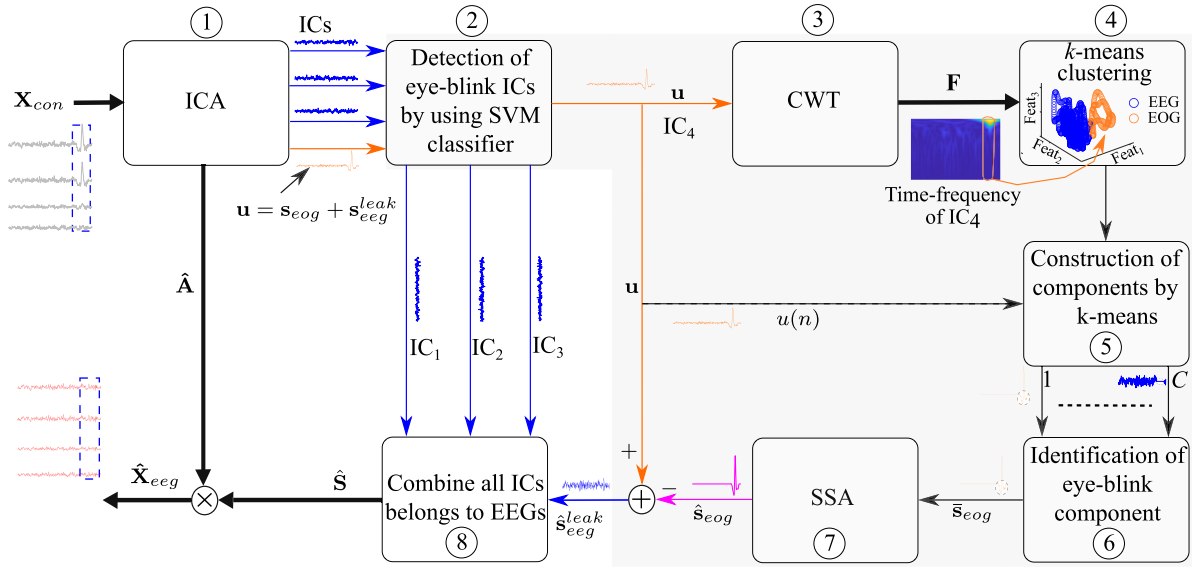
Fig. 2. Block diagram of the proposed framework.

constructed 10 such data matrices from five subjects. Using these data-sets and simple additive mixing model, we have constructed 180 ($= 18 \times 10$) synthetically contaminated multichannel EEG data $\mathbf{X}_{con}$. We have used 90 multichannel EEG epochs for build the SVM classifier and the remaining 90 EEG epochs are used for testing.

B. ERP-Based BCI EEG Data-1

The ERP-BCI dataset contains data from 12 subjects with sampling frequency $f_s = 2048Hz$. Each subject was asked to spell 20 characters using matrix speller and hence dataset has 20 trials per subject. For simulation study with real EEG signals, we have constructed 120 contaminated EEG signals as follows: first, before removing the dc drift and high frequency components from the raw EEG data, the EEG signals are down sampled to $256Hz$. Next, a band-pass filter is applied with cut-off frequencies 1 and $45Hz$. A 10s EEG epoch is segmented from lengthy EEG record of each subject such that at least one eye-blink artifact is present in the extract EEG epoch. We have constructed 120 EEG epochs of length 10s from 12 subjects EEG records. From this EEG dataset, sixty EEG epochs are used to build (train) SVM based classifier and the remaining 60 EEG epochs are used to test the classifier performance on real EEG data. For more details about the EEG data, please refer to [30], [31].

C. ERP-Based BCI Data-2

In order to evaluate the proposed method on lengthy EEG signals, we have considered publicly available covert overt EEG data. The covert and overt EEG data contains $P300$ evoked potentials of 10 healthy subjects recorded using two different paradigms. Using these two paradigms, the subjects attention in covert and overt conditions is studied. The EEG signals were recorded using 16 channels EEG device with sampling frequency of $256Hz$. The location of 16 EEG electrode on the scalp are F_3, F_z, F_4, FC_z, C_3, C_z, C_4, CP_3, CP_z, CP_4, P_3, P_z, P_4, PO_7, PO_8, and O_z

and shown in Fig. 1(b). In this study, each subject attended three recording sessions and each session comprises of six trials of 24s time duration for each condition. This results in total 36 trials (including both overt and covert conditions) per subject. As a result, a total of 360 EEG epochs were extracted from 10 subjects. However, we have selected a total of 120 eye-blink artifact affected EEG epochs from 360 trials. We have used 60 EEG epochs to build the classifier and 60 EEG trails were used for testing. Interested readers can refer to [32], [33] for more details about the EEG data.

III. PROPOSED METHOD

The proposed methodology to remove eye-blink artifact from fewer channel EEG data is described in Fig. 2. It comprises of three steps: (i) application of ICA to the contaminated EEG data, (ii) detection and (iii) denoising of ICs corresponding to eye-blink artifact. The novelty of the proposed method was highlighted in the shaded region. In what follows, the key steps are detailed in

A. ICA

Consider d number of statistically independent source components $\mathbf{s}^1, \mathbf{s}^2, \ldots, \mathbf{s}^d$, of length N samples each. These source signals were measured using d number of channels (assumed that the number of sources are equal to the number of channels). Then the measured EEG signals over the scalp are a linear mixture of sources and can be defined as

$$\begin{bmatrix} \mathbf{x}^1_{con} \\ \vdots \\ \mathbf{x}^d_{con} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1d} \\ \vdots & \vdots & \ddots & \vdots \\ a_{d1} & a_{d2} & \cdots & a_{dd} \end{bmatrix} \begin{bmatrix} \mathbf{s}^1 \\ \vdots \\ \mathbf{s}_d \end{bmatrix}$$

$$\mathbf{X}_{con} = \mathbf{AS} \qquad (1)$$

where $\mathbf{S} = [\mathbf{s}^1, \mathbf{s}^2, \ldots, \mathbf{s}^d]^T$ and $\mathbf{A}$ are the source and mixing matrices respectively. In fact, the source components and the mixing process information are unknown. The ICA algorithm will estimate the de-mixing matrix $\mathbf{W}$ ($\approx \mathbf{A}^{-1}$) from the mixed

EEG data $\mathbf{X}_{con}$. The ICs are extracted by multiplying the de-mixing matrix $\mathbf{W}$ with the contaminated EEG data $\mathbf{X}_{con}$ and given by

$$\mathbf{W}\mathbf{X}_{con} = \hat{\mathbf{S}} \approx \mathbf{S} \qquad (2)$$

After computing the ICs using (2), the IC corresponding to the eye-blink artifact, say $\mathbf{u}$, is identified automatically using SVM based classifier, will be discussed in Section III-B. In most of the traditional ICA methods [12], [13], identified eye-blink artifact ICs are set to zero. In fact, the remnants of EEG signal, we call it leaked EEG components ($\mathbf{s}_{eeg}^{leak}$), can be seen in the estimated eye-blink artifact ICs [17]. Setting them to zero in the reconstruction step of ICA will result in loss of original EEG information. Therefore, instead of setting the identified eye-blink artifact ICs to zero, we localize the eye-blink component region in the detected IC using time-frequency transform and $k-$means clustering algorithms.

The application of wavelet transforms and the selection of mother wavelets for analyzing EEG signals has been extensively studied in [36]–[38]. As we want to localize the eye-blink component (the artifact) region in the detected IC (eye-blink artifactual IC), in the proposed method, we map it into its time-frequency representation using CWT. Although several mother wavelets exists in-literature, the Morlet wavelet preserves the time and frequency resolution [39] and helps in localizing the eye-blink component in the detected IC. Hence, we employed CWT with Morlet wavelet to map the detected eye-blink artifact IC into its time-frequency representation. CWT with Morlet wavelet is a promising tool and recently, it has been also used to decode finger movement from the EEG signal [40]. To output the time-frequency representation of a signal, in CWT, the convolution of the given signal and wavelet will be performed. In each convolution step of CWT, the wavelet function is stretched or compressed by some quantity and results two-dimensional (time-frequency) representation of the signal. Hence, in this method, the identified eye-blink artifact IC is represented into its time-frequency domain using CWT. In fact, such representation maps each sample of the eye-blink artifact IC into a high dimensional feature vector. The $k-$means algorithm cluster (group) these features and provides the clustering information. With the clustering information we construct the partially denoised eye-blink artifact IC say $\bar{\mathbf{s}}_{eog}$, will be discussed in Section III C. The constructed eye-blink artifact IC say $\bar{\mathbf{s}}_{eog}$ is further processed using SSA [41] method to remove the EEG components superimposed on the eye-blink artifact region in the IC. Thus it results in the denoised eye-blink artifact IC $\hat{\mathbf{s}}_{eog}$. The denoised eye-blink artifact IC $\hat{\mathbf{s}}_{eog}$ is subtracted from $\mathbf{u}$ to retrieve the EEG information reside on the eye-blink artifact IC. The ICs associated to the EEG, $\hat{\mathbf{s}}_{eeg}$ and the residual EEG, $\hat{\mathbf{s}}_{eeg}^{leak} = \mathbf{u} - \hat{\mathbf{s}}_{eog}$ components are concatenated to form the source matrix $\hat{\mathbf{S}}$. Finally, the corrected EEG signal $\hat{\mathbf{X}}_{eeg}$ is obtained by multiplying the estimated source matrix $\hat{\mathbf{S}}$ and the mixing matrix $\hat{\mathbf{A}}$, which is inverse of de-mixing matrix $\mathbf{W}$. The procedure for the eye-blink artifact detection and the denoising is discussed in the following two subsections respectively.

B. Detection of Artifactual ICs

Detection of eye-blink artifact associated ICs is a critical step in ICA based artifact removal methods. Therefore, in this work, we have employed SVM based classifier to detect eye-blink artifact ICs automatically. For that, a pre-trained SVM based classifier is build as follows: first, the time-domain features, Hjorth mobility [42] and Kurtosis, of ICs extracted from the EEG data using ICA are computed. Next, ICs associated to eye-blink artifact and the EEG are manually detected and labelled as $+1$ and -1, respectively. Finally, using the computed features and the labelled information, the SVM classifier was built. For new EEG data (unseen EEG data), the pre-trained classifier will classify the IC belongs to the eye-blink artifact as positive class ($+1$) and the IC belongs to EEG as negative class (-1). The following section describe the denoising of artifactual IC that is classified as positive class ($+1$).

C. Denoising of Artifactual ICs

The shaded region in the block diagram shown in Fig. 2 describes the key steps involved in the denoising of the eye-blink artifact IC. Consider that $\mathbf{u}$ is an eye-blink artifact IC classified by SVM classifier and is defined as

$$\mathbf{u} = \mathbf{s}_{eog} + \mathbf{s}_{eeg}^{leak} \qquad (3)$$

where, $\mathbf{s}_{eog}$ is the true eye-blink artifact component and $\mathbf{s}_{eeg}^{leak}$ is the leaked EEG component present in the identified eye-blink artifact IC. In the denoising process of eye-blink IC, we try to extract the true eye-blink artifact $\mathbf{s}_{eog}$ from $\mathbf{u}$. The estimated eye-blink artifact $\hat{\mathbf{s}}_{eog}$ is then subtracted from $\mathbf{u}$, resulting the residual EEG component $\hat{\mathbf{s}}_{eeg}^{leak}$. To extract of eye-blink artifact $\hat{\mathbf{s}}_{eog}$ from $\mathbf{u}$, the identified N sampled eye-blink artifact IC, $\mathbf{u}$ should be mapped into its time-frequency representation. Using CWT, the identified eye-blink IC, $\mathbf{u}$ is mapped into its time-frequency representation and that results a matrix $\mathbf{F}$ of size $T \times N$ (feature matrix). Each column vector of $\mathbf{F}$ represent a feature vector of each sample of $\mathbf{u}$. In other words, the nth sample of $\mathbf{u}$ is represented as a feature vector $\mathbf{f}^n$ of size $T \times 1$. Next, the $k-$means clustering algorithm will be applied on the feature matrix $\mathbf{F}$ to cluster the N feature vectors into C clusters (or groups). After clustering the feature matrix $\mathbf{F}$, C number of uni-variate signals are derived based on the $k-$means clustering information. Using the $k-$means clustering information, the ith signal can be obtained as follows,

$$\hat{s}^i(n) = \begin{cases} u(n) & \text{if } \mathbf{f}^n \in Cluster^i, \ i = 1, 2, \ldots, C \\ 0 & \text{if } \mathbf{f}^n \notin Cluster^i, \ n = 1, 2, \ldots, N \end{cases} \qquad (4)$$

After decomposing the eye-blink artifact IC $\mathbf{u}$ into C (in this study, we set the number of clusters C to 2) number of signals using (4), the Hjorth mobility of each signal is computed. In general, the Hjorth mobility value is low for the eye-blink artifact associated component. Based on this, we identify the eye-blink artifact associated component from C signals and denoted it as $\bar{\mathbf{s}}_{eog}$. However, due to the operation in (4), we could see that the sample values of $\bar{\mathbf{s}}_{eog}$ changes from zero to non-zero value and vise versa at the onset

and the offset of eye-blink. Moreover, the EEG information superimposed on the eye-blink activity should be smoothed out, as $\bar{\mathbf{s}}_{eog}(n) = u(n)$ in the artifact region. Therefore, smoothing of the component $\bar{\mathbf{s}}_{eog}$ has to be performed.

D. SSA

SSA is a subspace based technique and it can decompose the given uni-variate time-series signal into low-frequency trend, the oscillating and the noise components. SSA has been used widely to process the EEG signals [43], [44]. Here, we employ SSA as a low-pass (smoothing) filter to remove any EEG remnants that reside on the eye-blink activity portion and also to smooth the edges at the onset and offset of eye-blink activity. This results in estimated eye-blink artifact IC $\hat{\mathbf{s}}_{eog}$. Finally, the residual EEG component $\hat{\mathbf{s}}_{eeg}^{leak}$ is obtained by subtracting the estimated eye-blink artifact IC from $\mathbf{u}$. This residual EEG component is concatenated with the ICs associated to EEG ($\hat{\mathbf{s}}_{eeg}$) and multiplied with the estimated mixing matrix $\hat{\mathbf{A}}$ (inverse of the de-mixing matrix $\mathbf{W}$) to obtain the corrected EEG signal $\hat{\mathbf{X}}_{eeg}$. In this paper, the SSA technique has been used to smooth the eye-blink artifact IC obtained after CWT and $k-$means operation.

E. Performance Measures

1) *Relative Root Measure Square Error (RRMSE):* This measure is often used to assess the performance of artifact removal methods on synthetic EEG data. The RRMSE between the two multichannel EEG data $\mathbf{X}$ and $\hat{\mathbf{X}}$, can be defined as

$$RRMSE = \sqrt{\frac{\frac{1}{Nc \times N} \sum_{i=1}^{N_C} \sum_{n=1}^{N} (\mathbf{X}(i,n) - \hat{\mathbf{X}}(i,n))^2}{\frac{1}{Nc \times N} \sum_{i=1}^{N_C} \sum_{n=1}^{N} \mathbf{X}^2(i,n)}} \times 100(\%)$$

(5)

where N_c and N represents the number of channels and the number of samples in each channel. $\mathbf{X}$ and $\hat{\mathbf{X}}$ represent the ground truth and the corrected EEG signals, respectively. The low RRMSE value indicates a good removal of artifact by the method.

2) *Correlation Coefficient (CC):* The CC measure is also used to evaluate the performance of an artifact removal technique on synthetic EEG data and indicates the strong relationship between the two signals. The CC between the ith channels of the true and the corrected EEG datasets $\mathbf{x}_i$ and $\hat{\mathbf{x}}_i$ can be defined as

$$CC_i = \frac{cov(\mathbf{x}_i, \hat{\mathbf{x}}_i)}{\sigma_{\mathbf{x}_i \mathbf{x}_i} \sigma_{\hat{\mathbf{x}}_i \hat{\mathbf{x}}_i}}$$

where $cov(\cdot)$ represents the co-variance between the two signals $\mathbf{x}_i$ and $\hat{\mathbf{x}}_i$ and $\sigma_{(\cdot)}$ variance of the signal itself. Then, the average CC value between the true and the corrected EEG signals is

$$CC = \frac{1}{N_C} \sum_{i=1}^{N_C} CC_i$$

(6)

The CC value close to one indicates a good removal of eye-blink artifact from the EEG data.

3) *Phase Locking Value (PLV):* PLV is a statistical measure often used to assess the synchronization of neural activity between two EEG channels [45], [46]. In this paper, we have used the PLV measure to account the phase changes between the contaminated and the corrected EEG signals at all frequency levels. Moreover, PLV is in-sensitive to the amplitude changes. The PLV between two EEG signals $\mathbf{x}$ and $\mathbf{y}$ at each frequency point can be defined as

$$PLV(f) = \frac{1}{N} \left| \sum_{n=1}^{N} exp(i\{\phi_x(f,n) - \phi_y(f,n)\}) \right|$$

(7)

where $\phi_x(f,n)$ is the instantaneous phase of the signal $\mathbf{x}$ and N is the number of samples. As the eye-blink artifact contaminates the EEG signal between $0-12Hz$ band the PLV value expected to be less than one in this band and should be equal to one above $12Hz$. The mean PLV (averaged over all channels) curve with respect to frequency is plotted in the simulation results with synthetic and real EEG datasets.

Moreover, statistical analysis was also performed to show the significance of results of the proposed method over the existing methods. As the data is non-Gaussian distribution (based on the Shapiro-wilk test), we performed non-parametric statistical test (Kruskal Wallis). For all statistical analysis the significance of statistical tests was set at $p < 0.05$.

IV. RESULTS

A. Parameter Selection for All Methods

For all ICA based methods, we employed *runica EEGLAB* function with default settings to extract the ICs from the mixed (contaminated) EEG data. For better removal of eye-blink artifacts, the parameters of the proposed and the existing methods are set as follows. The parameters for MAICA method, the correlation parameters such as $MinOriginal$, $MaxOriginal$, and $MaxOutputs$ are set to $0.96, 0.9$ and 0.9 respectively. The interested readers can refer [27], [28] for more details about the significance of MAICA parameters. We set the parameters of w-ICA (number of decomposition levels and the wavelet transform) based on the recommendation in [19]. In case of ASR method, the performance mainly depends on the selection of cutoff parameter. Hence, based on the recommendation in [11], the cutoff parameter of ASR method is set to 12 (for synthetic EEG data), 20 (for ERP-BCI data) and 10 (for covert and overt EEG data). In case of proposed method, the number of clusters is set to 2. Morlet wavelet transform is employed to represent the eye-blink artifact IC in time-frequency feature matrix. However, based on the recommendation in [39], the number of cycles of Morlet wavelet transform is selected to be 6. With a frequency resolution of $0.5Hz$, we have computed the wavelet coefficients for the frequency band from 1 to $45Hz$. Thus it results a time-frequency representation of eye-blink artifact IC of size $89 \times N$, where N is the number of samples of the IC. The obtained time-frequency matrix will be applied as input to the $k-$means clustering algorithm. As discussed in methods Section III-D, we have employed SSA to smooth the eye-blink artifact constructed by CWT and the $k-$means clustering algorithm. The window (embedding) length in SSA
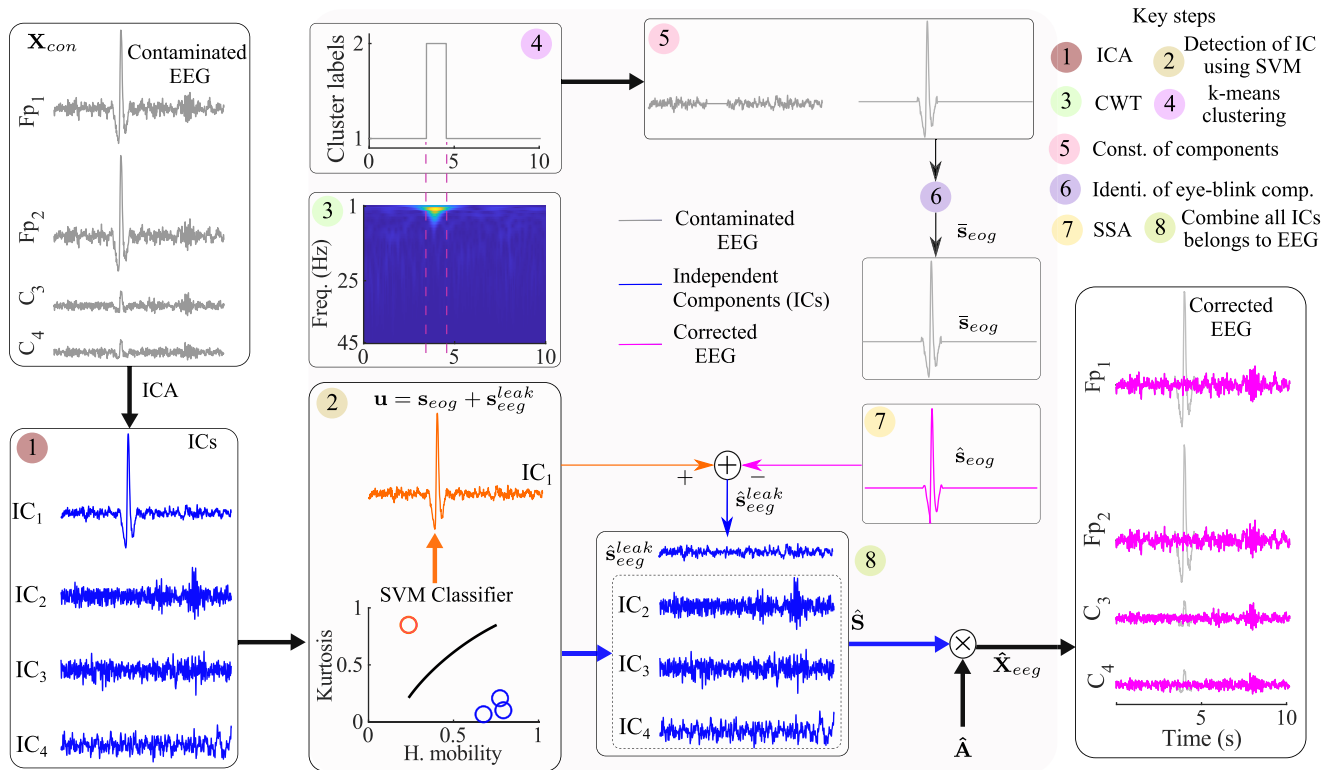
Fig. 3.   Graphical representation of eye-blink artifact removal from synthetically contaminated four channels EEG data using the proposed method.

is selected to be 32 and threshold is set to $8Hz$, as the most of the energy of eye-blink artifact is concentrated in $1 - 8Hz$ frequency band.

B. Building and Evaluating the Classifier Performance

In order to build the classifier that can identify the eye-blink artifact ICs, we have considered sixteen channels EEG data of length $10s$. After applying the ICA[1] algorithm to the 16 channel EEG data, we have identified the eye-blink artifact ICs manually and labeled them as positive class ($+1$) and the ICs associated to EEG are labeled as negative class (-1). The same procedure is performed for $2, 4$ and 8 EEG channels setting for all 90 EEG epochs (training data). This results in a total of $2700 = (2 + 4 + 8 + 16) \times 90$ ICs. Next, we compute two time domain features, Hjorth mobility and Kurtosis for all IC's. With the class labels and the features of ICs, the SVM based classifier (model) was built. The same procedure is employed to build SVM based classifier for MAICA method. However, in case of MAICA method, the input data size to ICA algorithm doubles due to its parameters. As result, the number of ICs accounted for building classifier for MAICA method are $5400 = (4 + 8 + 16 + 32) \times 90$. Similar procedure is employed to build the SVM classifier for the two real EEG datasets.

To build the model, classifier regularization parameter C_{SVM} and the kernel coefficient γ values were first identified by implementing a grid search using 10 fold cross-validation. Next, with the best C_{SVM} and γ_{SVM} values, the model

---

[1]We have used the EEGLAB Toolbox *runica* command to decompose the components from mixed EEG data.

---

was built using the training data. These models are used to detect the eye-blink artifact ICs from the ICA estimated source components. Note that for each EEG dataset, we build two models, one for ICA and other for MAICA methods. Table I shows the performance of the SVM based classifier in terms of accuracy, specificity and sensitivity for EEG records. We notice from our study that with two time domain features, the classifier accuracy and specificity are greater than 99% for ICs obtained by ICA and MAICA methods. In this study, we have used the LIBSVM Toolbox to build SVM based classifier [47].

We also conducted a study to detect eye-blink artifact associated ICs based on the threshold. To detect the artifact ICs, we used the lower limit of the 95% confidence interval (CI) of the mean for thresholding the ICs based on the Hjorth mobility and the upper limit of the 95% CI of the mean for thresholding the ICs based on the Kurtosis. All the ICs whose Hjorth mobility and the Kurtosis values lies below $T1$ and above $T2$, respectively are identified as artifactual IC. For more details about the setting the thresholds, please refer to [19]. Table II shows the performance of detecting the artifact ICs based on the threshold. We observed from this study that SVM based artifact ICs detection showed 0.54% and 0.68% improvement in accuracy and specificity over the threshold based artifact IC detection in ICA method. The SVM based artifact ICs detection for MAICA method showed 2.52% and 2.9% improvement in accuracy and specificity, respectively compared to the threshold based artifact IC detection. However, we haven't seen much improvement in the sensitivity of the SVM based classifier over the threshold based approach. We also observed that, the combined use of two
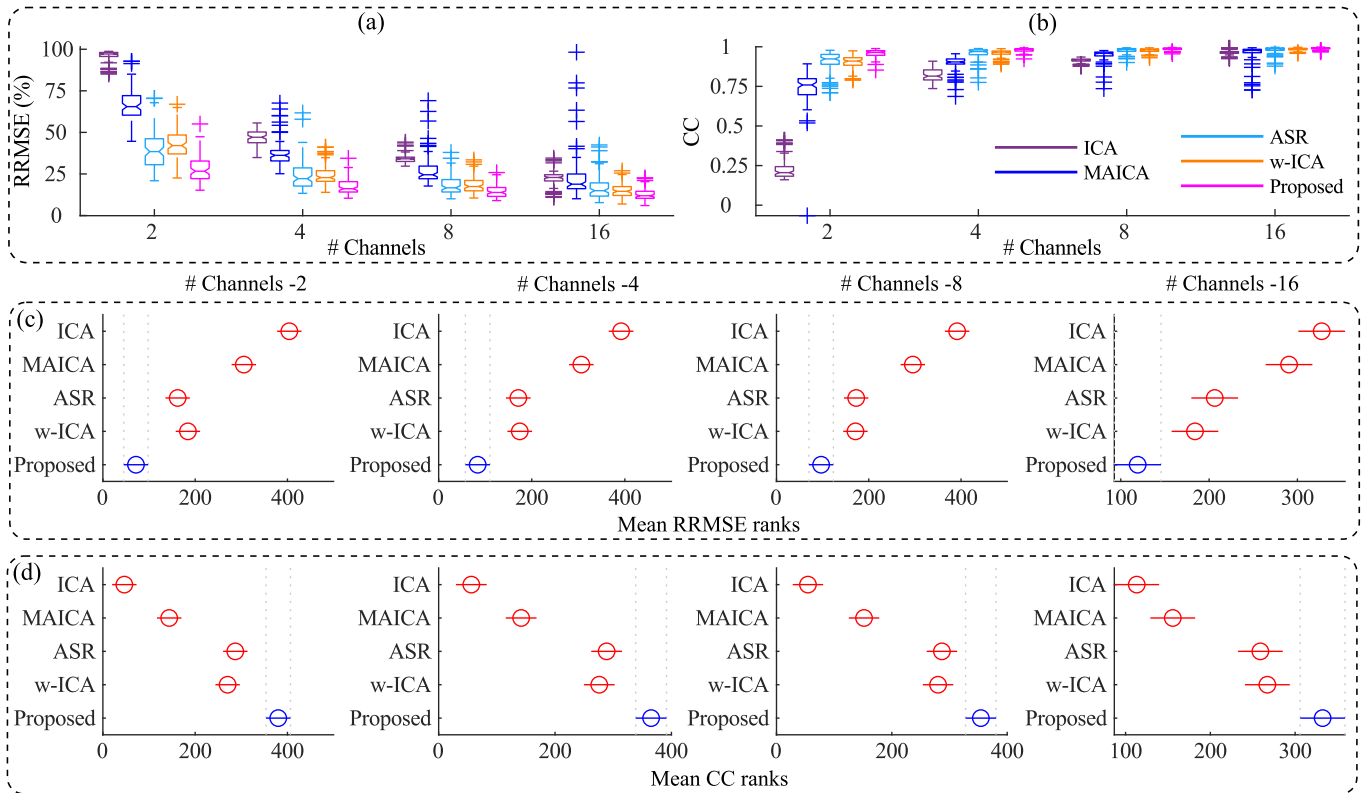
Fig. 4. (a) RRMSE and (b) CC values for different number of EEG channels. (c) and (d), respectively, the statistical analysis of RRMSE and CC values of existing and the proposed methods for $p < 0.05$.

TABLE I

SVM BASED CLASSIFIER PERFORMANCE IN DETECTING THE ARTIFACT ICS USING TWO TIME-DOMAIN FEATURES OF THE ICS

Methods	MAICA						ICA					
Features	H. mobility and Kurtosis			mMSE and Kurtosis [19]			H. mobility and Kurtosis			mMSE and Kurtosis [19]		
	Acc.	Spec.	Sens.	Acc.	Spec.	Sens.	Acc.	Spec.	Sens.	Acc.	Spec.	Sens.
Synthetic EEG Data	99.50	99.78	95.81	99.79	99.94	97.90	100	100	100	99.92	100	99.44
ERP-BCI Data1	99.66	99.67	99.58	98.30	98.54	95.04	99.77	100	98.34	99.22	99.35	98.34
ERP-BCI Data2	99.41	99.61	96.87	98.16	98.68	91.40	99.55	99.74	98.36	99.55	99.67	98.77
Average	**99.52**	**99.68**	**97.42**	98.75	99.05	94.78	**99.77**	**99.91**	**98.90**	99.56	99.67	98.85

time domain features, Hjorth mobility and Kurtosis, showed better performance in detecting the artifact ICs as compared to the existing modified multiscale sample entropy (mMSE) and Kurtosis time domain features.

C. Results With Synthetic EEG Data

We evaluated the performance of the proposed method on synthetic EEG data. Fig. 3 describes the graphical representation of removing the eye-blink artifact from synthetically contaminated EEG signals ($\mathbf{X}_{con}$) using the proposed method. First, the ICs are obtained from the four channel contaminated EEG data $\mathbf{X}_{con}$ using ICA. Next, by using the pre-trained classifier, the eye-blink artifact IC is detected based on the two time-domain features (Hjorth mobility and Kurtosis) of the ICs and is denoted as $\mathbf{u}$. After that the CWT of detected eye-blink artifact IC is computed and fed to the k−means clustering algorithm. Using the clustering labels and (4), two components are derived, as the number of clusters are set to 2. From these two components, the eye-blink artifact

associated component is identified based on their Hjorth mobility and denoted as $\bar{\mathbf{s}}_{eog}$. In fact, EEG information still resides on the eye-blink component (the artifact region) in $\bar{\mathbf{s}}_{eog}$. To filter out the EEG information from it, we applied SSA to $\bar{\mathbf{s}}_{eog}$, thus it results in $\hat{\mathbf{s}}_{eog}$ and it is subtracted from $\mathbf{u}$ to obtain $\hat{\mathbf{s}}_{eeg}^{leak}$. Finally, the obtained $\hat{\mathbf{s}}_{eeg}^{leak}$ is concatenated with the ICs associated to the EEG (IC_2, IC_3 and IC_4) and multiplied with the estimated mixing matrix $i.e.\hat{\mathbf{A}}$, which is inverse of de-mixing matrix $\hat{\mathbf{W}}$, to obtain the corrected EEG signal ($\hat{\mathbf{X}}_{eeg}$).

The proposed method is applied to the synthetically contaminated EEG signals (total of 90 records) to evaluate its performance in terms of RRMSE and CC. Fig. 4 (a) and (b) shows the comparison of the proposed and the existing methods in terms of RRMSE and CC, respectively for different EEG channels configurations. The mean RRMSE and CC values of the proposed method are superior as compared with the existing methods. We conducted Kruskal Wallis test to observe the significance of mean RRMSE and CC ranks
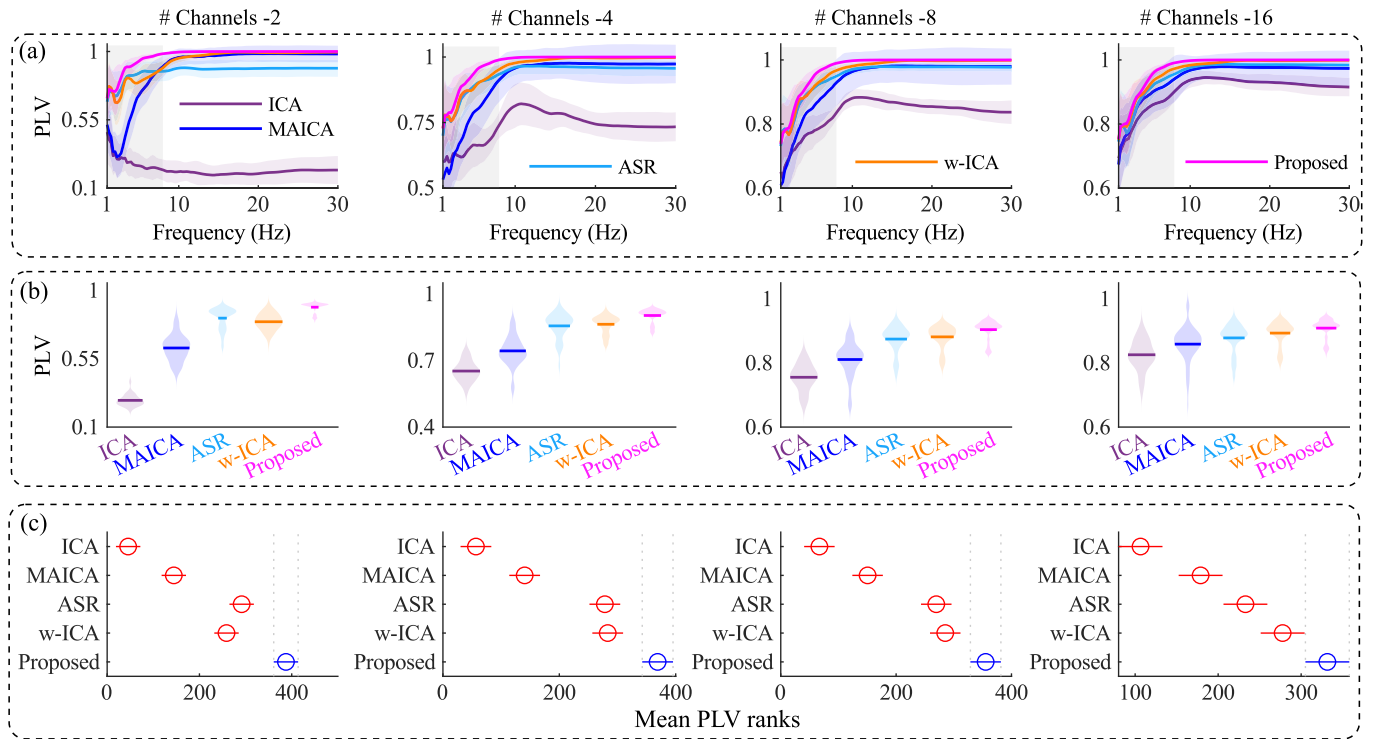
Fig. 5. (a) Performance of the proposed and the existing methods on Synthetic EEG data in terms of PLV vs frequency under different EEG channels setting. (b) the mean PLV values in the frequency band $1-8Hz$ (the shaded region in (a)) of each method in different EEG channels setting. (c) Statistical significance of mean PLV values ranks.

TABLE II

THRESHOLD BASED EYE-BLINK IC DETECTION PERFORMANCE USING TIME DOMAIN FEATURES OF ICs AND COMPARED WITH THE EXISTING APPROACH. HERE, T1, T2 AND T3 REPRESENTS THE THRESHOLDS FOR HJORTH MOBILITY, KURTOSIS AND mMSE, RESPECTIVELY

Methods	MAICA						ICA					
Features	H. mobility and Kurtosis			mMSE and Kurstosis [19]			H. mobility and Kurtosis			mMSE and Kurstosis [19]		
Synthetic EEG Data	Th1=0.2289, Th2=6.7991			Th3=0.9604, Th2=6.7991			Th1=0.3030, Th2=7.6156			Th3=1.0539, Th2=7.6156		
	Acc.	Spec.	Sens.	Acc.	Spec.	Sens.	Acc.	Spec.	Sens.	Acc.	Spec.	Sens.
	96.11	95.81	100	95.24	94.87	100	98.59	98.37	100	98.70	98.50	100
ERP-BCI Data1	Th1=0.4181, Th2=7.7501			Th3=0.9236, Th2=7.7501			Th1=0.3678, Th2=9.5584			Th3=1.0563, Th2=9.5584		
	Acc.	Spec.	Sens.	Acc.	Spec.	Sens.	Acc.	Spec.	Sens.	Acc.	Spec.	Sens.
	96.58	96.33	100	94.80	94.43	100	99.44	99.48	99.17	99.27	99.29	99.17
ERP-BCI Data2	Th1=0.2355, Th2=8.0501			Th3=1.0485, Th2=8.0501			Th1=0.2076, Th2=10.3524			Th3=1.2060, Th2=10.3524		
	Acc.	Spec.	Sens.	Acc.	Spec.	Sens.	Acc.	Spec.	Sens.	Acc.	Spec.	Sens.
	98.33	98.20	100	95.55	95.21	100	99.66	99.74	99.18	99.38	99.54	98.36
Average	**97.00**	**96.78**	**100**	95.19	94.83	100	**99.23**	**99.19**	**99.45**	99.11	99.11	99.17

of the proposed over the existing methods, as shown in Fig. 4(c) and (d). It can be noticed that the mean RRMSE and CC ranks of the proposed method are significantly different from the existing methods mean RRMSE and CC ranks.

Moreover, we have also evaluated the performance of proposed method over the existing methods in terms of phase changes in the frequency domain. Fig. 5 (a) shows the average PLV curves with respect to frequency for different EEG channel conditions. In fact, the eye-blink artifact is removed efficiently by all the methods. However, while removing the artifact, most of the existing methods alters the EEG low/high frequency components. As most of the eye-blink artifact energy is concentrated in $1-8Hz$ band, we have computed mean PLV values in this band for each EEG record. Fig. 5 (b) shows the mean PLV values in the frequency band

$1-8Hz$ (the shaded region Fig 5(a)) for different EEG channel condition. Note that, the mean PLV value of each EEG record is computed by adding all the PLV values in the band $1-8Hz$ and dividing it by number of frequency bins in that band. Fig. 5 (c) shows statistical significance of mean PLV ranks of the proposed method in the $1-8Hz$ band over the existing methods. The mean PLV ranks of the proposed method are significantly different from the mean PLV ranks of the existing methods.

D. Results With ERP-BCI Data-1

We have evaluated the performance of all the methods with ERP-BCI EEG dataset [30], [31]. The proposed and the existing methods are applied to the 4 channel EEG data
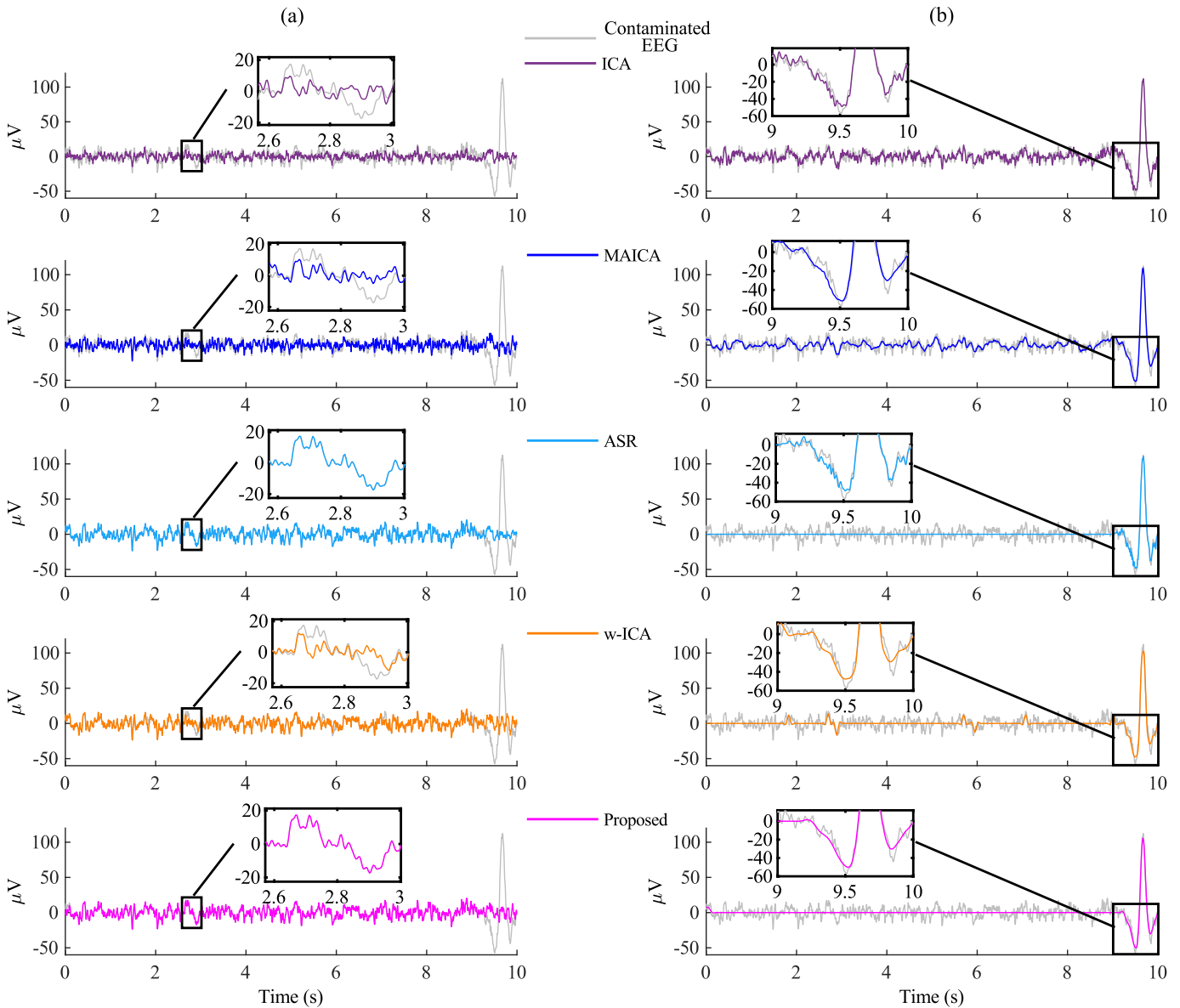
Fig. 6. (a) The corrected EEG signal ($\hat{\mathbf{X}}_{eeg}$) and (b) the estimated eye-blink artifact ($\hat{\mathbf{X}}_{Con}$-$\hat{\mathbf{X}}_{eeg}$) using all methods from Fp_1 channel in 4 channel configuration setting.

for performance evaluation. As the eye-blink artifact is more predominant in pre-frontal EEG channels, here, we only show the corrected EEG and the eye-blink artifact components from the Fp_1 EEG channel. Fig. 6(a) and (b) shows the corrected EEG signal and the eye-blink artifact component respectively using the existing and the proposed methods. The zoomed plots showed in Fig. 6(a) and (b) shows the affect of artifact removal methods in non-artifact and artifact regions of the EEG signal in time-domain.

It is evident from these plots that ICA, MAICA and w-ICA methods alters the non-artifact regions while removing the eye-blink artifact from the EEG data. Although ASR method doesn't affect the non-artifact regions of the signal, the EEG information superimposed on the artifact region is removed. In order to show the EEG signals spectral changes caused by artifact removal methods, we have applied the proposed and the existing methods on 60 real EEG ERP-BCI datasets and

computed the PLV between the contaminated and the corrected EEG signals at each frequency. The average PLV plots of each method (in different EEG channel configurations) with respect to the frequency are shown in Fig 7(a).

For all existing methods, the spectral changes can be clearly seen in the PLV plots after $8Hz$, where the PLV values are less than 1. Whereas the mean PLV values obtained by the proposed method are near to 1 after $8Hz$. Fig 7(b) shows the mean PLV values of each method in $1 - 8Hz$ frequency band (the shaded region Fig 7(a)) in different EEG channels setting. Fig 7(c) shows the statistical significance of mean PLV values of proposed method over the existing method on Dataset-1. When the proposed method is applied on the Dataset-1, with the exception of the ASR method for 2 channel case, the mean PLV values of the proposed method are statistically significant over all the existing methods.
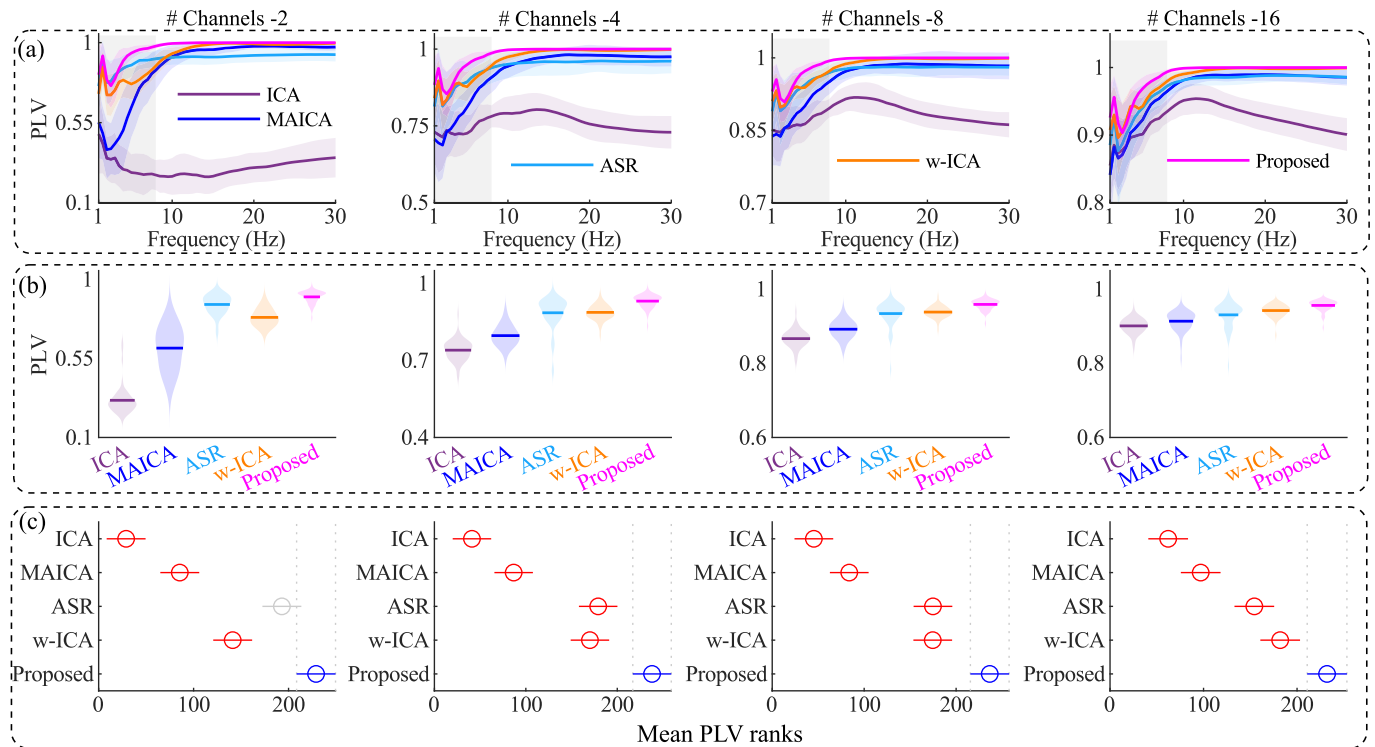
Fig. 7. (a) Performance of the proposed and the existing methods on ERP-BCI data in terms of PLV vs frequency under different EEG channels setting. (b) the mean PLV values in the frequency band $1 - 8Hz$ (the shaded region in (a)) of each method in different EEG channels setting. (c) Statistical significance of mean PLV values ranks.

E. Results With ERP-BCI Data-2

To further evaluate the performance of proposed method, we have also conducted simulation study with covert and overt EEG data under different EEG channels configurations setting. In this study, the performance of the proposed and the existing methods are applied on 60 EEG records of length $24s$. Fig. 8(a) shows the performance of the proposed and the existing methods in removing the eye-blink artifact in terms of PLV. For all EEG channel configurations, the PLV values obtained by the proposed method are near to one as compared with the existing methods. The mean PLV values in $1 - 8Hz$ band are also showed in Fig. 8(b). The mean PLV values of the proposed method in $1 - 8Hz$ band good as compared with the existing methods. Moreover, we have also conducted statistical significance of mean PLV values of proposed method, shown in Fig. 8(c). In case of Dataset-2, except for the ASR method with 2 and 4 channels cases, the mean PLV values of the proposed method are statistically significant over the existing methods.

F. Computation Time

We have also studied the computation time required for each algorithm with input EEG data by varying signal length and the number of EEG channels. All the algorithms were implemented in MATLAB R2020b on an Intel(R) Core(TM) i5-8600 CPU @ 3.19GHz, Windows 10 64-bit operating system, and 16GB RAM. Fig. 9(a) and (b) shows the average computational time for all algorithms when they are executed 1000 times with the input EEG data of $10s$ and $5s$

lengths, respectively. It is observed from Fig. 9(a) that the computation time of proposed method for the number of EEG channels ≤ 8 is high as compared with the existing methods. Even though the computation time of MAICA method is better than the proposed method for the number of EEG channels ≤ 8, we observe a drastic increase in the computation time when the number of EEG channels are greater than 8. Interestingly, it is observed that the ASR method is computationally efficient as compared with other artifact removal methods for input EEG data of $10s$ length (for the case of the number channels > 4). However, when the proposed method is applied to EEG data of length $5s$, we observed from Fig. 9(b) that the computation time of proposed method is less than MAICA and ASR methods for the number of channels ≥ 4. Although the computation time of w-ICA method is less than proposed method, the loss of valuable EEG information in the corrected EEG signal makes it undesirable.

V. DISCUSSION

In this study, the performance of the proposed method is evaluated on three datasets: one synthetic and two real EEG datasets. It was observed from the results that the performance of the ASR method is good compared to the other methods and similar with performance of proposed method (for few channels configuration only). However, the performance of ASR is dependent on the number of channels and it decreases as the number of channels increases. For example, the performance of the ASR method is superior than w-ICA method for the case of 2 and 4 channels configuration, however there is a decrease in performance for the case of 8 and 16 channel configurations
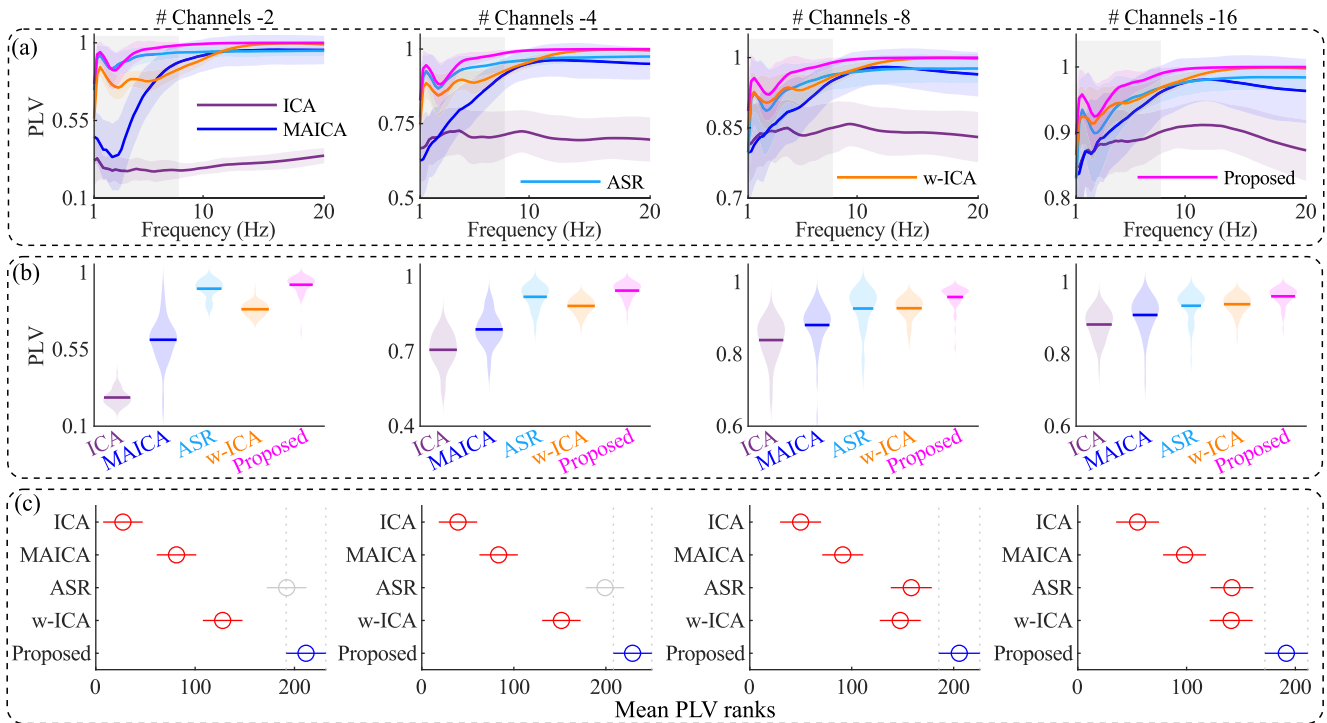
Fig. 8. (a) Performance of the proposed and the existing methods on Covert and overt ERP-BCI data in terms of PLV vs frequency under different EEG channels setting. (b) the mean PLV values in the frequency band $1 - 8Hz$ (the shaded region in (a)) of each method in different EEG channels setting (c) Statistical significance of mean PLV values ranks.
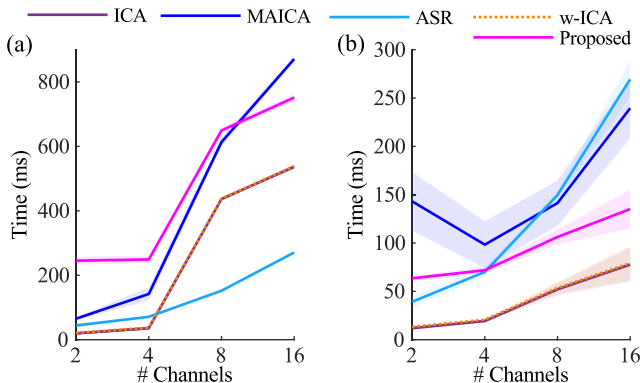

Fig. 9. Comparison of computational time of each method for EEG signals with (a) $10s$ and (b) $5s$ lengths. Note: the time complexity of ICA and w-ICA are almost equal.

(see Fig. 5(b), 7(b) and 8(b)). In contrast, the performance of proposed method is consistent irrespective of the number of EEG channels (Fig. 4, 5, 7 and 8). To validate the significance of the results obtained by the proposed method, we conducted the statistical analysis of the results. It was noticed that the results obtained by the proposed method are significantly different from the results obtained by the existing methods. Although we does not observe the significance between the results obtained by the proposed and the ASR methods for the case of 2 and 4 channel configurations, the mean PLV values of the proposed method are better than the mean PLV values of the ASR method.

Our analysis shows that the existing methods alters the EEG low/high frequency information while removing the eye-blink artifact from the EEG data. Anyway, the loss of EEG low/high

frequency information by the artifact removal technique can be problematic in both clinical and non-clinical applications. For example, detection of the limb movements in subject with spinal cord injury (SCI) [48], [49] and the attention/fatigue level of drivers in vehicle automation [50]. However, there is no such loss of EEG low/high frequency components after the removal of eye-blink artifact with the proposed method. The reason is that when the eye-blink artifact IC (in which the eye-blink component is dominant) is mapped into its time-frequency representation using CWT, the distribution of energy in the artifact region significantly differs from the non-artifact region. Such significant difference allows the $k-$mean clustering algorithm to localize the non-artifact and artifact regions in the eye-blink artifact IC (as shown in steps 3 and 4 in Fig. 3). As a result of these features, the proposed method denoises the eye-blink artifact IC more efficiently.

In general, the performance of the proposed method depends on how well the ICA estimates the eye-blink artifact component from the EEG data. In fact, this will further depend on the strength of the eye-blink artifact component present in the mixed EEG signal. When the strength of eye-blink artifact is relatively small, $k-$means clustering algorithm fails to distinguish the features representing the non-artifact region from the features of the artifact region. Therefore, the proposed method alters the non-artifact regions when the eye-blink component strength is small. From our simulation analysis, we understand that the proposed method can work efficiently in a condition where the eye-blink artifact is marginally strong in at least one of the EEG channels.

The EEG signals carry vital information. However, while designing the artifact removal method, it is important that

the method should process the given data without loosing valuable EEG information with less computation time. When accounting the computation time of each method, the proposed method showed higher computation time for fewer EEG channels. This is due to the mapping of detected eye-blink artifact IC into its time-frequency representation using CWT. In fact, the computational complexity of the artifact removal methods may not be burden with the recent advancements. Even in applications where the computational complexity is an issue, still the proposed method can be used by processing the EEG data shorter epoch by epoch, as the computational time of proposed method for shorter EEG epochs is less than ASR method. However, we do not see any improvement in the performance of proposed method when other transformation techniques such as, stock-well transform (SWT) are employed in place of CWT.

VI. Conclusion

This paper integrated CWT, $k-$means and SSA algorithms for accurate filtering of eye-blink artifactual IC without any loss of neural information. Furthermore, the SVM based classifier developed with two simple time domain features showed accuracy greater than 99% in the detection of eye-blink artifacts ICs. Performance of the proposed method is evaluated on one synthetic and two real EEG datasets under different EEG channels setting. Results show that unlike existing methods, the performance of proposed method in terms of mean RRMSE and CC in time domain and the mean PLV in the frequency band $1-8Hz$ is consistent for different number of EEG channels setting. Moreover, it was also observed that for shorter epochs, the computational time of the proposed method is less as compared to the existing MAICA and ASR methods for the number of EEG channels ≥ 4. From the analysis it can be observed that the proposed method can remove artifact efficiently if at least one EEG channel contains the eye-blink artifact where it is marginally strong, which is the limitation of the proposed method. The focus of the current study was only limited to eye-blink artifact removal. However, the selection number of clusters and identification of artifact component (step 6) forms the key steps of the proposed method and will be extended to multiple artifact removal, in our future work.

References

[1] S. Panwar, S. D. Joshi, A. Gupta, and P. Agarwal, "Automated epilepsy diagnosis using EEG with test set evaluation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 6, pp. 1106–1116, Jun. 2019.

[2] J. M. Antelis, L. Montesano, A. Ramos-Murguialday, N. Birbaumer, and J. Minguez, "Decoding upper limb movement attempt from EEG measurements of the contralesional motor cortex in chronic stroke patients," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 1, pp. 99–111, Jan. 2017.

[3] N. K. Al-Qazzaz, S. H. B. M. Ali, S. A. Ahmad, M. S. Islam, and J. Escudero, "Discrimination of stroke-related mild cognitive impairment and vascular dementia using EEG signal analysis," *Med. Biol. Eng. Comput.*, vol. 56, no. 1, pp. 137–157, Jan. 2018.

[4] Y.-K. Wang, T.-P. Jung, and C.-T. Lin, "EEG-based attention tracking during distracted driving," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 23, no. 6, pp. 1085–1094, Nov. 2015.

[5] A. Jafarifarmand and M. A. Badamchizadeh, "EEG artifacts handling in a real practical brain–computer interface controlled vehicle," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 6, pp. 1200–1208, Jun. 2019.

[6] K. A. Robbins, J. Touryan, T. Mullen, C. Kothe, and N. Bigdely-Shamlo, "How sensitive are EEG results to preprocessing methods: A benchmarking study," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 5, pp. 1081–1090, May 2020.

[7] D. Hagemann and E. Naumann, "The effects of ocular artifacts on (lateralized) broadband power in the EEG," *Clin. Neurophysiol.*, vol. 112, no. 2, pp. 215–231, Feb. 2001.

[8] S. Halder *et al.*, "Online artifact removal for brain-computer interfaces using support vector machines and blind source separation," *Comput. Intell. Neurosci.*, vol. 2007, pp. 1–10, Apr. 2007.

[9] P. He, G. Wilson, and C. Russell, "Removal of ocular artifacts from electro-encephalogram by adaptive filtering," *Med. Biol. Eng. Comput.*, vol. 42, no. 3, pp. 407–412, 2004.

[10] C. A. E. Kothe and T.-P. Jung, "Artifact removal techniques with signal reconstruction," U.S. Patent 1 489 544 0, Apr. 28, 2016.

[11] C.-Y. Chang, S.-H. Hsu, L. Pion-Tonachini, and T.-P. Jung, "Evaluation of artifact subspace reconstruction for automatic artifact components removal in multi-channel EEG recordings," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 4, pp. 1114–1121, Apr. 2019.

[12] T. P. Jung *et al.*, "Removing electroencephalographic artifacts by blind source separation," *Psychophysiology*, vol. 37, no. 2, pp. 163–178, Mar. 2000.

[13] 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, no. 4, pp. 1443–1449, Feb. 2007.

[14] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: A noise-assisted data analysis method," *Adv. Adapt. Data Anal.*, vol. 1, no. 1, pp. 1–41, 2009.

[15] K. Zeng, D. Chen, G. Ouyang, L. Wang, X. Liu, and X. Li, "An EEMD-ICA approach to enhancing artifact rejection for noisy multivariate neural data," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 6, pp. 630–638, Jun. 2015.

[16] G. Wang, C. Teng, K. Li, Z. Zhang, and X. Yan, "The removal of EOG artifacts from EEG signals using independent component analysis and multivariate empirical mode decomposition," *IEEE J. Biomed. Health Informat.*, vol. 20, no. 5, pp. 1301–1308, Sep. 2016.

[17] N. P. Castellanos and V. A. Makarov, "Recovering EEG brain signals: Artifact suppression with wavelet enhanced independent component analysis," *J. Neurosci. Methods*, vol. 158, no. 2, pp. 300–312, Dec. 2006.

[18] C. Y. Sai, N. Mokhtar, H. Arof, P. Cumming, and M. Iwahashi, "Automated classification and removal of EEG artifacts with SVM and wavelet-ICA," *IEEE J. Biomed. Health Informat.*, vol. 22, no. 3, pp. 664–670, May 2017.

[19] R. Mahajan and B. I. Morshed, "Unsupervised eye blink artifact denoising of EEG data with modified multiscale sample entropy, kurtosis, and wavelet-ICA," *IEEE J. Biomed. Health Informat.*, vol. 19, no. 1, pp. 158–165, Jan. 2014.

[20] N. Mammone, F. La Foresta, and F. C. Morabito, "Automatic artifact rejection from multichannel scalp EEG by wavelet ICA," *IEEE Sensors J.*, vol. 12, no. 3, pp. 533–542, Mar. 2011.

[21] S. Barua, M. U. Ahmed, C. Ahlstrom, S. Begum, and P. Funk, "Automated EEG artifact handling with application in driver monitoring," *IEEE J. Biomed. Health Informat.*, vol. 22, no. 5, pp. 1350–1361, Sep. 2018.

[22] N. Al-Qazzaz, S. H. B. M. Ali, S. Ahmad, M. Islam, and J. Escudero, "Automatic artifact removal in EEG of normal and demented individuals using ICA–WT during working memory tasks," *Sensors*, vol. 17, no. 6, p. 1326, Jun. 2017.

[23] V. Mihajlovic, B. Grundlehner, R. Vullers, and J. Penders, "Wearable, wireless EEG solutions in daily life applications: What are we missing?" *IEEE J. Biomed. Health Informat.*, vol. 19, no. 1, pp. 6–21, Jan. 2015.

[24] P. Sawangjai, S. Hompoonsup, P. Leelaarporn, S. Kongwudhikunakorn, and T. Wilaiprasitporn, "Consumer grade EEG measuring sensors as research tools: A review," *IEEE Sensors J.*, vol. 20, no. 8, pp. 3996–4024, Apr. 2019.

[25] O. R. Pinheiro, L. R. G. Alves, and J. R. D. Souza, "EEG signals classification: Motor imagery for driving an intelligent wheelchair," *IEEE Latin Amer. Trans.*, vol. 16, no. 1, pp. 254–259, Jan. 2018.

[26] F. Faradji, R. K. Ward, and G. E. Birch, "A self-paced two-state mental task-based brain-computer interface with few EEG channels," in *New Frontiers in Brain-Computer Interfaces*. London, U.K.: IntechOpen, 2019.

[27] I. Rejer and P. Górski, "Benefits of ICA in the case of a few channel EEG," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 7434–7437.

[28] I. Rejer and P. Górski, "MAICA: An ICA-based method for source separation in a low-channel EEG recording," *J. Neural Eng.*, vol. 16, no. 5, Oct. 2019, Art. no. 056025.

[29] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.

[30] L. Citi, R. Poli, and C. Cinel, "Documenting, modelling and exploiting P300 amplitude changes due to variable target delays in Donchin's speller," *J. Neural Eng.*, vol. 7, no. 5, Oct. 2010, Art. no. 056006.

[31] A. L. Goldberger *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, Jun. 2000.

[32] P. Aricò, F. Aloise, F. Schettini, S. Salinari, D. Mattia, and F. Cincotti, "Influence of P300 latency jitter on event related potential-based brain–computer interface performance," *J. Neural Eng.*, vol. 11, no. 3, Jun. 2014, Art. no. 035008.

[33] F. Aloise *et al.*, "A covert attention P300-based brain–computer interface: Geospell," *Ergonomics*, vol. 55, no. 5, pp. 538–551, May 2012.

[34] M. Torkamani-Azar, S. D. Kanik, S. Aydin, and M. Cetin, "Prediction of reaction time and vigilance variability from spatio-spectral features of resting-state EEG in a long sustained attention task," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 9, pp. 2550–2558, Sep. 2020.

[35] O. G. Lins, T. W. Picton, P. Berg, and M. Scherg, "Ocular artifacts in EEG and event-related potentials I: Scalp topography," *Brain Topogr.*, vol. 6, no. 1, pp. 51–63, Sep. 1993.

[36] N. K. Al-Qazzaz, S. H. B. M. Ali, S. A. Ahmad, M. S. Islam, and J. Escudero, "Selection of mother wavelet functions for multi-channel EEG signal analysis during a working memory task," *Sensors*, vol. 15, no. 11, pp. 29015–29035, 2015. [Online]. Available: https://www.mdpi.com/1424-8220/15/11/29015

[37] M. I. Al-Kadi, M. B. I. Reaz, and M. A. M. Ali, "Compatibility of mother wavelet functions with the electroencephalographic signal," in *Proc. IEEE-EMBS Conf. Biomed. Eng. Sci.*, Dec. 2012, pp. 113–117.

[38] J. Rafiee, M. A. Rafiee, N. Prause, and M. P. Schoen, "Wavelet basis functions in biomedical signal processing," *Expert Syst. Appl.*, vol. 38, no. 5, pp. 6190–6201, 2011.

[39] M. X. Cohen, "A better way to define and describe Morlet wavelets for time-frequency analysis," *NeuroImage*, vol. 199, pp. 81–86, Oct. 2019.

[40] J. B. Salyers, Y. Dong, and Y. Gai, "Continuous wavelet transform for decoding finger movements from single-channel EEG," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 6, pp. 1588–1597, Jun. 2019.

[41] N. Golyandina, V. Nekrutkin, and A. A. Zhigljavsky, *Analysis of Time Series Structure: SSA and Related Techniques*. Boca Raton, FL, USA: CRC Press, 2001.

[42] B. Hjorth, "EEG analysis based on time domain properties," *Electroencephalogr. Clin. Neurophysiol.*, vol. 29, no. 3, pp. 306–310, 1970.

[43] A. K. Maddirala and R. A. Shaik, "Removal of EOG artifacts from single channel EEG signals using combined singular spectrum analysis and adaptive noise canceler," *IEEE Sensors J.*, vol. 16, no. 23, pp. 8279–8287, 2016.

[44] A. K. Maddirala and K. C. Veluvolu, "SSA with CWT and k-means for eye-blink artifact removal from single-channel EEG signals," *Sensors*, vol. 22, no. 3, p. 931, 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/3/931

[45] C. Brunner, R. Scherer, B. Graimann, G. Supp, and G. Pfurtscheller, "Online control of a brain-computer interface using phase synchronization," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 12, pp. 2501–2506, Dec. 2006.

[46] V. Gonuguntla, Y. Wang, and K. C. Veluvolu, "Event-related functional network identification: Application to EEG classification," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 7, pp. 1284–1294, Oct. 2016.

[47] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 1–27, May 2011, doi: 10.1145/1961189.1961199.

[48] P. Ofner, A. Schwarz, J. Pereira, D. Wyss, R. Wildburger, and G. R. Müller-Putz, "Attempted arm and hand movements can be decoded from low-frequency EEG from persons with spinal cord injury," *Sci. Rep.*, vol. 9, no. 1, pp. 1–15, Dec. 2019.

[49] J. M. Cassidy *et al.*, "Low-frequency oscillations are a biomarker of injury and recovery after stroke," *Stroke*, vol. 51, no. 5, pp. 1442–1450, May 2020.

[50] W. Li, Q.-C. He, X.-M. Fan, and Z.-M. Fei, "Evaluation of driver fatigue on two channels of EEG data," *Neurosci. Lett.*, vol. 506, no. 2, pp. 235–239, Jan. 2012.