



**HAL**  
open science

## Online detection and removal of eye blink artifacts from electroencephalogram

Ashvaany Egambaram, Nasreen Badruddin, Vijanth Asirvadam, Tahamina Begum, Eric Fauvet, Christophe Stolz

► **To cite this version:**

Ashvaany Egambaram, Nasreen Badruddin, Vijanth Asirvadam, Tahamina Begum, Eric Fauvet, et al.. Online detection and removal of eye blink artifacts from electroencephalogram. Biomedical Signal Processing and Control, 2021, 69, pp.102887. 10.1016/j.bspc.2021.102887 . hal-03549148

**HAL Id: hal-03549148**

**<https://hal.science/hal-03549148>**

Submitted on 2 Aug 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

## Online Detection and Removal of Eye Blink Artifacts from Electroencephalogram

Ashvaany Egambaram<sup>a,c,e</sup>, Nasreen Badruddin<sup>a,c</sup>, Vijanth S Asirvadam<sup>b,c</sup>,  
Tahamina Begum<sup>d</sup>, Eric Fauvet<sup>e</sup>, Christophe Stolz<sup>e</sup>

<sup>a</sup>*Institute of Health and Analytics, Universiti Teknologi PETRONAS, Perak, Malaysia*

<sup>b</sup>*Institute of Autonomous Systems, Universiti Teknologi PETRONAS, Perak, Malaysia*

<sup>c</sup>*Department of Electrical and Electronic Engineering, Universiti Teknologi PETRONAS,  
Perak, Malaysia*

<sup>d</sup>*Department of Neuroscience, Universiti Sains Malaysia, Kelantan, Malaysia*

<sup>e</sup>*Laboratoire Electronique, Informatique et Image (Le2i), ERL VIBOT CNRS 6000,  
Universite de Bourgogne, France*

---

### Abstract

The most prominent type of artifact contaminating electroencephalogram (EEG) signals are the eye blink (EB) artifacts, which could potentially lead to misinterpretation of the EEG signal. Online identification and elimination of eye blink artifacts are crucial in applications such as Brain-Computer Interfaces (BCI), neurofeedback, and epilepsy diagnosis. In this paper, algorithms that combine unsupervised eye blink artifact detection (eADA) with modified Empirical Mode Decomposition (FastEMD) and Canonical Correlation Analysis (CCA) are proposed, i.e., FastEMD-CCA<sup>2</sup> and FastCCA, to automatically identify eye blink artifacts and remove them in an online setting. The average accuracy, sensitivity, specificity, and error rate for eye blink artifact removal with FastEMD-CCA<sup>2</sup> is 97.9%, 97.65%, 99.22%, and 2.1%, respectively, validated on a Hitachi dataset with 60 EEG signals, consisting of more than 5600 eye blink artifacts. FastCCA achieved an average of 99.47%, 99.44%, 99.74%, and 0.53% artifact removal accuracy, sensitivity, specificity, and error rate, respectively, validated on the Hitachi dataset too. FastEMD-CCA<sup>2</sup> and FastCCA algorithms are developed

---

*Email addresses:* [ashvaany.egambaram@utp.edu.my](mailto:ashvaany.egambaram@utp.edu.my) (Ashvaany Egambaram),  
[nasreen.b@utp.edu.my](mailto:nasreen.b@utp.edu.my) (Nasreen Badruddin), [vijanth\\_sagayan@utp.edu.my](mailto:vijanth_sagayan@utp.edu.my) (Vijanth S  
Asirvadam), [tahamina@usm.my](mailto:tahamina@usm.my) (Tahamina Begum), [eric.fauvet@u-bourgogne.fr](mailto:eric.fauvet@u-bourgogne.fr) (Eric  
Fauvet), [christophe.stolz@u-bourgogne.fr](mailto:christophe.stolz@u-bourgogne.fr) (Christophe Stolz)

and implemented in the C++ programming language, mainly to investigate the processing speed that these algorithms could achieve in a different medium. Analysis has shown that FastEMD-CCA<sup>2</sup> and FastCCA took about 10.7 and 12.7 milliseconds, respectively, on average to clean a 1-second length of EEG segment. As a result, they're a viable option for applications that require online removal of eye blink objects from EEG signals.

*Keywords:* Canonical Correlation Analysis (CCA), Electroencephalogram (EEG), Eye Blink Artifact, Modified Empirical Mode Decomposition (FastEMD).

---

## 1. Introduction

Electroencephalogram (EEG) signals are used extensively in the medical field for diagnosing epilepsy, sleep disorders, coma, encephalopathy, brain injury, and brain death [1–8]. EEG signals are also utilized for research purposes, for example, in neuroscience, cognitive science, cognitive psychology, neurolinguistics, and psychophysiology [9–11]. Although EEG signals play a vital role in many research and medical fields, artifacts often contaminate them. The most prominent type of artifacts contaminating the EEG signals is the eyeblink artifacts. Therefore, eyeblink artifact detection and removal is an essential preprocessing step for any EEG-based application.

Eye blink artifacts contaminate all the EEG channels or electrodes but are not apparent on channels located further from the eyes. They appear most prominently on the frontal electrodes, Fp1 and Fp2, as these electrodes are very close to the position of the eyes. Various eyeblink artifact removal algorithms are developed to date. They include algorithms of which are capable of manual, semi-automated, fully-automated artifact elimination, either from single-channel EEG or multichannel EEG. One of the most widely used techniques is the manual eye blink artifact rejection method, which involves manual inspection of the EEG signals to discard the eye blink artifact regions. This method may result in substantial loss of data as EEG segments that are be-

ing removed could contain useful neurological information. Alternatively, many semi-automatic and fully-automatic eyeblink artifact removal algorithms were developed to replace manual eyeblink artifact rejection method, which were discussed and reviewed in [12–17]. In addition to the existing automatic artifact removal algorithms, techniques addressing and implementing the online removal of eyeblink artifacts are essential [14].

In the brain-computer interface (BCI) devices, eye blink artifacts are the most significant and apparent type of artifact [18, 19]. If not correctly identified and removed, eye blink artifacts may be accidentally used as a source that could mislead classification and the controllability of the BCI device. This scenario could lead to a drop in the BCI system’s performance during real-time applications. For this reason, BCIs related to severe motor disabilities are in demand for accurate eye blink artifact recognition and elimination during online operations, which will make the BCI device more robust [20]. Similarly, in neurofeedback, the frequency of eye blink artifacts often overlaps with those used for neurofeedback training. This could potentially manipulate the feedback signal, which may strongly influence and invalidate the learning outcome [21]. Another study by Sherlin et al. [22] stated that neurofeedback training could misinterpret eye blink artifacts, causing incorrect learning process. Therefore, the authors have strongly suggested that real-time eye blink artifact detection and removal is required to avoid ”artifact-driven” feedback.

In real-time epilepsy monitoring and seizure detection units [23–25], eyeblink artifacts often lead to misinterpretation of artifactual segments as an epileptiform activity. Epileptiform refer to waves and spikes that may be associated with epilepsy. Misinterpreted eye blink artifacts as an epileptiform activity may result in incorrect medication and treatment [26]. Acar et al. have stated that artifacts originating from eye blinks and eye movements often undermine efforts to localize epileptic foci from the EEG signal of an epilepsy patient [27]. In research related to Alzheimer’s disease, EEG recording during the eyes-open state is not preferred mainly due to contamination from eye blink artifacts, and artifact avoidance is impractical. Whenever EEG signals are recorded during

the eyes-open state, excessive eye blink artifacts are present; thus, artifact-free segments have to be patched together to obtain sufficient duration of EEG signal for analysis. However, patching EEG segments together produces discontinuous neural information, which may introduce incorrect interpretation on the analysis of Alzheimer's EEG signal [28]. In research related to cognitive development, researchers prefer EEG portions associated with the cognitive process of interest to be free of eye blink artifacts so that the data analysis is meaningful [29]. Although automatic artifact removal algorithms are available, it is clear that studies addressing and implementing online removal of artifacts are essential. This is particularly relevant in BCI research for proper BCI output device controllability and in neurofeedback in generating distinct feedback signals. It is therefore mandatory to remove eye blink artifacts from EEG signals in an online manner for correct EEG signal interpretation in any EEG-based application.

As discussed, EEG based applications such as BCI, neurofeedback, and epilepsy monitoring systems relies on the availability of instant instructions from the brain and real-time execution of the devices. These systems must be workable while the users perform their daily tasks or during real-time monitoring, which means they should operate in an online manner. Therefore, it is mandatory for any algorithm that deals with BCI, neurofeedback, and epilepsy monitoring to perform EEG acquisition, artifact correction, feature extraction, and classification online. Since these applications demand online signal processing, artifact removal methods and algorithms should also process data online. This paper focuses on online eyeblink artifact elimination from EEG signals as useful EEG instructions can be fed to BCI or neurofeedback applications only if artifacts contaminating the EEG signals are correctly identified and removed. The online artifact removal algorithm should have a low computational cost so as not to introduce an undesirable time delay in an online implementation. It should also be able to perform window or block processing on the EEG signals to reduce computational complexity. This will automatically improve the processing speed of the online algorithm, thus not introducing an unacceptable time delay to the entire application. The computational environment in

which the eyeblink artifact removal algorithm is implemented is crucial in online applications. MATLAB is currently the most popular tool in research settings, but MATLAB alone may not be a viable medium for online implementation.

In this paper, the proposed algorithms, FastEMD-CCA<sup>2</sup> and FastCCA are implemented and evaluated in an inexpensive computing environment, C++ programming language, to investigate the feasibility of the approaches in accommodating online applications. Implementation of the algorithms in C++ is mainly to investigate the computation time or the processing speed of the proposed algorithms, FastEMD-CCA<sup>2</sup> and FastCCA, could achieve, in line to support an online application. The developed algorithms are compared with one of the state-of-the-art methods in MATLAB, i.e., FORCe, due to its effectiveness in removing eyeblink artifacts.

## 2. Materials and Methods

### 2.1. EEG Recording and Analysis

The EEG dataset used for evaluation in this paper were recorded at Hitachi's Hatayoma facility in Japan. Following receipt of written informed consent, EEG signals from volunteers were collected in accordance with the regulations of the internal review board of Hitachi, Ltd.'s Central Research Laboratory. 20131021-0138 is the approval number. These EEG signals were explicitly obtained for the purpose of conducting a mental stress analysis. Since eye blink artifacts contaminate all recorded signals, the dataset is appropriate for this research work. Fourteen free electrodes were used to record these EEG signals, placed on the scalp following the 10-20 system. The EEG signals were obtained from 10 participants, each of whom had six recordings, resulting in 60 EEG signals. The participants are between the ages of 30 and 55. All the recorded signals are of varying lengths and were captured at a sampling rate of 256 Hz. During EEG recording, the EOG electrodes that capture eye blink events are not used to record the EOG signals for convenience purposes. Validation, whether the eye blink artifacts are removed, turns out to be difficult due to the unavailability of

eye blink ground truths. Thus, an experts advice, Neuroscientist Dr. Tahamina Begum, is required to substantiate if the approaches can effectively remove eye blink artifacts. The total number of eye blink artifacts found in this dataset is more than 5600, identified through manual inspection. The number of eye blink artifacts contaminating each of the EEG signal varies, ranging from 20 to 172 occurrences.

## *2.2. Existing Online Artifact Removal Techniques*

Some of the most related works which may be viable for online applications are discussed in the following subsections.

### *2.2.1. Wavelet Neural Network*

Using a combination of Wavelet Transform (WT) and Artificial Neural Network (ANN), Nguyen et al. [30] have published their work on ocular artifact elimination, naming their technique as Wavelet Neural Network (WNN). Initially, the neural network is trained to classify artifacts using a separate artifact/EOG recording and non-artifacts with simulated EEG signals. Once the network is trained, contaminated EEG signals are subjected to WT to obtain wavelet coefficients, which are then passed to the ANN classifier for artifact identification and correction. Corrected wavelet coefficients are then reconstructed to get a clean version of the EEG signal. The authors have mentioned that this algorithm is computationally efficient; therefore, it may be a reliable solution for real-time artifact removal. Though it is computationally efficient, this algorithm requires an additional artifact/EOG recording to train the ANN classifier, which may add up some time delay for its implementation in real-time. Moreover, the algorithm is only capable of removing artifacts from a single-channel EEG signal, which is not practical for a real-world application.

### *2.2.2. FORCe*

The Fully Online and Automated Artifact Removal for Brain-Computer Interfacing (FORCe) is a software plugin GUI developed by Daly et al. [31]. This plugin works based on the combination of WT, ICA and thresholding. It is

designed to perform in an automated online environment and to remove several types of artifacts, such as the eyeblink artifacts, cardiac artifacts and muscle artifacts. First, WT is added to each channel of an EEG signal for a 1-second epoch. The approximation coefficients from WT are then subjected to ICA to produce a collection of independent components or ICs. The artifactual ICs are then categorised using a set of threshold parameters, with ICs exceeding these thresholds being labelled as eye blink and cardiac artifacts, and thus excluded. To get a set of cleaned approximation coefficients, the inverse of ICA is performed on the remaining non-artifactual ICs. Then, to suppress/remove muscle artifacts, soft thresholding is applied to the ICA approximation coefficients and the detail coefficients acquired through WT. Finally, the algorithm produces EEG epochs that are free of artifacts.

While the algorithm is able to remove eyeblink, cardiac and muscle artifacts online, the algorithm relies completely on manually pre-determined thresholds to classify if an IC is artifactual. Furthermore, the selection of thresholds was based on the analysis performed only on the EEG signals of two participants. Since artifact patterns or characteristics may vary for every individual, manually adjusted and pre-determined thresholds based on signals of two participants may not be suitable to detect and remove artifactual ICs of a wider range of EEG datasets. The authors have also stated that the running time of the algorithm would linearly increase with an increasing number of channels. So, this would add up some time delay to its implementation in online operations, especially in applications requiring additional number of channels, for example, in seizure detection units to localize epileptic foci.

### *2.2.3. Real-time Source-mapping Toolbox*

Tonachini et al. [32] recently published an online automated artefact rejection (REST) toolbox utilising artefact subspace reconstruction (ASR), PCA, online recursive ICA (ORICA), and an IC classifier. ASR is an automated, variance-based algorithm, that learns the statistical properties of an artifact-free EEG segment. Once the learning is complete, PCA is applied to transform



contaminated EEG segments into PCs, which are then compared with the learnt data. PCs that are exceeding initially calibrated/learnt data are removed, where transient and large-amplitude artifacts get removed in this stage. Remaining PCs are re-projected back to acquire a partially cleaned EEG segment. Next, ORICA is performed on partially cleaned EEG segments from PCA, which produces a set of ICs. These ICs are categorized into eye movement ICs and non-eye movement ICs using an altered version of the EyeCatch classifier. The classification is done by getting the correlation value of the ICs with the IC scalp maps contained in the library of EyeCatch. ICs that exceed a fixed correlation value is removed and a clean EEG segment is reconstructed.

Despite the fact that the authors claim that this algorithm works in an online environment to remove eye blink, cardiac, and muscle artifacts, they also note that ASR had negligible effect on eye blink artifact removal. On the other hand, ORICA took 26 seconds to converge well on a blink-related IC to be removed, which is too long for an online algorithm. Additionally, the authors have pointed out that the altered version of EyeCatch classifier has introduced some instability to the correlation values used in classifying the artifactual ICs. Hence, an online implementation is certainly intolerable with the significant amount of time consumed by ORICA to identify an eyeblink artifact related IC, and the instability introduced by the EyeCatch classifier. Thus, it is concluded that this algorithm may not be suitable to eliminate eyeblink artifacts in an online manner, although it can effectively remove cardiac and muscle artifacts.

### *2.3. Review of Proposed Methods*

#### *2.3.1. Proposed Eyeblink Artifact Detection Algorithm (eADA)*

A novel algorithm, (eADA), to automatically identify eyeblink artifacts with adaptable and varying threshold values, without any supervision on the EEG signal is proposed. The detailed description of how eADA works can be found in [33]. The idea behind designing an eyeblink artifact detection algorithm is to assist the subsequent artifact removal algorithm. The eyeblink artifacts contaminate the EEG signal at random points of the EEG signal, which tanta-

mount to a very short period in time compared to the entire length of the EEG signal. The artifact correction algorithm does not have to work on long EEG segments if accurate locations of the eyeblink artifacts are identified in advance. Consequently, distortion to the artifact-free segments of the EEG signal can be avoided.

### 2.3.2. Implementation of CCA in Artifact Removal

Canonical Correlation Analysis (CCA) is a technique that is based on the blind source separation (BSS) concept. BSS distinguishes a collection of source signals from a set of mixed signals without any prior knowledge of the source signals or the weighted mixing elements, as the name implies. The linear relationship between two multidimensional variables is measured. Conceptually, BSS assumes the contaminated or the set of mixed signals as the combination of clean EEG sources and artifacts blended together with a mixing formula. Thus, BSS attempts to isolate EEG sources and artifacts apart with an unmixing formula, which can be useful for artifact elimination from EEG signals.

The observed EEG signal,  $\mathbf{x}(\mathbf{t})$  is the first multidimensional variable, while the second multidimensional variable is obtained by taking a temporally delayed component of the observed EEG signal,  $\mathbf{y}(\mathbf{t}) = \mathbf{x}(\mathbf{t} - \mathbf{1})$ . As BSS implies, the observed EEG signal  $\mathbf{x}(\mathbf{t})$  is a combination of sources  $\mathbf{S}(\mathbf{t})$ , where  $\mathbf{S}(\mathbf{t})$  consists of EEG sources and artifactual components, mixed through a weighted mixing matrix,  $\mathbf{W}$ :

$$\begin{aligned}\mathbf{x}(\mathbf{t}) &= \mathbf{W}_x \mathbf{S}(\mathbf{t}) \\ \mathbf{y}(\mathbf{t}) &= \mathbf{W}_y \mathbf{S}(\mathbf{t}).\end{aligned}\tag{1}$$

The source signals can be obtained by projecting the weighted de-mixing matrix onto the observed EEG signals as in Eq.(2):

$$\begin{aligned}\mathbf{A} &= \mathbf{W}_x^{-1} \\ \mathbf{B} &= \mathbf{W}_y^{-1} \\ \mathbf{S}_x(\mathbf{t}) &= \mathbf{x} \mathbf{A}(\mathbf{t}) \\ \mathbf{S}_y(\mathbf{t}) &= \mathbf{y} \mathbf{B}(\mathbf{t}).\end{aligned}\tag{2}$$

Hence, sources of the two multidimensional variables,  $\mathbf{S}_x(\mathbf{t})$  and  $\mathbf{S}_y(\mathbf{t})$  can be estimated if the weighted de-mixing matrices,  $\mathbf{A}$  and  $\mathbf{B}$  are known.

In CCA, the sources are named as canonical variates or canonical components,  $\mathbf{U}$  and  $\mathbf{V}$ . Canonical variates for a multi-channel EEG signal can be represented through a linear combination between the de-mixing matrices with mean removed EEG variables,  $\hat{\mathbf{x}}$  and  $\hat{\mathbf{y}}$ , with  $n$  number of EEG sample points in one channel, and  $p$  are the number of channels of  $\mathbf{x}(\mathbf{t})$  and  $\mathbf{y}(\mathbf{t})$ :

$$\begin{aligned} U_1 &= \hat{x}_1 A_{11} + \hat{x}_2 A_{12} + \dots + \hat{x}_p A_{1p} \\ V_1 &= \hat{y}_1 B_{11} + \hat{y}_2 B_{12} + \dots + \hat{y}_p B_{1p} \end{aligned} \quad (3)$$

Eq. (3) can be generalized as below, similar to Eq. (2):

$$\begin{aligned} \mathbf{U} &= \hat{\mathbf{x}}\mathbf{A} \\ \mathbf{V} &= \hat{\mathbf{y}}\mathbf{B}. \end{aligned} \quad (4)$$

The purpose of CCA is to find the de-mixing matrices  $\mathbf{A}$  and  $\mathbf{B}$  such that the correlation between canonical variates is maximized. The correlation between the canonical variates,  $\mathbf{U}$  and  $\mathbf{V}$  is called the canonical correlation,  $\rho$ . The  $\rho$  between  $\mathbf{U}$  and  $\mathbf{V}$  should be maximized, or as large as possible:

$$\begin{aligned} \rho_i &= \text{corr}(U_i, V_i) \\ \rho_i &= \frac{U_i^T V_i}{\sqrt{U_i^T U_i} \sqrt{V_i^T V_i}} \end{aligned} \quad (5)$$

where  $\rho_i$  is the  $i$ -th canonical correlation,  $U_i$  and  $V_i$  are the  $i$ -th canonical variates. From Eq. (5), canonical variate pairs are derived, where  $(U_1, V_1)$  is the first canonical variate pair; similarly,  $(U_2, V_2)$  is the second canonical variate pair, so on and so forth. The de-mixing matrices  $A_1 = [A_{11}, A_{12}, \dots, A_{1p}]$  and  $B_1 = [B_{11}, B_{12}, \dots, B_{1p}]$  are estimated to maximise the coefficient of canonical correlation between the first pair of canonical variates  $U_1$  and  $V_1$ . The canonical correlation of the second pair is computed in the same way. The condition that the second pair of canonical variates are uncorrelated or orthogonal to the first pair and other pairs of canonical variates in the subspace.

As discussed earlier, some rows of the canonical variates obtained through CCA represent the clean EEG sources and one canonical variates row represents the artifact. So, by forcing the artifactual canonical variates to zero and projecting back the artifact-free canonical components, a clean EEG section free of artifacts can be reconstructed. The implementation of eyeblink artifact removal from EEG using CCA is summarized below.

1. For elaboration purpose, a data set,  $\mathbf{s}(\mathbf{t})$  with a length of 5 seconds is used.
2. Synthetic eyeblink artifacts are added to the EEG dataset to produce a contaminated EEG signal. The synthetic eyeblink artifacts,  $eb(t)$  can be simulated through exponential functions with different amplitudes, as in Eq. (6).

$$eb(t) = 40e^{-(10t-10)^2} + 40e^{-(10t-30)^2} + 32e^{-(10t-45)^2} + 28e^{-(10t-70)^2} \quad (6)$$

Contaminated EEG signal can be obtained by adding the EEG dataset  $s(t)$  with synthetic eyeblink artifacts, as in Eq. (7).

$$\mathbf{x}(\mathbf{t}) = \mathbf{e}\mathbf{b}(\mathbf{t}) + \mathbf{s}(\mathbf{t}) \quad (7)$$

3. The second multidimensional data set,  $\mathbf{y}(\mathbf{t})$  is taken, with  $\mathbf{y}(\mathbf{t}) = \mathbf{x}(\mathbf{t}-1)$ .
4. Mean removed signals,  $\hat{\mathbf{x}}(t)$  in Fig. 1 and  $\hat{\mathbf{y}}(t)$  in Fig. 2 are obtained by removing the respective mean from  $\mathbf{x}(\mathbf{t})$  and  $\mathbf{y}(\mathbf{t})$ .
5. Next, the weighted de-mixing matrices  $\mathbf{A}$  and  $\mathbf{B}$  are estimated.
6. Canonical variates  $\mathbf{U}$  are computed by projecting the estimated de-mixing weight matrix  $\mathbf{A}$  onto the mean removed signal  $\hat{\mathbf{x}}$  as in Eq. (4). The resulting canonical variate  $\mathbf{U}$ , is shown in Fig. 3.
7. The eyeblink artifact components are well distinguished from the neural components as they behave as the least cross-correlated components among the canonical variate vectors. This row of canonical components

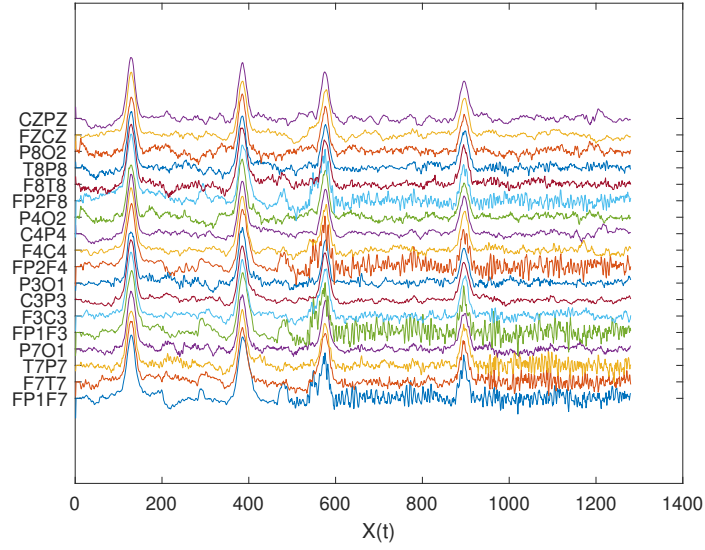


Fig. 1: First multidimensional data set,  $\hat{x}(t)$

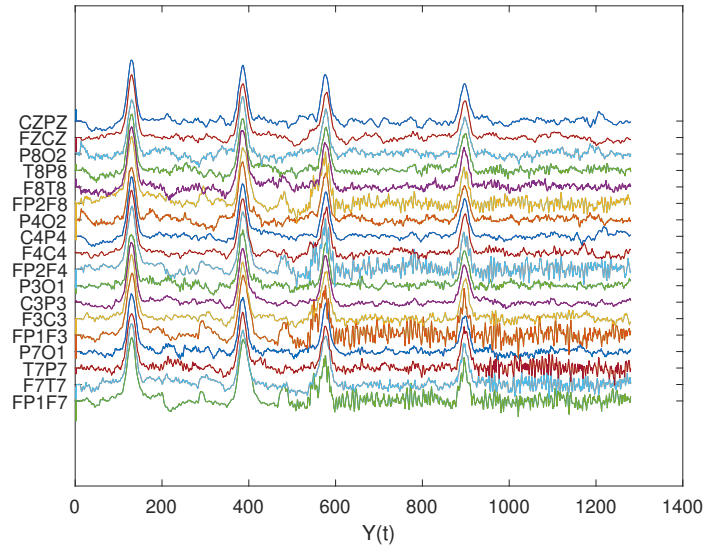


Fig. 2: Second multidimensional data set,  $\hat{y}(t)$

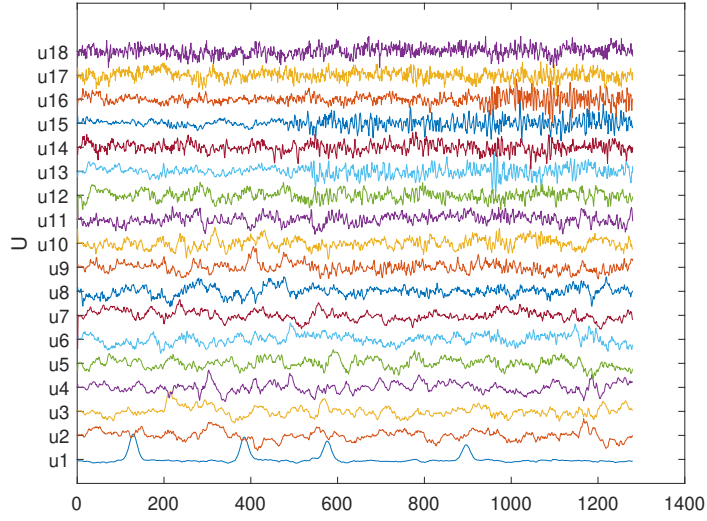


Fig. 3: Canonical components of  $\hat{\mathbf{x}}$

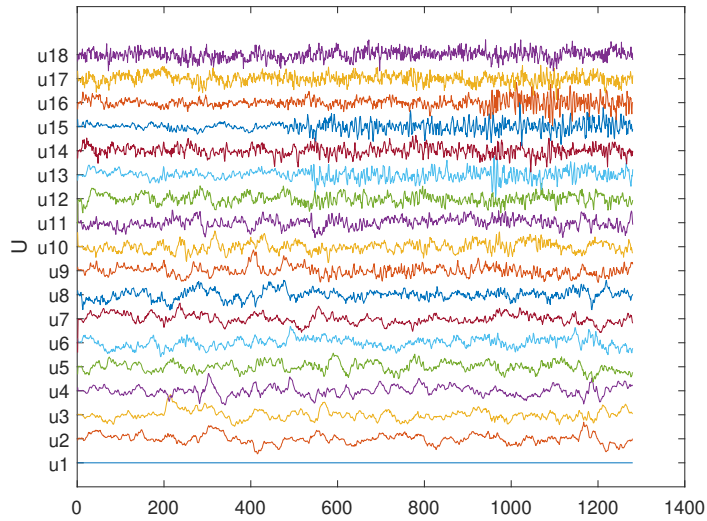


Fig. 4: Non-artifactual Canonical components

is excluded out by forcing it to become zero in order for it to behave non-artifactual,  $\mathbf{U}_{\text{clean}}$  as shown in Fig. 4.

Finally, a clean EEG signal that is free from artifacts can be reconstructed as shown in Fig. 5, by taking the inverse of the de-mixing matrix,  $\mathbf{A}^{-1}$ , into the non-artifactual canonical components,  $\mathbf{U}_{\text{clean}}$ :

$$\mathbf{X}_{\text{clean}} = \mathbf{A}^{-1}\mathbf{U}_{\text{clean}} \quad (8)$$

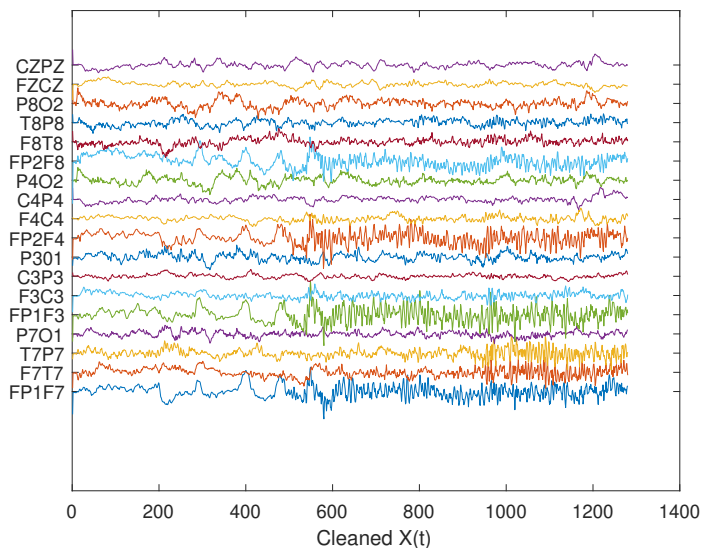


Fig. 5: Artifact-free EEG dataset

### 2.3.3. Modified Empirical Mode Decomposition

Empirical Mode Decomposition (EMD) decomposes a raw signal into several oscillating trends in a recursive manner through the interpolation and subtraction process, producing high and low oscillating decomposed trends. These decomposed trends are called Intrinsic Mode Functions (IMFs). In comparison to its predecessor, each IMF derived from the original signal has a lower oscillating trend. The original signal can be reconstructed by adding all IMFs and

the residual signal obtained from the decomposition. A signal can be broken down into several IMFs using traditional EMD until the final residue becomes a monotonic trend.

$$X(t) = \sum_{i=1}^{n-1} x_i(t) + R_n(t) \quad (9)$$

, where  $X(t)$  is the raw EEG signal,  $x_i(t)$  are  $i$  number of IMFs, and  $R_n(t)$  is the residual component, which is monotonous. The higher oscillations in the raw EEG signal belong to EEG components and will be isolated out in the first few IMFs; hence the sum of remaining IMFs would, by default, produce an eye blink artifact trend. Hence, partially reconstructing the higher oscillating trends, which are lower in amplitude, would yield the EEG trend. Alternatively, low oscillating trends with high amplitudes are summed together to attain the eye blink artifact trend.

Several modifications to the classical EMD algorithm are proposed, as described in [34] to resolve the algorithm's processing time inefficiency. The FastEMD algorithm is designed to decompose the raw EEG signal to up to 5 IMFs only, sufficient to segregate out the clean EEG signal and the eye blink artifact template. Consequently, this reduces the computation time, and the algorithm does not have to repeat itself to extract too many IMFs until a monotonic residue is acquired. However, the IMF selection is predetermined during reconstruction according to an assumption made. The first two IMFs correspond to EEG, and the remaining IMFs correspond to the eye blink artifact. Thus, adding up the first two IMFs gives an EEG trend; similarly, adding up the remaining IMFs would produce an eye blink artifact template. In this paper, the IMF selection process is automated with CCA's help to prevent any errors attributed to the assumption made. The automatic selection or classification of IMFs in FastEMD is required to categorize whether an IMF belongs to EEG or the eye blink artifact, subsequently extracting the eye blink artifact template. This can be accomplished by subjecting the row vectors of IMFs to CCA. Fig. 6 shows the canonical variates obtained by applying CCA on the IMFs.

The most pertinent artifactual canonical variate row is extracted out as the



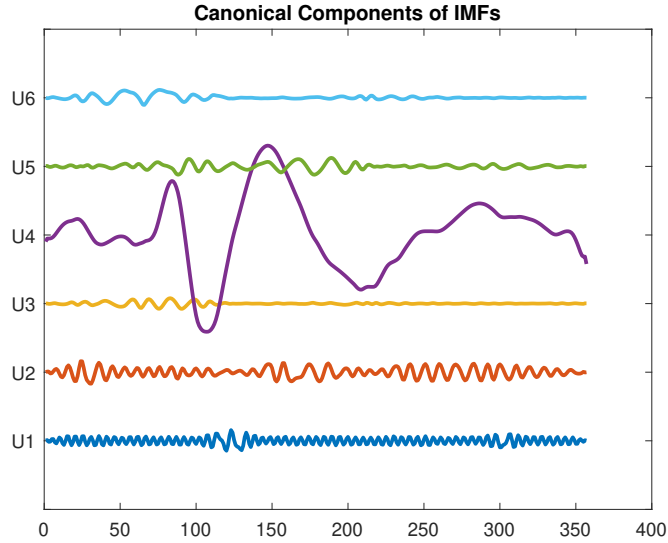


Fig. 6: Canonical Variates of the IMFs

eyeblink artifact template. The remaining canonical variates are non-artifactual sources, so they are used to reconstruct the clean EEG trend. The eyeblink artifact template and the clean EEG signal reconstructed using IMF’s canonical components are shown in Fig. 7.

#### 2.4. Proposed Eye blink Artifact Removal Algorithms

##### 2.4.1. FastEMD-CCA<sup>2</sup>

The combination of eye blink Artifact Detection Algorithm (eADA), modified Empirical Mode Decomposition (FastEMD) and Canonical Correlation Analysis (CCA) is proposed, producing FastEMD-CCA<sup>2</sup>, which can be used to remove eye blink artifacts in online applications. In FastEMD-CCA<sup>2</sup>, several eye blink artifact regions are searched, identified and saved using eADA until two eye blink artifact regions exhibit a correlation coefficient of more than 0.9. A high correlation, i.e., 0.9, between the eye blink artefact regions is thought to indicate repetitiveness or similarity in an individual’s blinking pattern, which can be assumed to be a general eye blinking pattern for that particular EEG

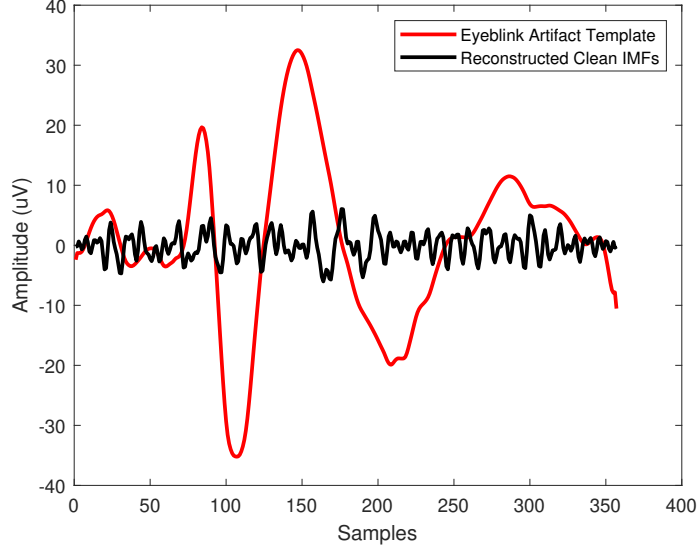


Fig. 7: Extracted Eyeblink Artifact Template and Reconstructed EEG Signal

signal. FastEMD is applied only on the latter eye blink artifact region to extract an eye blink artifact template, thus keeping the number of FastEMD applications as low as possible. This approach differs from what is used in traditional artifact removal techniques using EMD, in which traditional EMD is used to remove the artifacts whenever an artifact event is detected. In FastEMD-CCA<sup>2</sup>, the IMFs obtained through FastEMD are subjected to CCA as illustrated in subsection 2.3.3 to extract out a general eye blink artifact template. Remaining eye blink artifacts that contaminate an EEG signal are removed with the help of the extracted eyeblink artifact template, assuming every other eye blink artifacts within a subject exhibit a consistent pattern with the template. A sliding window with the length of the extracted eyeblink artifact template is moved along the EEG signal and each EEG window is cross-correlated with the general eyeblink artifact template extracted:

$$\text{Cross-correlation} = \frac{C_{X(t), X_{EB}(t)}}{\sigma_{X(t)} * \sigma_{X_{EB}(t)}} \quad (10)$$

where  $C_{X(t), X_{\text{EB}}(t)}$  is the covariance between the contaminated EEG signal,  $X(t)$ , and the eyeblink artifact template,  $X_{\text{EB}}(t)$ , while  $\sigma_{X(t)}$  and  $\sigma_{X_{\text{EB}}(t)}$  are the standard deviations of these signals.

Observations and validation on the Hitachi EEG signals revealed that the correlation between the eye blink artifact template and EEG windows contaminated with eye blink artifacts often lies in the range of 0.4 to 0.6, so a correlation value of more than 0.4 is used to indicate the presence of an eye blink artifact. Hence, EEG windows that exhibit a similarity score of more than 0.4 with the eye blink artifact template are confirmed as eye blink artifacts, thus subjected to CCA for artifact removal. Removal of the eye blink artifacts from the specified multichannel EEG window relies on eliminating the artifactual canonical components of CCA. CCA first estimates the canonical components that maximize temporal correlation within the given window. Then, among the canonical variate vectors, the most relevant artifactual canonical components ( $U_i, V_i$ ), typically the least cross-correlated canonical component, are forced to become zero in order for it to behave non-artifactual. The artifact-free canonical components are termed as  $\mathbf{U}_{\text{clean}}$ . Finally, a clean EEG segment is reconstructed by projecting the inverse of the de-mixing matrix,  $\mathbf{A}^{-1}$  into the non-artifactual source,  $\mathbf{U}_{\text{clean}}$  as explained in Eq. (8).

In this proposed technique, the extracted general eyeblink artifact template is utilized further by CCA for eyeblink artifact elimination. So CCA is used, first in selecting the eyeblink artifact related IMFs after FastEMD application, and second in eliminating the remaining eyeblink artifacts that are present in an EEG signal. Thus the entire algorithm is named FastEMD-CCA<sup>2</sup> as CCA is used twice. The flowchart of the proposed technique, FastEMD-CCA<sup>2</sup>, is shown in Fig. 8.

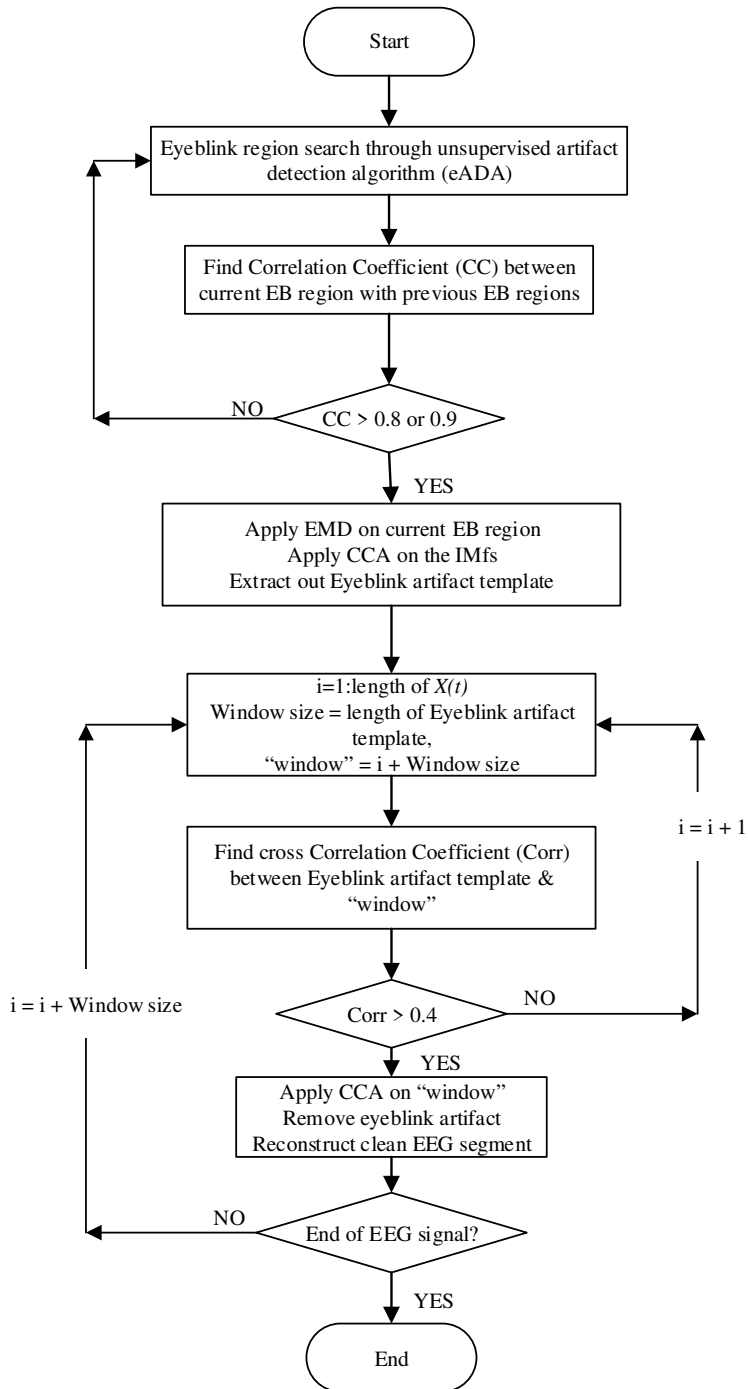


Fig. 8: Flowchart of the Proposed Technique FastEMD-CCA<sup>2</sup>

#### 2.4.2. *FastCCA*

In FastEMD-CCA<sup>2</sup>, the elimination of eye blink artifacts from the entire multichannel EEG signal is conducted through cross-correlation between EEG segments and the eye blink artifact template extracted out through FastEMD. EEG segments that are highly correlated with the eye blink artifact template are subjected to CCA for artifact removal. Since an automatic eADA is already developed and it could accurately identify the eye blink artifact locations, the proposed algorithm’s performance without dependency on the eye blink artifact template to determine the eye blink artifact locations is investigated. Therefore, another algorithm is proposed by combining eADA and CCA to develop the FastCCA algorithm, which is purely a correlation-based approach. In FastCCA, the unsupervised eye blink artifact detection algorithm, eADA is performed in windows of about 1.95s on Fp1 and Fp2 EEG channels. Once an eye blink artifact region is found on the Fp1 channel, this region’s multichannel EEG signal is subjected to CCA for artifact elimination. Then eADA is executed again to search for the next eye blink artifact region on the Fp1 channel. Multichannel artifact elimination is performed on the newly found eye blink artifact region via CCA. So, eADA and CCA are used repeatedly until all eye blink artifacts contaminating the EEG signal are identified and removed. The flowchart of the proposed FastCCA algorithm is shown in Fig. 9. In this approach, CCA is directly applied to EEG segments identified with eye blink artifacts through eADA. So this algorithm bypasses the requirement to have an eye blink artifact template extracted via FastEMD and CCA, which could reduce the computation time as well. This allows an adaptive detection and removal of eye blink artifacts for every event of blink, without the need to have a general template for artifact identification.

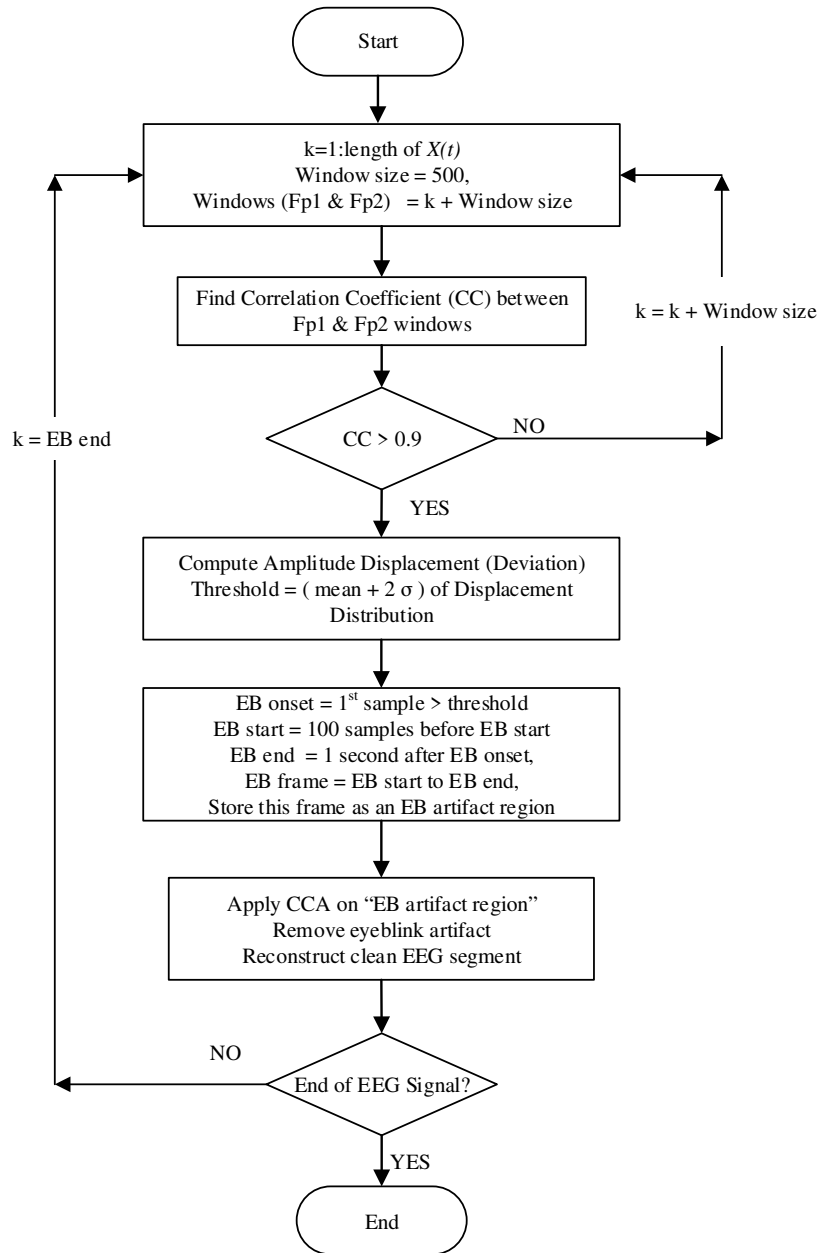


Fig. 9: Flowchart of the Proposed FastCCA Algorithm

The simplified workflow of both algorithms are shown in Fig. 10.

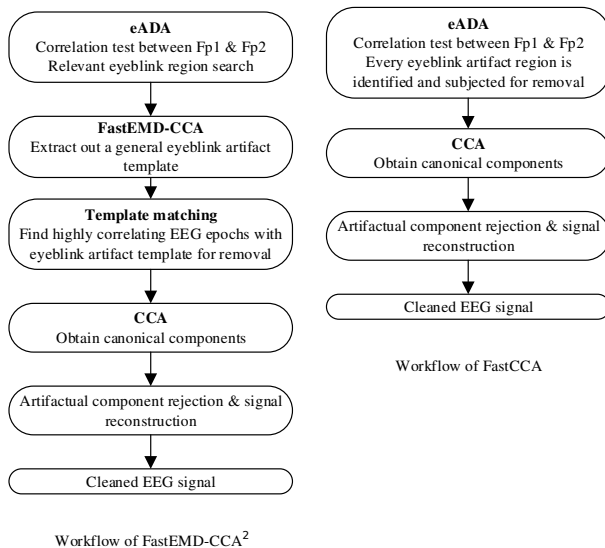


Fig. 10: Workflow of FastEMD-CCA<sup>2</sup> and FastCCA

### 2.5. Implementation Strategies for Online Artifact Removal

In order to achieve an online implementation of the proposed approaches in removing eye blink artifacts from EEG signals, two implementation procedures are proposed. First, the proposed algorithms are executed and processing is performed in small EEG windows, rather than applying the proposed algorithms to the entire EEG signal. Secondly, the proposed algorithms are implemented and evaluated with a compiled language, the C++ language on an Ubuntu Linux 14.04 (64-bit OS, 4GB RAM).

FastEMD-CCA<sup>2</sup> and FastCCA are initially implemented in MATLAB, then followed by C++. In MATLAB, the entire multichannel EEG signal is imported into MATLAB's workspace for processing. The algorithms are executed, and artifact correction is performed on overlapping windows, with each window being about 1.95s in length. In the C++ language, the algorithms are executed in Ubuntu to process the EEG signals in a simulated online setting. The proposed algorithms are designed to fill a buffer by streaming EEG recording on a sample-by-sample basis. After buffering a 1.95s EEG epoch, the samples in the buffer are subjected to unsupervised artifact detection and elimination.

## 2.6. Performance Evaluation

The competency of an online eyeblink artifact removal algorithm relies on the processing speed of the algorithm, where the time taken for artifact elimination from EEG signals should be acceptable for an online application. The time taken for a system to be considered as real-time is still debatable, but estimated to be between 6 and 20 milliseconds [35]. As stated earlier, the reason for the proposed algorithms implemented in C++ is to investigate the processing speed the algorithms can achieve. So, other performance measures apart from the computation time should remain the same as they are the same algorithms, regardless of the implementation platform. This is feasible only if the algorithms are correctly implemented in both MATLAB and C++. So, visual inspection and quantitative analysis in the time domain are performed on the EEG signals to verify the algorithms' effectiveness in both MATLAB and C++. On the whole, the algorithms should achieve instantaneous artifact correction without loss of neural information, to be useful in online applications. The proposed algorithms, FastEMD-CCA<sup>2</sup> and FastCCA, are then compared with one of the state-of-the-art algorithm, FORCE, as it was proven to remove eye blink artifacts from multi-channel EEG signals. The FORCE algorithm is evaluated only in MATLAB and not in C++ as it was originally developed and evaluated in MATLAB by [31]. The other two algorithms, i.e., WNN and REST, are not used for comparison because WNN could only remove artifacts from single-channel EEG signal, while REST had a negligible effect on eyeblink artifact removal.

### 2.6.1. Offline Evaluation through Visual Inspection

The proposed algorithms are evaluated in MATLAB and C++ using offline analysis on the online processed EEG signals after artifact correction. The analysis is conducted on 60 EEG signals, described in section 2.1. Since the ground truths are not available for real EEG signals, the effectiveness of the proposed algorithms in recognising and eliminating eyeblink artefacts while maintaining artifact-free EEG segments is checked using manual visual inspection (MVI)



with the assistance of an expert, Neuroscientist Dr. Tahamina Begum. The evaluation parameters are derived from various binary prediction measures [36, 37], evaluating the algorithms' accuracy, sensitivity, specificity, and error rate..

### 2.6.2. Offline Evaluation through Quantitative Evaluation

Supplementary to eyeblink artifact detection and rejection, the performance of the proposed algorithms in retaining the neural information is quantitatively assessed. This is accomplished by selecting artifact-free EEG segments at random from the 60 EEG recordings mentioned above. Each EEG signal is evaluated for one artifact-free EEG segment chosen randomly. Ideally, it is expected that these segments remain undistorted, even after the artifacts have been removed.

In the time domain, the correlation coefficient (CC), root mean square error (RMSE) and similarity index ( $\eta_{dB}$ ) [38] are used to measure how well the proposed algorithms have preserved the artifact-free EEG segments after artifact correction. CC in Eq. (11) measures the similarity between the original artifact-free EEG segment,  $X_{in}(t)$  with its corresponding reconstructed EEG segment,  $X_{out}(t)$  after artifact removal.

$$CC = \frac{C_{X_{in}(t), X_{out}(t)}}{\sigma_{X_{in}(t)} * \sigma_{X_{out}(t)}} \quad (11)$$

RMSE measures the reconstruction error between original artifact-free EEG segment and the reconstructed EEG segment. In Eq. (12),  $n$  is the number of sample points in the EEG segment:

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (X_{in}(t) - X_{out}(t))^2}{n}} \quad (12)$$

Similarity index ( $\eta_{dB}$ ) in Eq. (13), of the artifact-free EEG segment is computed to quantify the degree of neural information preservation.

$$\eta_{dB} = 10 \log_{10} \left[ \frac{\sum_{t=1}^n (1 - X_{in}(t) - X_{out}(t))}{n} \right] \quad (13)$$

### 2.6.3. Online Evaluation

The online performance of the proposed algorithms is evaluated in terms of the computation time taken by the algorithms in removing artifacts from EEG

signals. The processing speed taken by both algorithms in two different platforms, MATLAB R2018b in Windows 7 Professional (64 bit OS, 4GB RAM) and C++ language on an Ubuntu Linux 14.04 (64-bit OS, 4GB RAM), are recorded for evaluation purposes. The computation time is interpreted as the computing efficiency of the algorithms in cleaning the eyeblink artifacts online. Hence, it is used to evaluate the feasibility of the algorithms for online processing, whether they can achieve instantaneous artifact removal. Whichever algorithm that achieves a shorter computation time denotes better computing efficiency, making it a more suitable candidate for online eyeblink artifact removal applications.

### 3. Results and Discussions

#### 3.1. Offline Analysis Results through Visual Inspection

The EEG signals are subjected to online eyeblink artifact elimination via FORCe, FastEMD-CCA<sup>2</sup> and FastCCA in the MATLAB computing environment and C++ programming language. Visual comparison of Fp1 channel of EEG 3, before and after artifact correction using FORCe in MATLAB, FastEMD-CCA<sup>2</sup> and FastCCA in MATLAB and C++, are presented in Figs. 11, 12 and 13.

Screen-shots of EEG3's Fp1 channel display in the C++ programming language, after artifact correction through FastEMD-CCA<sup>2</sup> and FastCCA, are shown in Figs. 14 and 15.

Figs. 16, 17, 18, 19 and 20 visualize multichannel EEG signal of EEG 3, before and after artifact correction using FORCE in MATLAB, FastEMD-CCA<sup>2</sup> and FastCCA in MATLAB and C++.

The average error rate, accuracy, sensitivity, and specificity obtained through offline visual inspection on the corrected EEG signals are presented in Table 1.

##### 3.1.1. Discussion

FastEMD-CCA<sup>2</sup> has achieved similar performance measures in both MATLAB and C++, i.e., an error rate of 2.10%, accuracy of 97.9%, sensitivity of

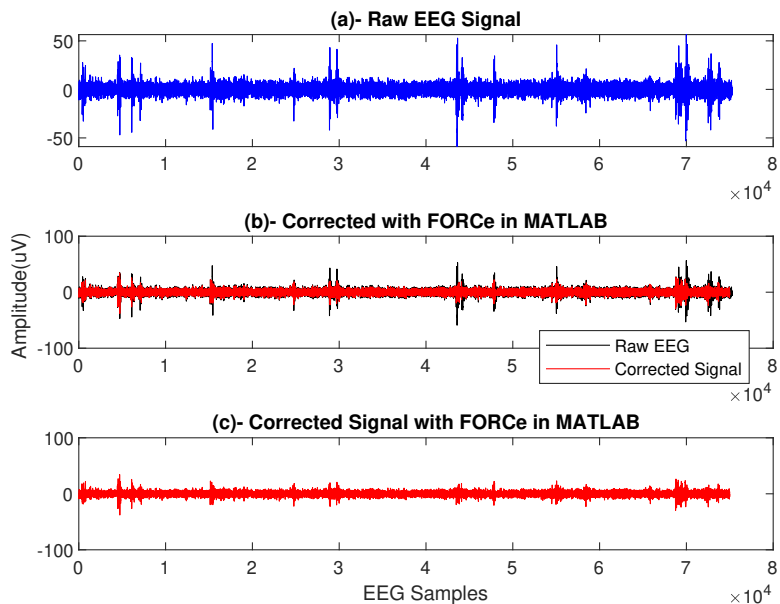


Fig. 11: Recovered EEG Signal from Fp1 Channel through FORCE in MATLAB

97.65%, and specificity of 99.22%. Similarly, FastCCA has achieved an error rate of 0.53%, accuracy of 99.47%, sensitivity of 99.44%, and specificity of 99.74%, in both MATLAB and C++. Achieving similar performance measures in removing the eye blink artifacts in both platforms is expected, proving that the algorithms are implemented correctly in both MATLAB and C++.

The accuracy represents the ability of the proposed algorithms in accurately identifying the eyeblink artifacts and uncontaminated EEG segments, in order for it to remove the eyeblink artifacts. The proposed algorithms have achieved an average of 97.9% and 99.47% of accuracy compared to 91.7% by FORCE. The proposed algorithms achieved a slightly higher accuracy level compared to the FORCE algorithm. Failing to correctly identify the eyeblink artifacts and EEG segments are interpreted through the error rate produced by the algorithms, where it counts the chance of the algorithms to miss an eyeblink artifact. FastEMD-CCA<sup>2</sup> produced an average error rate of 2.10%, FastCCA

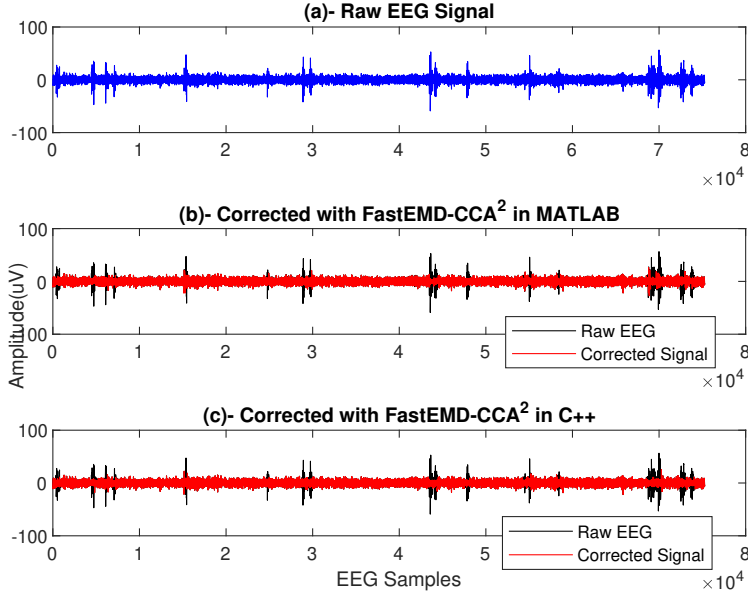


Fig. 12: Recovered EEG Signal from Fp1 Channel through FastEMD-CCA<sup>2</sup> in MATLAB and C++

with 0.53%, while FORCE yield 8.30%. This denotes that these algorithms are still susceptible to miss out eyeblink artifacts. FORCE records the highest error rate and lowest accuracy level on average among the three algorithms, which means FORCE has a higher probability of identifying eye blink artifacts or EEG segments erroneously.

Apart from error rate and accuracy, sensitivity is used to measure how precise the algorithms are in correctly identifying and removing the eye blink artifacts compared to the actual number of observed eye blink artifacts. Similarly, specificity is used to measure the algorithms' precision in properly recognizing EEG segments and retaining them. Again FORCE records the lowest sensitivity on average for both datasets, but not significantly low. The results indicate that the proposed algorithms, FastCCA and FastEMD-CCA<sup>2</sup> have achieved 99.44% and 97.65% of sensitivity respectively, which are 9.97% and 8.18% higher than

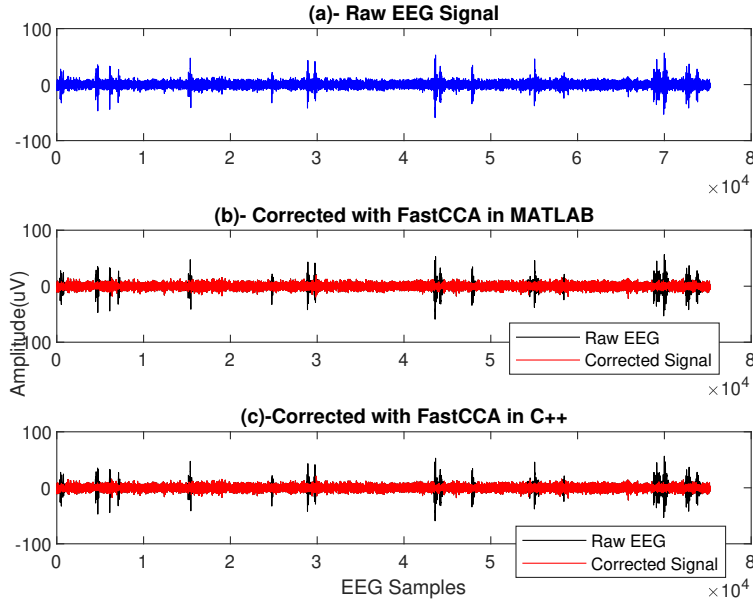


Fig. 13: Recovered EEG Signal from Fp1 Channel through FastCCA in MATLAB and C++

that of FORCE algorithm. This demonstrates that FastCCA and FastEMD-CCA2 were able to detect and remove eye blink artifacts more effectively than FORCE. Manually adjusted threshold values are used to classify or make a binary judgement if an IC is artifactual in FORCE during ICA application on the wavelet coefficients. So, having manually adjusted thresholds may lead to detection errors, thereby some artifacts are not removed.

Specificity is the ratio of undistorted artifact-free EEG segments before and after artifact elimination. The ideal expectation is to have these portions undistorted after artifacts have been removed. Through visual inspection, FastCCA, FastEMD-CCA<sup>2</sup> and FORCE records an average specificity of 99.74%, 99.22% and 98.65% respectively. This signifies that these algorithms do not introduce much distortion to the EEG signals' neural information under evaluation via visual inspection. However, visual inspection alone is not sufficient to confirm if the artifact-free EEG segments are undistorted. So, quantitative analysis on the

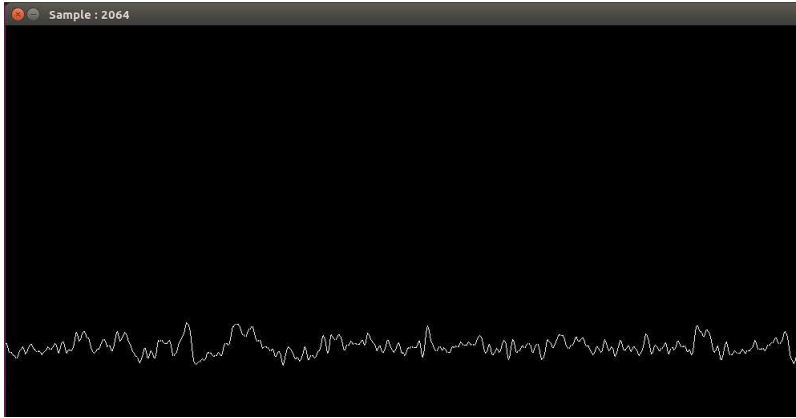


Fig. 14: Display of Recovered EEG Signal from Fp1 Channel through FastEMD-CCA<sup>2</sup> in C++

Table 1: Average Performance Metrics of FORCe, FastEMD-CCA<sup>2</sup> and FastCCA in MATLAB and C++

Average (%)						
Algorithm	FORCe		FastEMD-CCA <sup>2</sup>		FastCCA	
	MATLAB		MATLAB	C++	MATLAB	C++
<b>Error Rate</b>	8.30		2.10	2.10	0.53	0.53
<b>Accuracy</b>	91.70		97.9	97.9	99.47	99.47
<b>Sensitivity</b>	89.47		97.65	97.65	99.44	99.44
<b>Specificity</b>	98.65		99.22	99.22	99.74	99.74

artifact-free EEG segments is conducted as well, which is analysed and discussed in subsection 3.2.

Despite the fact that the FastEMD-CCA<sup>2</sup> and FastCCA algorithms achieve attractive performance scores, these algorithms still fail to detect and remove some of the eye blink artifacts. FastEMD-CCA<sup>2</sup> was unable to detect and remove nearly 137 eye blink artifact events, which accounts for about 2.10% of error rate. The undetected and unremoved eye blink artifacts are analyzed to determine the failure’s root cause, although the error rate is relatively low. In FastEMD-CCA<sup>2</sup>, a general eye blink template is assumed to closely correlates

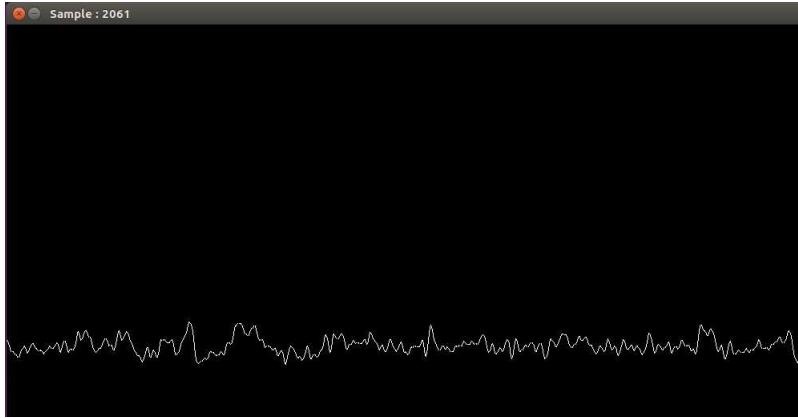


Fig. 15: Display of Recovered EEG Signal from Fp1 Channel through FastCCA in C++

with all eye blink artifact events in that EEG signal, so it is used as a reference template to locate the remaining eye blink artifacts. The invalidity of the assumption mostly causes failure cases. Theoretically, the correlation between the general eye blink artifact template and every EEG segment that is contaminated by eye blink artifact is expected to be  $\geq 0.9$ . In a practical scenario, this is not the case, whereby observations and validation on the real EEG signals revealed that the correlation often lies in the range of 0.4 to 0.6. Hence, using an artifact template with a low correlation threshold of 0.4 to rule out if an EEG segment is contaminated by eye blink artifact may fail occasionally. For instance, if an EEG segment contaminated with eye blink artifact exhibits a correlation of lower than 0.4 with the eye blink artifact template, this EEG segment will not be identified as containing an eye blink artifact, producing a false negative scenario. Similarly, if an uncontaminated EEG segment exhibits a correlation of more than 0.4 with the eye blink artifact template; this EEG segment will be identified as containing an eye blink artifact, although it does not, developing a false positive scenario. Subsequent application of CCA to this EEG segment to remove eye blink artifact is irrelevant and may cause loss of neural information. Whereas in FastCCA, CCA is directly applied to EEG segments identified contaminated by eye blink artifacts with the existing eADA

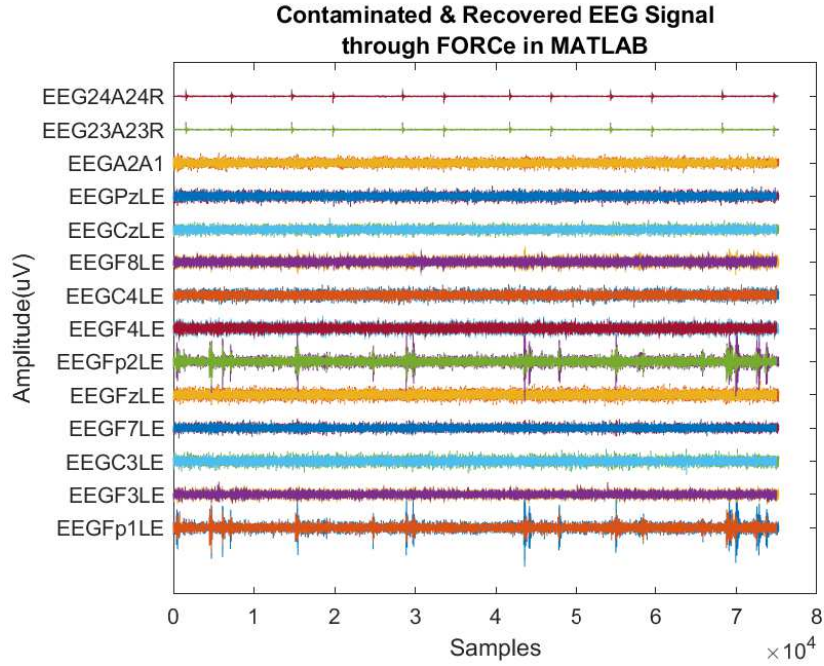


Fig. 16: Recovered Multichannel EEG Signal through FastEMD-CCA<sup>2</sup> in MATLAB

algorithm. This allows an adaptive detection and removal of eye blink artifacts for every event of blink.

This explains why the average accuracy, sensitivity, and specificity of FastEMD-CCA<sup>2</sup> are lower than FastCCA by 1.57%, 1.79%, and 0.52%, respectively. In summary, it can be concluded that in FastEMD-CCA<sup>2</sup>, eye blink artifacts that are irregular in pattern compared with the eye blink template extracted are not classified as eye blink artifacts, thereby not getting removed. Similarly, when an artifact-free EEG segment exhibits an artifact-like pattern, this segment is misclassified as an eye blink artifact and subjected to eye blink artifact rejection, although it isn't contaminated by an eye blink artifact. The high accuracy, sensitivity, and specificity level of both algorithms suggest that these approaches are reliable in detecting and removing eye blink artifacts in real-time; however, FastCCA is a better choice because of its better performance over FastEMD-CCA<sup>2</sup>.



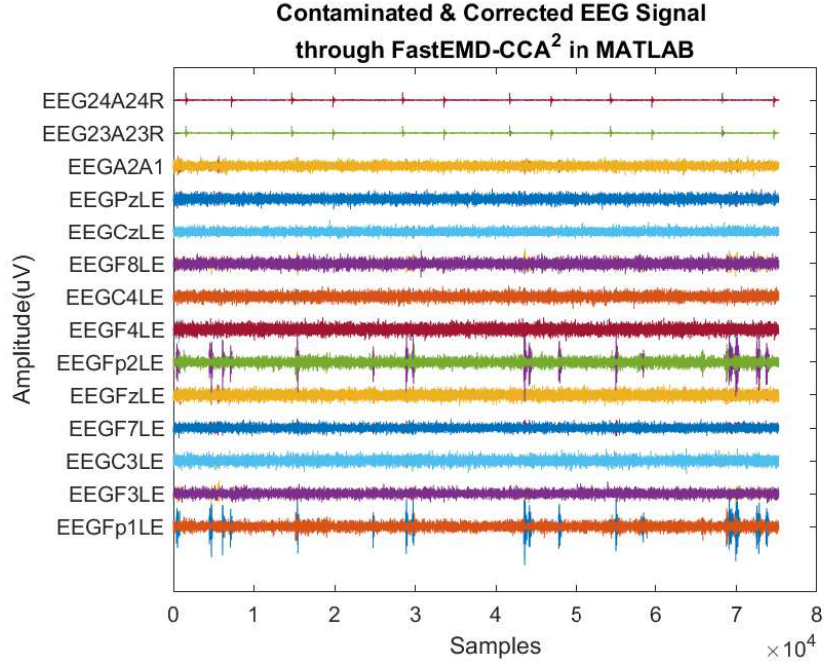


Fig. 17: Recovered Multichannel EEG Signal through FastEMD-CCA<sup>2</sup> in MATLAB

### 3.2. Offline Analysis Results through Quantitative Evaluation

The proposed algorithms' performance in preserving the underlying neural information of an EEG signal is evaluated on randomly selected artifact-free EEG segments after artifact correction. One artifact-free Fp1 EEG segment is evaluated on every EEG signal. Twenty-seven EEG signals out of the 60 EEG signals are excluded from this analysis. Too many eye blink artifacts contaminate them; thus, selecting a proper artifact-free EEG segment is challenging. Hence, only 33 EEG signals out of 60 EEG signals are used for quantitative analysis.

Figs. 21 and 22 show a visual comparison of a short segment of EEG recording (EEG 3), corrected using FORCE, FastEMD-CCA<sup>2</sup> and FastCCA in MATLAB and C++. From Fig. 21, it can be seen that the artifact is not entirely removed with FORCE; it was suppressed instead, while the proposed algorithms can remove eyeblink artifacts effectively. Fig. 22 shows a

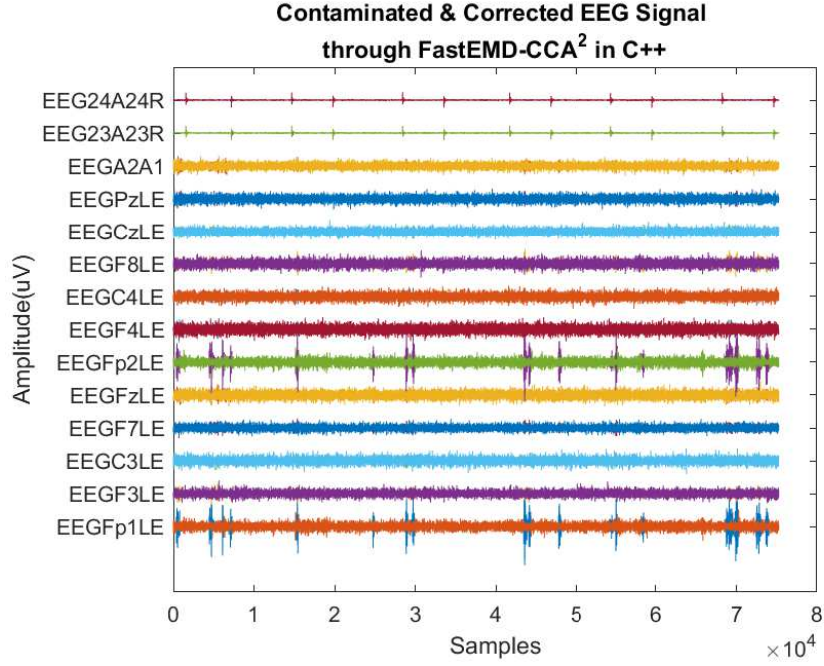


Fig. 18: Recovered Multichannel EEG Signal through FastEMD-CCA<sup>2</sup> in C++

portion of the reconstructed artifact-free segment that overlaps well with the raw EEG segment. Visual inspection on the removed eyeblink artifacts and the reconstructed artifact-free EEG segment shows that the proposed algorithms effectively remove eyeblink artifacts and preserve the neural information.

Table 2 presents the average CC, RMSE and Similarity Index ( $\eta_{dB}$ ) of the artifact-free Fp1 EEG segments.

Table 2: Average CC, RMSE and  $\eta_{dB}$  of FORCe, FastEMD-CCA<sup>2</sup> and FastCCA in MATLAB and C++

Algorithm	Average					
	FORCe		FastEMD-CCA <sup>2</sup>		FastCCA	
	MATLAB		MATLAB	C++	MATLAB	C++
Correlation Coefficient (CC)	0.9582		0.9992	0.9999	0.9993	0.9998
Root Mean Square Error (RMSE)	1.2454		0.1572	0.0154	0.1010	0.0227
Similarity Index ( $\eta_{dB}$ )	-0.7002		-0.3132	0.0010	-0.1377	0.0009

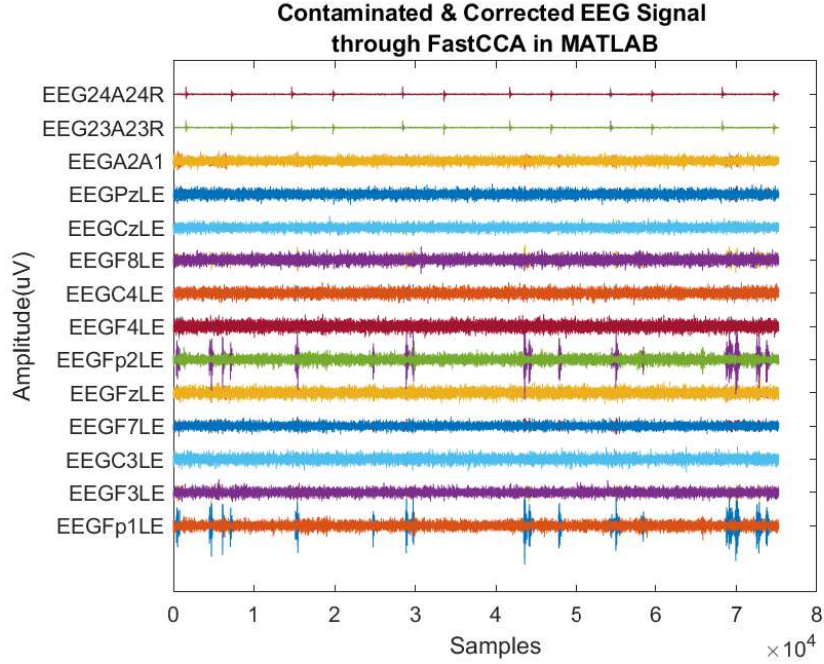


Fig. 19: Recovered Multichannel EEG Signal through FastCCA in MATLAB

### 3.2.1. Discussion

The artifact-free EEG segments in an EEG signal are expected to be unaffected, and the neural information in these segments should remain intact after artifact elimination is performed. Ideally, a perfect artifact removal algorithm will produce CC value of 1, RMSE, and  $\eta_{dB}$  of 0, but this is impractical in real scenarios. Hence, the proposed algorithms' effectiveness in preserving the underlying EEG information can be deduced through CC value that approaches very close to 1, RMSE, and  $\eta_{dB}$  close to 0. There hasn't been any researches that outline a threshold for CC because a closer CC to 1 denotes the best reconstruction after artifact removal.

Among the three algorithms, FORCE achieved the least average CC and  $\eta_{dB}$ , i.e., 0.9582 and -0.7002, respectively. On top of that, it records the highest RMSE of 1.2454, which denotes that FORCE introduces a higher amount of error to the EEG signals after eye blink artifacts have been removed. The average

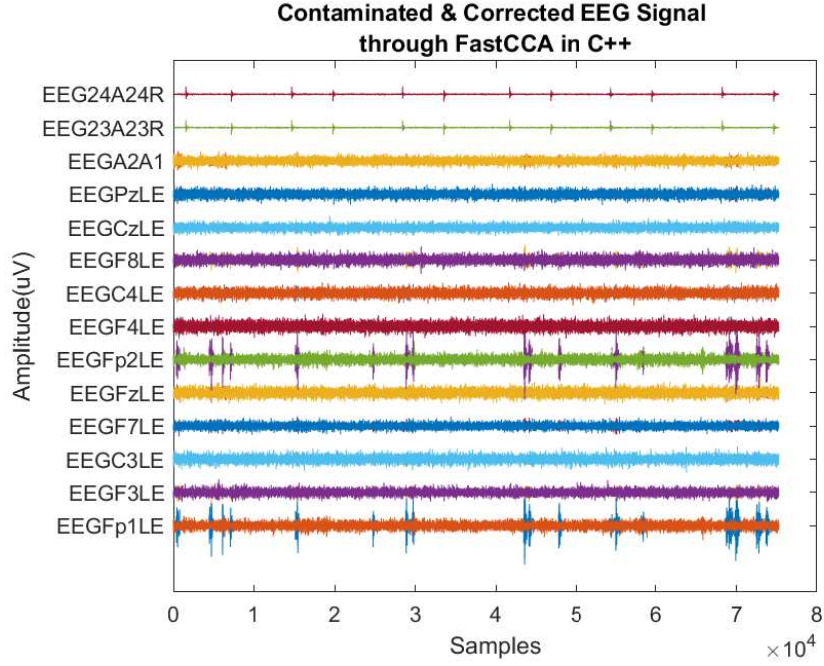


Fig. 20: Recovered Multichannel EEG Signal through FastCCA in C++

RMSE and  $\eta_{dB}$  for FastEMD-CCA<sup>2</sup> and FastCCA are very low compared to FORCE, RMSE ranges between 0.0154 to 0.1572 and  $\eta_{dB}$  range between -0.3132 to 0.0010. FastEMD-CCA<sup>2</sup> in MATLAB recorded the highest average RMSE, but it produced the least RMSE in C++. The average CC value for FastEMD-CCA<sup>2</sup> is 0.9992 and 0.9999 in MATLAB and C++, respectively. FastCCA has achieved an average CC value of 0.9993 in MATLAB and 0.9998 in C++. Quantitative analysis of the artifact-free EEG segments has shown a slight difference in MATLAB and C++ results. This may be due to the QR decomposition and SVD within CCA used in MATLAB and C++ are from different toolboxes and libraries, producing decomposition difference. This difference in decomposition eventually causes a slight difference during EEG reconstruction. However, the average CC values for both proposed algorithms in MATLAB and C++ are comparable, and the difference among them is considered insignificant. In general, the individual CC values produced by FastEMD-CCA<sup>2</sup> and FastCCA in

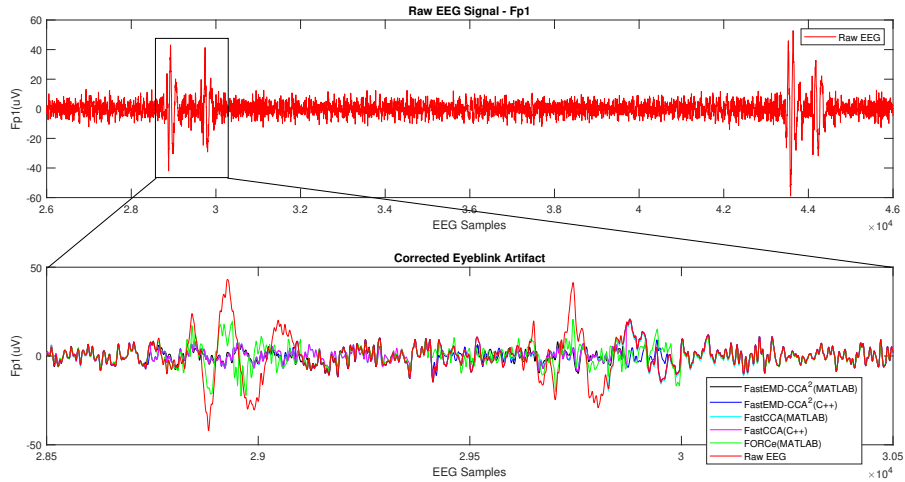


Fig. 21: Corrected Eyeblink Artifact Regions through FORCE, FastEMD-CCA<sup>2</sup> and FastCCA in MATLAB and C++

reconstructing artifact-free segments are more than 0.99 for all 33 analysed EEG signals, in both MATLAB and C++. This indicates that both algorithms do not introduce much distortion to the artifact-free EEG segments during reconstruction in comparison to FORCE, thus reliable in preserving neural information of an EEG signal.

The quantitative results of FastEMD-CCA<sup>2</sup> and FastCCA suggest that they have caused an insignificant loss of neural information to the EEG signals in comparison to FORCE. In FastEMD-CCA<sup>2</sup>, the eADA algorithm is first implemented to identify the first few eyeblink artifact locations, and an eyeblink artifact template is extracted out through FastEMD and CCA. The eyeblink artifact template is cross-correlated with sliding EEG windows, where highly correlated EEG windows are subjected to eyeblink artifact removal. CCA is applied to remove artifacts, only on EEG windows that were confirmed to be contaminated by eyeblink artifacts. Hence, this is the reason that the artifact-free EEG segments are unaffected in FastEMD-CCA<sup>2</sup>, causing negligible loss of neural information during the reconstruction of the neural signal. In FastCCA,

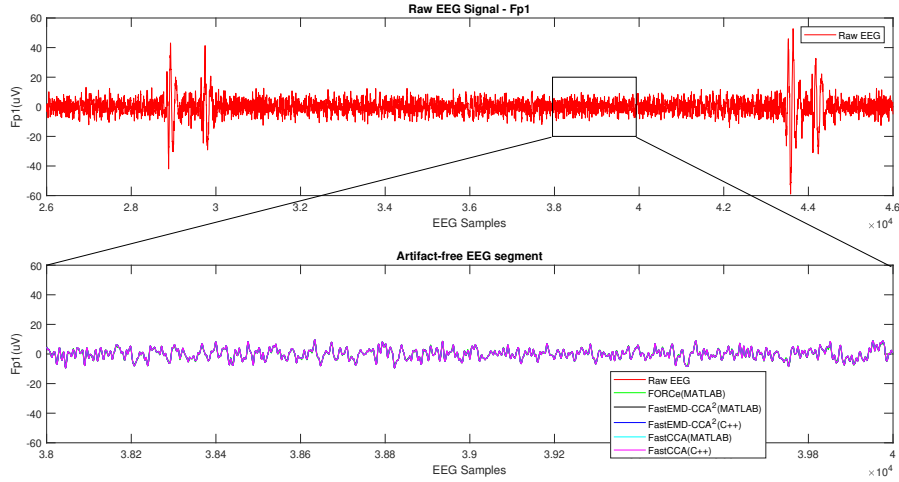


Fig. 22: Reconstructed Artifact-free EEG through FORCE, FastEMD-CCA<sup>2</sup> and FastCCA in MATLAB and C++

CCA is used to remove eye blink artifacts on eye blink artifact locations identified through eADA. Similar to FastEMD-CCA<sup>2</sup>, the artifact-free EEG segments are unaffected, preventing loss of neural information during the reconstruction of a clean EEG signal.

### 3.3. Online Analysis Results

The average computation time taken by these algorithms in MATLAB computing environment and C++ programming language to remove eyeblink artifacts from 14-channel EEG signals are tabulated in Table 3.

Table 3: Average Computation Time of FORCE, FastEMD-CCA<sup>2</sup> and FastCCA in MATLAB and C++

Average Processing Time (s)						
Algorithm	FORCE		FastEMD-CCA <sup>2</sup>		FastCCA	
	MATLAB	C++	MATLAB	C++	MATLAB	C++
Clean Entire EEG	85.10	6.78	3.35	6.87	3.96	
Clean 1s of EEG	0.2725	0.0217	0.0107	0.0220	0.0127	

The proposed algorithms are carried out on every 1.95s epoch of the recorded EEG signal to remove any eye blink artifact that is present. FORCE took about 85.10s in MATLAB to remove eye blink artifacts from 14-channel EEG signals with an average signal length of 312s (Hitachi dataset). This is equivalent to an average processing time of 272.5 milliseconds to process and remove any eye blink artifact from a 1-second length of 14-channel EEG signal (256 sample points x 14 EEG channels). This high processing time of FORCE may not accommodate an online application, although it can remove eye blink artifacts from EEG signals. The average computation time of FastEMD-CCA<sup>2</sup> is 6.78s in MATLAB and 3.35s in C++. It takes an average of 21.7 milliseconds in MATLAB and 10.7 milliseconds in C++ to process and remove any eye blink artifact from a 1-second length of 14-channel EEG signal. The average computation time of FastCCA is 6.87s in MATLAB and 3.96s in C++. The average processing time to process and clean a 1-second length of 14-channel EEG signal by FastCCA is 22 milliseconds in MATLAB and 12.7 milliseconds in C++. Comparing all three algorithms, FastEMD-CCA<sup>2</sup> achieves a shorter processing time of 10.7 milliseconds in C++, followed by FastCCA, with a processing time of 12.7 milliseconds in C++. It has to be noted that although FastCCA is better than FastEMD-CCA<sup>2</sup> as demonstrated by the qualitative and quantitative measurements, it is slower than FastEMD-CCA<sup>2</sup> in terms of processing time. This can be due to the difference in the eye blink artifacts' detection mechanism in these algorithms. In FastEMD-CCA<sup>2</sup>, a general eye blink artifact template is extracted out and used to identify the remaining eye blink artifact events in that EEG signal. Whereas in FastCCA, every eye blink artifact event is continuously searched by the eADA algorithm to be removed by CCA, so justifying the time difference between these algorithms. Considering the average processing time, FastEMD-CCA<sup>2</sup> is a better choice for online applications compared to FastCCA, but again the processing time of FastEMD-CCA<sup>2</sup> is not significantly better than FastCCA. So, both proposed algorithms can be used for applications requiring online removal of eye blink artifacts with high accuracy.

Results clearly show that the eyeblink artifact removal in C++ is well-suited

for real-time implementation. Both proposed algorithms took about 10 to 13 milliseconds on average to clean an EEG epoch of 1s length (256 x 14 EEG sample points) in C++ environment. This is because the proposed algorithms in C++ are designed to stream the EEG signal on a sample by sample basis into a buffer, before the eyeblink artifacts are located and removed. The streaming of EEG sample points into the buffer, and the cleaning process are performed in parallel. Therefore the processing speed of proposed algorithms in C++ is purely dependent on the speed of EEG sample point acquisition and instantaneous processing, hence the processing is considered online. Considering the average processing time, FastEMD-CCA<sup>2</sup> is a better choice for online applications compared to FastCCA, but again the processing time of FastEMD-CCA<sup>2</sup> is not significantly better than FastCCA. So, both proposed algorithms can be used for applications requiring online removal of eyeblink artifacts with high accuracy.

#### 4. Conclusion

FastEMD-CCA<sup>2</sup> and FastCCA are proposed, analysed and discussed in this paper. The performance of these algorithms in removing eye blink artifacts in real-time, and how well the algorithms are able to preserve artifact-free EEG segments, without distorting the neural signal is evaluated. The algorithms are implemented in MATLAB and C++, both in a real-time setting. It appears that both algorithms have achieved instantaneous eye blink artifact elimination, with very fast processing time in C++. The artifact removal accuracy of FastCCA is 1.57% higher than FastEMD-CCA<sup>2</sup> in both MATLAB and C++, indicating FastCCA is slightly better in artifact removal, regardless of the implementation medium, MATLAB or C++. On the other hand, quantitative analysis has pointed out that both FastEMD-CCA<sup>2</sup> and FastCCA are effective in retaining underlying neural information, with insignificant distortion to the artifact-free EEG segments. Thus, any of these two algorithms are suitable to be used in EEG-based applications requiring eye blink artifact removal without loss of



neural information in real-time, with FastCCA is slightly desirable in terms of artifact removal accuracy.

### **Acknowledgment**

The authors would like to thank the Ministry of Education, Malaysia for supporting this research through the Fundamental Research Grant Scheme, FRGS (FRGS/2/2014/TK03/UTP/02/1) and the Higher Institution Centre of Excellence (HICoE) Scheme. The authors would also like to thank Dr. Tahamina Begum from Department of Neurosciences, Hospital Universiti Sains Malaysia for validating the EEG signals after eye blink artifact correction.

### **References**

- [1] V. Alvarez, Therapeutic Coma for the Treatment of Status Epilepticus Timing, Choice of Drug, and Impact On, Springer Nature 2018doi:10.1007/s10309-018-0194-x.
- [2] C. Braiman, E. A. Fridman, M. M. Conte, H. U. Voss, C. S. Reichenbach, T. Reichenbach, N. D. Schiff, C. Braiman, E. A. Fridman, M. M. Conte, H. U. Voss, C. S. Reichenbach, Cortical Response to the Natural Speech Envelope Correlates with Neuroimaging Evidence of Cognition in Severe Brain Injury Report Cortical Response to the Natural Speech Envelope Correlates with Neuroimaging Evidence of Cognition in Severe Brain Injury, *Current Biology* (2018) 1–7doi:10.1016/j.cub.2018.10.057.  
URL <https://doi.org/10.1016/j.cub.2018.10.057>
- [3] J. Claassen, F. S. Taccone, Recommendations on the Use of EEG Monitoring in Critically Ill Patients : Consensus Statement from the Neurointensive Care Section of the ESICM, *Intensive Care Medicine* (2013) 1337–1351doi:10.1007/s00134-013-2938-4.

- [4] U. R. Acharya, S. V. Sree, A. Peng, C. Alvin, J. S. Suri, Use of Principal Component Analysis for Automatic Classification of Epileptic EEG Activities in Wavelet Framework, *Expert Systems With Applications* 39 (10) (2012) 9072–9078. doi:10.1016/j.eswa.2012.02.040.  
URL <http://dx.doi.org/10.1016/j.eswa.2012.02.040>
- [5] E. J. Stepanski, J. K. Wyatt, Use of Sleep Hygiene in the Treatment of Insomnia, *Sleep Medicine Reviews* 7 (3). doi:10.1053/smr.2001.0246.
- [6] R. Flink, B. Pedersen, A. Guekht, K. Malmgren, R. Michelucci, B. Neville, F. Pinto, U. Stephani, C. Ozkara, Review article Guidelines for the use of EEG methodology in the diagnosis of epilepsy International League Against Epilepsy : Commission Report Commission on European Affairs : Subcommittee on European Guidelines (2002) 1–7.
- [7] B. Vivien, X. Paqueron, P. L. Cosquer, O. Langeron, P. Coriat, B. Riou, Detection of Brain Death Onset Using the Bispectral Index in Severely Comatose Patients, *Intensive Care Medicine* (2002) 419–425doi:10.1007/s00134-002-1219-4.
- [8] G. B. Young, The EEG in Coma, *Journal of Clinical Neurophysiology* 17 (5) (2000) 473–485.
- [9] R. I. Tivadar, M. M. Murray, A Primer on Electroencephalography and Event-Related Potentials for Organizational Neuroscience, *Organizational Research Methods* (2018) 1–26doi:10.1177/1094428118804657.
- [10] S. M. Peterson, E. Furuichi, D. P. Ferris, Effects of virtual reality high heights exposure during beam-walking on physiological stress and cognitive loading, *PLOS ONE* (2018) 1–17.
- [11] C. Miniussi, Æ. G. Thut, Combining TMS and EEG Offers New Prospects in Cognitive Neuroscience, *Brain Topography* (May 2014). doi:10.1007/s10548-009-0083-8.

- [12] X. Jiang, G. B. Bian, Z. Tian, Removal of Artifacts from EEG signals: A review, *Sensors (Switzerland)* 19 (5) (2019) 1–18. doi:10.3390/s19050987.
- [13] M. M. N. Mannan, M. Y. Jeong, M. A. Kamran, Hybrid ICA-Regression: Automatic Identification and Removal of Ocular Artifacts from Electroencephalographic Signals, *Frontiers in Human Neuroscience* 10 (5) (2016) 1–17. doi:10.3389/fnhum.2016.00193.
- [14] F. Minguillon, Jesus and Lopez-Gordo, M Angel and Pelayo, Trends in EEG - BCI for Daily - Life : Requirements for Artifact Removal, *Biomedical Signal Processing and Control* (October). doi:10.1016/j.bspc.2016.09.005.
- [15] S. Kanoga, Y. Mitsukura, Review of Artifact Rejection Methods for Electroencephalographic Systems, *Electroencephalography* doi:10.5772/68023.
- [16] K. Islam, A. Rastegarnia, Z. Yang, Methods for Artifact Detection and Removal from Scalp EEG : A Review, *Neurophysiologie Clinique / Clinical Neurophysiology* (November 2017). doi:10.1016/j.neucli.2016.07.002.  
URL <http://dx.doi.org/10.1016/j.neucli.2016.07.002>
- [17] J. A. Uriguen, B. Garcia-Zapirain, EEG Artifact Removal State of the Art and Guidelines, *Journal of neural engineering* 12 (3) (2015) 031001.
- [18] M. Fatourechi, A. Bashashati, R. K. Ward, G. E. Birch, EMG and EOG artifacts in brain computer interface systems: A survey, *Clinical Neurophysiology* 118 (3) (2007) 480–494. doi:10.1016/j.clinph.2006.10.019.
- [19] A. Bashashati, B. Nouredin, R. K. Ward, P. Lawrence, G. E. Birch, Effect of eye-blinks on a self-paced brain interface design, *Clinical Neurophysiology* 118 (7) (2007) 1639–1647. doi:10.1016/j.clinph.2007.03.020.
- [20] I. I. Goncharova, D. J. Mcfarland, T. M. Vaughan, J. R. Wolpaw, EMG Contamination of EEG : Spectral and Topographical Characteristics, *Clinical Neurophysiology* 114 (2003) 1580–1593. doi:10.1016/S1388-2457(03)00093-2.

- [21] R. J. Huster, Z. N. Mokom, S. Enriquez-geppert, C. S. Herrmann, Brain Computer Interfaces for EEG Neurofeedback : Peculiarities and Solutions, International Journal of Psychophysiology.
- [22] L. H. Sherlin, C. Kerson, Investigations in Neuromodulation , Neurofeedback and Applied Neuroscience Neurofeedback and Basic Learning Theory : Implications for Research and Practice, Journal of Neurotherapy (December 2012) (2011) 37–41. doi:10.1080/10874208.2011.623089.
- [23] A. Mansouri, S. P. Singh, K. Sayood, Online EEG Seizure Detection and Localization, Algorithms.
- [24] F. Forooghifar, A. Aminifar, L. Cammoun, I. Wisniewski, C. Ciomas, P. Ryvlin, D. Atienza, A Self-Aware Epilepsy Monitoring System for Real-Time Epileptic Seizure Detection, Mobile Networks and Applications.
- [25] O. Hanosh, Real-Time Epileptic Seizure Detection during Sleep using Passive Infrared ( PIR ) Sensors, IEEE Sensors Journal PP (c) (2019) 1. doi:10.1109/JSEN.2019.2907664.
- [26] T. K. M.M. Hartmann, K. Schindler, T.A. Gebbink, G. Gritsch, PureEEG : Automatic EEG Artifact Removal, Neurophysiologie Clinique / Clinical Neurophysiology 44 (5) (2014) 479–490. doi:10.1016/j.neucli.2014.09.001. URL <http://dx.doi.org/10.1016/j.neucli.2014.09.001>
- [27] E. Acar, C. Aykut-bingol, H. Bingol, R. Bro, Multiway Analysis of Epilepsy Tensors, Bioinformatics 23 (2006) (2007) 10–18. doi:10.1093/bioinformatics/btm210.
- [28] H. Adeli, S. Ghosh-dastidar, N. Dadmehr, A Spatio-Temporal Wavelet-Chaos Methodology for EEG-Based Diagnosis of Alzheimer’s Disease, Neuroscience Letters 444 (2008) 190–194. doi:10.1016/j.neulet.2008.08.008.
- [29] M. A. Bell, K. Cuevas, Using EEG to Study Cognitive Development : Issues and Practices, Journal of Cognition and Development 13 (3) (2012) 281–294. doi:10.1080/15248372.2012.691143.

- [30] H.-A. T. Nguyen, J. Musson, F. Li, W. Wang, G. Zhang, R. Xu, C. Richey, T. Schnell, F. D. McKenzie, J. Li, EOG Artifact Removal Using a Wavelet Neural Network, *Neurocomputing* 97 (2012) 374–389.
- [31] I. Daly, R. Scherer, M. Billinger, G. Müller-Putz, FORCe: Fully Online and Automated Artifact Removal for Brain Computer Interfacing, *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 23 (5) (2015) 725–736.
- [32] L. Pion-Tonachini, S.-H. Hsu, C.-Y. Chang, T.-P. Jung, S. Makeig, Online Automatic Artifact Rejection Using the Real-time EEG Source-mapping Toolbox (REST), in: 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2018, pp. 106–109.
- [33] A. Egambaram, N. Badruddin, V. S. Asirvadam, E. Fauvet, C. Stolz, T. Begum, Unsupervised Eye Blink Artifact Identification in Electroencephalogram, in: TENCON 2018-2018 IEEE Region 10 Conference, IEEE, 2018, pp. 2148–2152.
- [34] A. Egambaram, N. Badruddin, V. S. Asirvadam, T. Begum, E. Fauvet, C. Stolz, FastEMDCCA algorithm for unsupervised and fast removal of eyeblink artifacts from electroencephalogram, *Biomedical Signal Processing and Control* 57 (2020) 101692. doi:10.1016/j.bspc.2019.101692.  
URL <https://doi.org/10.1016/j.bspc.2019.101692>
- [35] S. Kudrle, M. Proulx, P. Carrieres, M. Lopez, Fingerprinting for solving A/V synchronization issues within broadcast environments, *SMPTE Motion Imaging Journal* (2010) 279–296doi:10.5594/m001370.
- [36] D. B. Stone, G. Tamburro, P. Fiedler, J. Haueisen, S. Comani, Automatic Removal of Physiological Artifacts in EEG: The Optimized Fingerprint Method for Sports Science Applications, *Frontiers in Human Neuroscience* 12 (2018) 96.

- [37] S. O'Regan, S. Faul, W. Marnane, Automatic Detection of EEG Artefacts Arising from Head Movements Using EEG and Gyroscope Signals, *Medical Engineering & Physics* 35 (7) (2013) 867–874.
- [38] M. H. Soomro, N. Badruddin, M. Z. Yusoff, M. A. Jatoi, Automatic Eye-Blink Artifact Removal Method Based on EMD-CCA, in: 2013 ICME International Conference on Complex Medical Engineering, CME 2013, no. April, 2013, pp. 186–190. doi:10.1109/ICCME.2013.6548236.