Multidimensional Feature Optimization Based Eye Blink Detection Under Epileptiform Discharges

Meng Wang, Jianhui Wang, Xiaonan Cui, Tianlei Wang, Tiejia Jiang, Feng Gao, and Jiuwen Cao, *Senior Member, IEEE*

***Abstract*— *Objectives:* Eye blink artifact detection in scalp electroencephalogram (EEG) of epilepsy patients is challenging due to its similar waveforms to epileptiform discharges. Developing an accurate detection method is urgent and critical. *Methods:* In this paper, we proposed a novel multi-dimensional feature optimization based eye blink artifact detection algorithm for EEGs containing rich epileptiform discharges. An unsupervised clustering algorithm based on smoothed nonlinear energy operator (SNEO) and variational mode extraction (VME) is proposed to detect epileptiform discharges in the frontal leads. Then, multi-dimensional time/frequency EEG features extracted from forehead electrodes (FP1 and FP2 channels) combining with the improved VME (IVME) threshold are derived for EEG representation. A variance filtering method is further applied for discriminative feature selection and a machine learning model is finally learned to perform detection. *Results:* Experiments on EEGs of 16 subjects from the Children's Hospital of Zhejiang University School of Medicine (CHZU) show that our method achieves the highest average sensitivity, specificity and accuracy of 95.04, 89.52, and 93.01, respectively. That outperforms 5 recent and state-of-the-art (SOTA) eye blink detection algorithms. *Significance:* The proposed method is robust in eye blink artifact detection for EEGs containing high-frequency epileptiform discharges. It is also effective in dealing with individual differences in EEGs, which is usually ignored in conventional methods.**

***Index Terms*—Epileptiform discharge, improved variational mode extraction, variance filtering, multi-dimensional EEG feature optimization.**

Manuscript received December 25, 2021; revised March 12, 2022 and March 23, 2022; accepted March 25, 2022. Date of publication April 1, 2022; date of current version April 12, 2022. This work was supported in part by the National Key Research and Development Program of China under Grant 2021YFE0100100 and Grant 2021YFE0205400, in part by the National Natural Science Foundation of China under Grant U1909209, in part by the Open Research Projects of Zhejiang Laboratory under Grant 2021MC0AB04, in part by the Natural Science Key Foundation of Zhejiang Province under Grant LZ22F030002, and in part by the Key Research and Development Program of Zhejiang Province under Grant 2020C03038. *(Meng Wang and Tiejia Jiang contributed equally to this work.) (Corresponding author: Jiuwen Cao.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Children's Hospital, Zhejiang University School of Medicine, and registered in the Chinese Clinical Trail Registry (ChiCTR1900028804).

Meng Wang, Jianhui Wang, Xiaonan Cui, and Tianlei Wang are with the Machine Learning and I-Health International Cooperation Base of Zhejiang Province and the Artificial Intelligence Institute, Hangzhou Dianzi University, Hangzhou 310018, China (e-mail: 15968138702@163.com; 15968194550@163.com; xiaonan18@foxmail.com; tianleiwang@hdu.edu.cn).

Tiejia Jiang and Feng Gao are with the Department of Neurology, The Children's Hospital, Zhejiang University School of Medicine, National Clinical Research Center for Child Health, Hangzhou 310003, China (e-mail: jiangyouze@zju.edu.cn; epilepsy@zju.edu.cn).

Jiuwen Cao is with the Machine Learning and I-Health International Cooperation Base of Zhejiang Province and the Artificial Intelligence Institute, Hangzhou Dianzi University, Hangzhou 310018, China, and also with the Research Center for Intelligent Sensing, Zhejiang Laboratory, Hangzhou 311100, China (e-mail: jwcao@hdu.edu.cn).

Digital Object Identifier 10.1109/TNSRE.2022.3164126

I. Introduction

EYE blink artifact detection and elimination are important in EEG analysis [1], [2], particularly for EEG-based nervous system diseases analysis [3]–[7], such as epilepsy, cerebellitis, etc. Eye blinks are generally caused by blinking or eye movement, and are mainly obvious in the front of the scalp. The signal strength of electrooculogram (EOG) usually deceases when the lead position to the eye increases. Eye blinks are usually difficult to avoid during EEG recording [8]–[10]. The artifact detection and removal are always challenging. The reason behind is that the characteristics of eye blinks are similar to neuron discharges or other normal EEGs caused by nerve system diseases [11], such as epileptiform discharges of epilepsy patients.

Adaptive filter is one of the representative methods in eye blink detection [9], [11]–[13], but suffers the limitation of requiring additional correlated reference electrodes [11]. A multi-window derivation and summation method (MSDW) exploiting the width statistic of eye blink waveforms is presented in [9]. Winner filter can overcome the additional reference channel issue, but it requires initial calibration and cannot be applied online. In [14], a minimum noise estimate filter is developed to remove eye blink artifacts by an iterative threshold method. In [15], the maximum overlap wavelet transform decomposition (MODWT) is used to perform time-frequency analysis on EEGs, which has been applied to filter single channel eye blinks and muscle artifacts. Based on the stationary wavelet transform (SWT), a method [16] on signal skewness has been developed to remove eye blinks. But due to the influence of individual differences, the skewness threshold for each individual fluctuates significantly, leading to poor performance. Without using prior knowledge and reference electrodes, blind source separation algorithm (BSS) is regarded as the most common and effective method in eye blink artifact detection for multi-channel EEGs [17], [18]. Principal component analysis (PCA) [19] and independent component analysis (ICA) [20] are the most recognized BSS algorithms in artifact detection. In [19], it is shown that PCA is better than regression-based blink artifacts elimination. But EEGs and eye blinking artifacts may not always meet the premise of PCA, such as obeying Gaussian distribution, orthogonal principal components. Different to PCA, ICA is built on the assumption that the sources are independent to each other, and achieves the decomposition by a linear transformation. It is stated in [21] that the bidirectional pollution effect of EEG and eye blinks has little effect to ICA. But the statistical independence requirement is built on large amount of EEG data, directly leading to a sharp increase in computing burden [12], [22]. Manual intervention on exact source amount and data distribution is another issue in ICA. To alleviate this deficiency, ICA combining with multiple empirical model decomposition (IMEMD) is developed in [23].

Machine learning based intelligent algorithms recently became popular in eye blink detection. In [24], the genetic algorithm is applied to EEG feature optimization and the Parzen window detector is employed for eye blink artifacts detection. [25] uses discrete wavelet transform combining with a meta-heuristic threshold optimization to achieve artifacts detection. In [26], ICA is combined with support vector machine (ICA-SVM) to remove blink artifacts, which is further extended to wavelet ICA (WICA-SVM) in [27]. Autoencoder (AE) has been applied for eye blink EEG feature learning in [28], [29], where SVM (SVM-AE) and recursive least square adaptive filter are respectively adopted as the detector. Aiming at low-cost single channel EEG application, a computationally-efficient algorithm with variational mode extraction (VME) and discrete wavelet transform is proposed for eye blink detection and removal in [30]. Deep belief network combining with AE has shown significant performance in [31]. Recently, an unsupervised clustering method based on a cascaded hybrid thresholding and Gaussian mixture model (GMM) is developed for eye blink detection for healthy and patients' EEGs, respectively [32].

Although rich results on eye blink detection and elimination have been reported in the past, it is noted that most are on clean or normal EEGs. Practically, eye blink detection in EEGs of patients suffered from nervous system diseases is more worthy of study and challenging. Eye blink artifacts usually have more serious misleading impact on the EEG based diagnosis of nervous disease. For instance, for frontal epilepsy, the pathological EEG discharges appearing in the frontal electrodes have very similar waveforms to eye blinks, making the detection crucial in disease analysis. Existing representative methods, including adaptive filtering [12], generic algorithm [24], MSDW [9], SVM-AE [28], VME [30], Revised VME [33], iterative template [34], ICA-SVM [26], WICA-SVM [27] and GMM [32], are generally not suitable, as shown in the comparative study presented in Table I. To address this issue, a multi-dimensional EEG feature optimization (MDF) based epileptiform discharges and eye blink artifacts detection algorithm is recently developed in [35]. An unsupervised clustering is preliminarily applied to screen out epileptiform discharges on EEGs of epilepsy patients, and then a machine learning model is trained for eye blink detection. But MDF and all aforementioned methods generally become poor when EEGs contain consecutive and high-frequency epileptiform discharges, as compared in Table I.

TABLE I
COMPARATIVE ANALYSIS OF EYE BLINK DETECTION METHODS

Methods	Adapt to High-Frequency Discharges	Epileptiform Discharge Screening	Automatic	Run On a few Channels	Differential Treatment
Adaptive filtering [12]			√	√	
Generic method [24]			√	√	√
MSDW [9]			√	√	√
SVM-AE [28]			√	√	
VME [30]			√	√	
Revised VME [33]			√	√	
Iterative template [34]			√	√	√
ICA-SVM [26]			√		
WICA-SVM [27]			√		
GMM [32]			√	√	√
MDF [35]		√	√	√	√
Proposed method	√	√	√	√	√

Inspired by MDF [35] and VME [30], a novel eye blink detection algorithm that is robust to EEG epileptiform discharge affection is developed in this paper. To enhance the epileptiform discharges and eye blink artifacts clustering performance, the hybrid features taken from raw EEGs and filtered EEGs by SNEO and VME are firstly derived for epileptiform discharges filtering. An improved VME (IVME) based robust threshold algorithm is then developed to overcome the less effective issue of VME in handling EEGs containing continuous eye blink artifacts, large offset, or multiple epileptiform discharges. Finally, a feature optimization selection with variance filtering is performed on the improved VME threshold and multi-dimensional time-frequency EEG features [35] for machine learning based detection model training. Fig. 1 shows the general flowchart of our method. Comparing with existing state-of-the-art (SOTA) eye blink detection algorithms, the main contributions of the paper are threefold:

- A robust K-means clustering algorithm is developed to screen out epileptiform discharges to reduce the affection to eye blink detection, where 20 discriminative statistical features, including the maximum value of the monotonically increasing slope (MSMI), the absolute average of the first-order difference (AAF), etc., are extracted from the partial slope sequences of raw EEG and filtered EEG by SNEO and VME for clustering.
- An improved VME (IVME) algorithm that can adapt to consecutive eye blink artifacts, large EEG offset, frequent epileptiform discharges is developed. A robust threshold that exploits the properties of eye blink duration and peak range ratio is derived for eye blink artifact characterization.
- Combining the improved VME threshold feature with multi-dimensional time-frequency EEG characteristics in [35], a variance filter based optimization that can

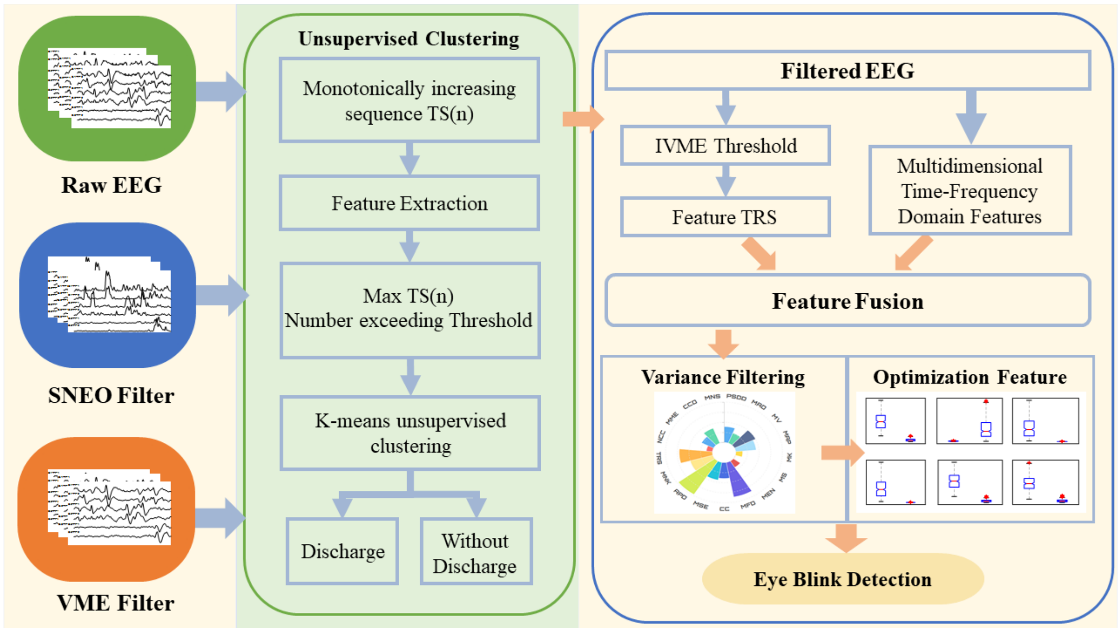


Fig. 1. The proposed eye blink artifact detection algorithm.

TABLE II
SPECIFICATIONS OF THE CHZU EYE BLINK DETECTION DATABASE

Subjects	Ages	Descriptions	Epileptiform discharge	Eye blink/ normal EEG	Duration
S1	3y10m	Healthy	0	170/233	15h 48min
S2	5y5m	Epilepsy	112	341/187	15h 47min
S3	6y1m	Cerebellitis	0	275/334	12h 26min
S4	8y2m	Epilepsy	0	172/274	16h 18min
S5	7y9m	Epilepsy	0	166/197	1h 55min
S6	4y2m	Bronchitis	0	153/414	15h 5min
S7	9y0m	Epilepsy	43	56/239	2h 1min
S8	8y8m	Epilepsy	0	315/406	1h 55min
S9	12y5m	Tic disorder	0	265/494	15h 5min
S10	7y9m	Dizzy	0	178/490	2h 1min
S11	7y5m	Epilepsy	104	102/104	2h 4min
S12	7y5m	Epilepsy	68	180/73	15h 57min
S13	9y4m	Epilepsy	134	210/94	1h 34min
S14	4y2m	Epilepsy	179	310/145	15h 34min
S15	12y0m	Epilepsy	230	315/122	16h 05min
S16	11y1m	Epilepsy	104	279/143	15h 32min

effectively adapt to subject EEG difference is developed for feature selection and machine learning based eye blink artifact detection model training.

We evaluate the proposed eye blink detection algorithm on the long-term recorded EEGs of 16 children from the Children's Hospital of Zhejiang University School of Medicine (CHZU), where among these subjects, 11 patients with epilepsy, 1 with cerebelitis, 1 with bronchitis, 1 with Tic disorder, 1 with dizzy, and the rest are healthy. Comparisons to many state-of-the-art (SOTA) eye blink artifact detection algorithms are presented to show the robustness of the proposed algorithm in dealing with EEGs containing epileptiform discharges.

II. DATABASE AND METHODOLOGY

A. Database

The EEG data used in the paper are recorded from 16 subjects in the Children's Hospital of Zhejiang University School of Medicine (CHZU) [36]. Table II shows the detailed specifications of the database. The age ranges from 3 to 12 years old. The EEG duration for 9 children is longer than 15 hours, 1 is longer than 12 hours, and the rest are longer than 1 hour. The EEG recording follows the international 10-20 standard, by a 16-bit A/D converter to collect 20 to 64 channel records with 1000 Hz. Among all subjects, 8 subjects contain epileptiform discharges. The sample numbers of epileptiform discharges and eye blink artifacts of each subject are presented in the table. It is noted that all data are marked by experienced neuroscientists from CHZU.

B. Methodology

The proposed eye blink detection algorithm generally contains three parts: 1) Epileptiform discharge filtering, an unsupervised K-means clustering based on hybird features extracted from the original EEG and the filtered EEGs by SNEO and VME is developed for epileptiform discharge filtering, 2) Multi-dimensional EEG feature extraction, an improved VME threshold feature combining with multi-dimensional EEG features in [35] is proposed for signal representation, 3) Feature optimization and eye blink detection, the variance filtering method is applied to select discriminative features for model learning and SVM is trained for eye blink artifact detection.

1) Epileptiform Discharge Filtering: Eye blink artifacts normally exist in the frontal region of brain, usually exhibit strong power in the two prefrontal lobe electrodes (FP1, FP2) in EEG. The signal amplitude is around 100∼300 uV and the wave width is generally between 250∼450 ms. Eye blink artifact has a very similar waveform to epileptiform discharges, spikes, and other signature waves, where they are critical in nervous system disease analysis, such as epilepsy, encephalitis, etc, as shown in Fig. 2. The appearance of eye blinks usually affects the analysis, making eye blinks detection

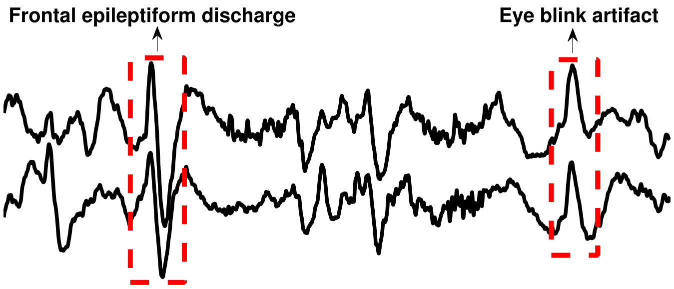


Fig. 2. Frontal epileptiform discharges and Eye blink artifact.

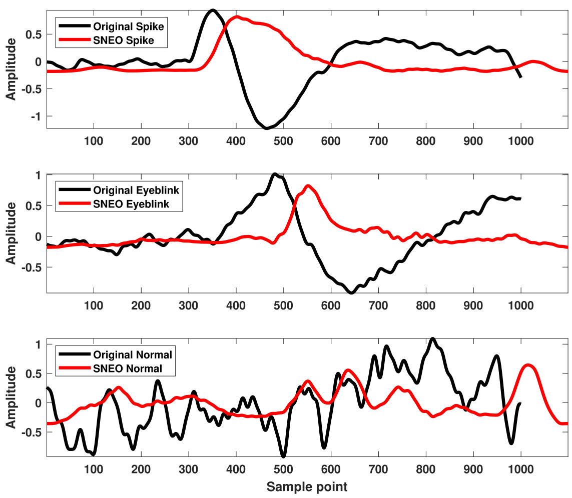


Fig. 3. Filtered EEGs by SNEO.

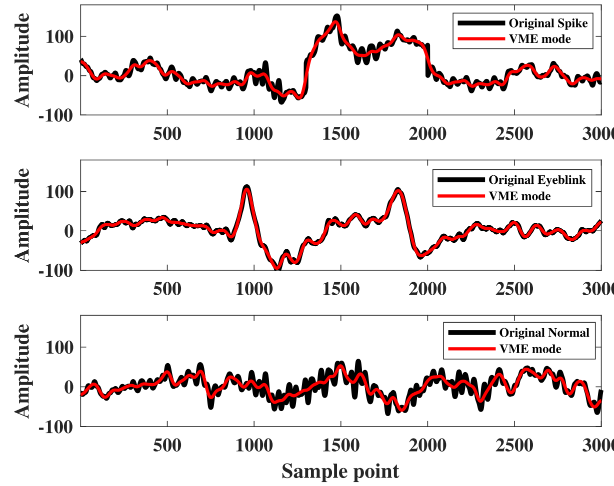


Fig. 4. Filtered EEGs by VME.

from epileptiform discharges very critical. In [35], we presented a primary study on epileptiform discharge screening for eye blink detection with unsupervised learning, where 8-dimensional features extracted from the slope sequence of filtered EEG by SNEO are adopted as the input patterns for EEG representation. *But it is found that the clustering method in [35] is less effective in handling EEGs containing frequently appeared epileptiform discharges in continuous signal segments.* The reason behind is that EEG is typical non-stationary random signal with plentiful local fluctuations, which seriously affects the characteristics of slope and skewness.

To address this issue, a hybrid filtering method using SNEO [37] and VME based signal transformation [30] is applied to EEG processing after 50 Hz notch filtering and 0.5-30 Hz band-pass filtering. It is noted that eye blink artifacts and epileptiform discharges normally jointly affect the FP1 and FP2 channels. Therefore, only these two channels are used for clustering. SNEO can highlight the high instantaneous energy at the high frequency of EEGs and is effective on spikes and sharp waves detection. The SNEO filtering is

$$S(i) = w(t) * (x(i) * x(i) - x(i-1) * x(i+1)) \qquad (1)$$

where $x(i)$ is the EEG, $w(t)$ represents the Barrett window function. Fig. 3 shows the filtered spike, eye blink artifact, and normal EEGs by SNEO, respectively. As observed, SNEO has a sensitive response at the position of epileptiform discharges when comparing with eye blink artifacts and normal EEGs. That is SNEO transformation can significantly enhance the smoothness of the original EEGs, making the epileptiform discharges and eye blinks more obvious and prominent.

Inherited from variational mode decomposition (VMD), an elegant VME algorithm with a simple version of VMD has been developed in [30]. VME can obtain the specific desired mode of interest of the signal by approximating the center frequency band. The basis of VME is to decompose the input signal $x(n)$ into the desired signal $u_d(n)$ and the residual $u_r(n)$ as $x(n) = u_d(n) + u_r(n)$. Setting the center frequency w, the approximate center frequency w_d can be obtained by iterative updation, and the expected signal u_d is iteratively approximated, with details as

$$\hat{u}_d^{m+1}(\omega) \leftarrow \frac{\hat{x}(\omega) + \alpha^2\left(\omega - \omega_d^m\right)^4 \hat{u}_d^m(\omega) + \frac{\hat{\lambda}(\omega)}{2}}{\left[1 + \alpha^2\left(\omega - \omega_d^m\right)^4\right]\left[1 + 2\alpha\left(\omega - \omega_d^m\right)^2\right]} \qquad (2)$$

$$\omega_d^{m+1} \leftarrow \frac{\sum_0^\infty \omega \left|\hat{u}_d^{m+1}(\omega)\right|^2}{\sum_0^\infty \left|\hat{u}_d^{m+1}(\omega)\right|^2} \qquad (3)$$

Here, α is the bandwidth of the adjustment filter, and λ is the Lagrangian multiplier, which iteratively optimized by

$$\hat{\lambda}^{m+1}(\omega) \leftarrow \hat{\lambda}^m(\omega) + \tau\left[\hat{x}(\omega) - \left(\hat{u}_d^{m+1}(\omega) + \hat{x}_r(\omega)\right)\right] \qquad (4)$$

The VME signal $V(i)$ can be obtained by

$$V(i+1) = \frac{x(i) + a^2 \cdot (i - (w_d^m)^4 \cdot V(i) + \lambda(i)/2}{([1 + a^2 \cdot (i - (w_d)^m)^4] \cdot [1 + 2 \cdot a \cdot (i - (w_d)^m)^2])} \qquad (5)$$

where a is the compactness coefficient, and $V(0)$ is set to be 0. For details of VME, one can refer to [30]. Fig. 4 shows the filtered signals by VME. As clearly observed, the local high-frequency fluctuations after VME transformation are significantly reduced. The smoothed signal is more significant for feature extraction and clustering.

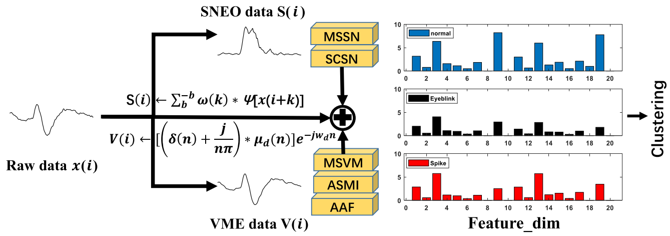


Fig. 5. The proposed unsupervised epileptiform discharge filtering.

TABLE III
THE EEG FEATURES FOR EYE BLINKS AND EPILEPTIFORM DISCHARGES CLUSTERING

Feature	Calculation	Signal
MSMI	$\max\{TS(n)\}$	EEG
SCMI	$count(TS(n){>}k)$	EEG
MSSN	$\max\{TS(n)\}$	SNEO
SCSN	$count(TS(n){>}k)$	SNEO
MSVM	$\max\{TS(n)\}$	VME
ASMI	$\frac{1}{N}\sum_{n=1}^{N} TS(n)$	VME
AAF	$\frac{1}{N-1}\sum_{j=1}^{(n-1)} \lvert V(j+1)-V(j)\rvert$	VME
MAF	$\max(\lvert V(j+1)-V(j)\rvert)$	VME
AAS	$\frac{1}{N-2}\sum_{j=1}^{(n-2)} \lvert V(j+2)-V(j)\rvert$	VME
MAS	$\max(\lvert V(j+2)-V(j)\rvert)$	VME

The steepness of EEGs, which has been shown a discriminative pattern of epileptiform discharges, is applied for eye blinks and epileptiform discharges clustering. Similar to [35], the monotonically increasing partial slope sequence $TS(n)$ is firstly obtained. Then, the monotonically increasing partial slope count STS of each signal exceeding a given threshold is calculated by

$$STS = count(TS(n) > k) \tag{6}$$

where k is a preset threshold. The operation is performed on the original EEG, the filtered EEG by SNEO and VME. Based on the partial slope sequence and the slope count, 20-dimensional features on the 2 frontal channels are extracted (FP1 and FP2, 10 features on each channel), where for the original EEG, the maximum value of the monotonically increasing slope (MSMI) and the slope count of monotonic increasing part (SCMI) are extracted, and for SNEO, these two similar features are also derived, in short as MSSN and SCSN, respectively. The effectiveness of these features has been verified in [35]. For VME, 6 features, including the maximum value of the monotonically increasing slope (MSVM), the average value of the monotonically increasing partial slopes (ASMI), the absolute average of the first-order difference (AAF), the absolute maximum of the first-order difference (MAF), the absolute average of the second-order difference (AAS), the absolute maximum of the second-order difference (MAS), are extracted. Table III shows the detailed calculations of all these features.

For illustration, Fig. 6 compares the distributions of 4 VME features obtained from EEGs with/without epileptiform discharges, where the estimated probability density distributions associated with the 4 features are also depicted for comparison. The distribution differences between epileptiform discharges and other EEGs are significant (Same observations can be found in the rest 2 VME features, MSVM and ASMI, the plots are omitted here due to page limitation). To perform epileptiform discharges filtering, the unsupervised K-means algorithm on the 20 features is performed, as shown in Fig. 5. Fig. 7 visualizes the clustering performance comparison between our algorithm and MDF. Comparing with our method, there has a clear overlap between the EEG categories with/without epileptiform discharges by MDF.

2) *Multi-Dimensional Feature Optimization:*

a) Improved VME (IVME) algorithm: After the above epileptiform discharge filtering, a novel multi-dimensional feature extraction and optimization algorithm is developed for EEG representation in eye blink artifact detection. First of all, the 16 discriminative features in [35] are also inherited for eye blink detection in this paper, where the details can be referred to [35]. Additionally, a new feature based on VME [30] with improvement on robust and adaptive thresholding is developed. VME can extract the approximate value of eye blinks from contaminated EEG, and is effective in locating eye blink position without any implementation calibration or artifact reference. VME first extracts the desired mode $m(n)$ from the original EEG, and then calculates the local maximum value of the desired mode to achieve eye blink detection by thresholding. The threshold is defined as [30]

$$\Theta = \frac{median(|m(n)|)}{0.6745}\sqrt{2\log N} \tag{7}$$

where N is the EEG segment length. Since eye blink generally lasts for 0.2 to 0.4 second, N is set be 0.5 second to cover the eye blink duration.

Although promising eye blink detection performance can be achieved in [30], we find that VME is less effective in dealing with data containing continuous multiple eye blink artifacts or having offset phenomenon in EEG segment. Fig.8 (a) shows the threshold derived by VME [30] on EEG with multiple continuous eye blinks. Due to the overall large amplitude, there have 3 eye blinks been miss detected. To alleviate this deficiency, we develop an improved VME (IVME) algorithm for robust threshold calculation. Instead of using the original VME filtered EEG, the EEG difference sequence within a given interval is first obtained. Then, the improved threshold is derived by

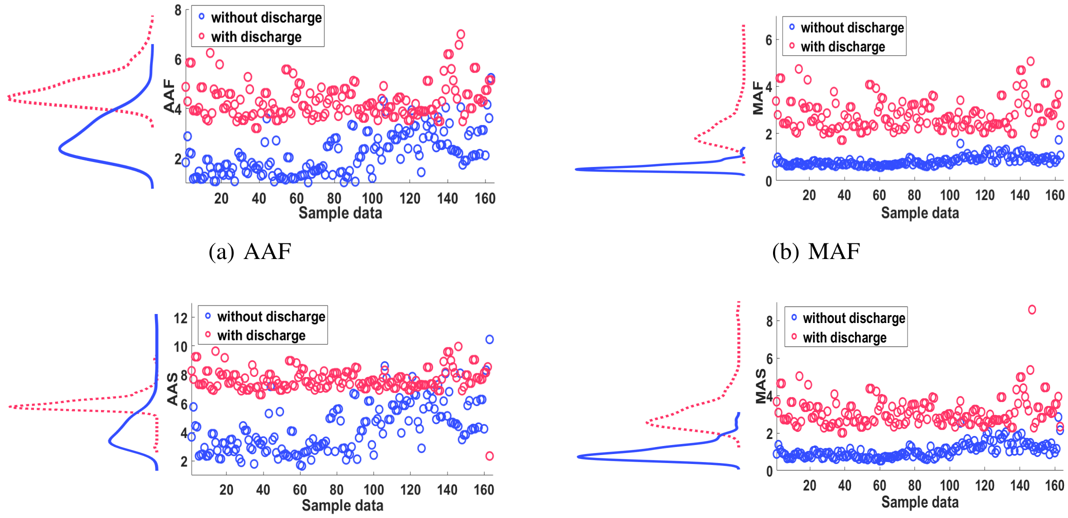
$$\Theta = \frac{median(|m(n+\Delta n) - m(n)|)}{0.6745}\sqrt{2\log N} \tag{8}$$

where Δn denotes the interval, which we calculate it by

$$\Delta n = T \cdot t_1 \cdot r \tag{9}$$

Here, T is the sampling frequency, t_1 represents the eye blink duration (we set it to be 0.3 s in this paper), and r is the ratio between the duration of the eye blink staring point to the peak with respect to the full duration of a eye blink. According to the property of eye blink wave, r is generally set to be 1/4.

Improved VME can effectively address the miss detection issue in VME when multiple consecutive eye blinks are



(a) AAF (b) MAF

(c) AAS (d) MAS

Fig. 6. Distributions of VME features: (a) AAF, (b) MAF, (c) AAS, (d) MAS.

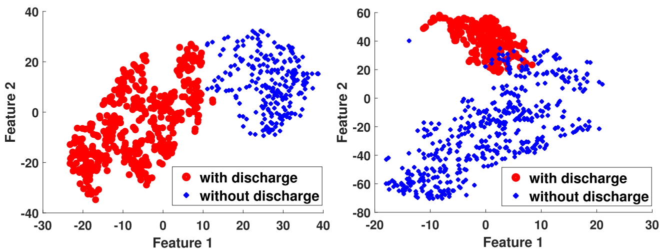
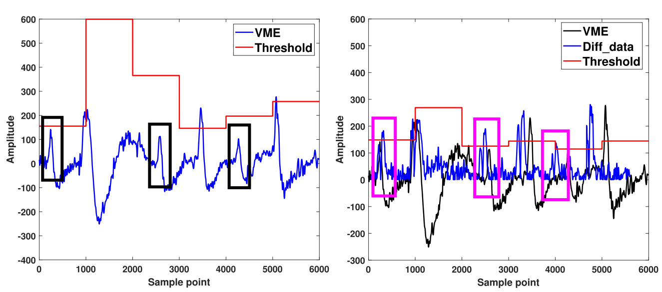


Fig. 7. Clustering comparison between our proposed algorithm(left) and MDF [35](right).



(a) VME (b) Proposed IVME

Fig. 8. (a) Threshold for eye blink detection obtained by VME [30], it can be seen that some of the eye blink artifacts with low amplitude are missed, as marked in the black boxes. (b) The blue line is the difference signal. It can be seen that the IVME threshold is more robust, that can detect all eye blinks, especially for those missed signals by VME, as shown in the pink box in (b).

presented. As shown in Fig. 8 (b), all eye blinks missed by VME can be correctly detected by improved VME. The EEG difference signal can reflect the amplitude change of the offset signal, thus reduce the high-amplitude phenomenon of the non-blinking EEG. Moreover, for consecutive eye blinks, the difference signals not only can retain high amplitude at the eye blinks, but also remain low amplitude at other EEGs. To well characterize EEGs in eye blink detection, the threshold (TRS)

Algorithm 1 IVME Threshold Calculation

Input: EEG $x(i)$, a, w_d, k, Δn, N
Output: IVME based TRS

index $count = 0$
$y(n) \leftarrow VME(x(i), a, w_d)$
$m(n) = |y(n + \Delta n) - y(n)|$
Calculate the VME threshold θ by

$$\theta = \frac{median(m(n))}{0.6745}\sqrt{2 \log N}$$

for $i = 1$ to N **do**
  if $x(i + 1) > \theta$ **then**
    $count = count + 1$
  end if
end for
TRS = $count$

by improved VME is extracted as a pattern, where TRS calculation is summarized in Algorithm 1.

b) Feature optimization by variance filtering: Combining the improved VME threshold with the 16 discriminative features in [35] for eye blink representation, the variance filtering is further applied to feature optimization in this section. Comparing with other feature selection, such as Relief, variance filtering is confirmed effective in eye blink EEG feature optimization. The details of variance filtering on eye blink feature optimization can be referred to [35]. It is noted that for each subject, the finally selected optimal features will be different due to the individual discrepancy. Fig. 9 visualizes the variance filtering based EEG feature optimization. Meanwhile, our proposed eye blink detection algorithm is briefly summarized in Algorithm 2. Particularly, SVM with the radial basis function (RBF) kernel is used as the classifier for eye blink detection in this paper.

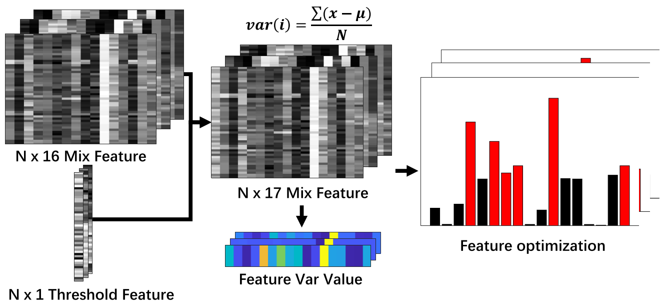


Fig. 9. Variance filtering based feature optimization.

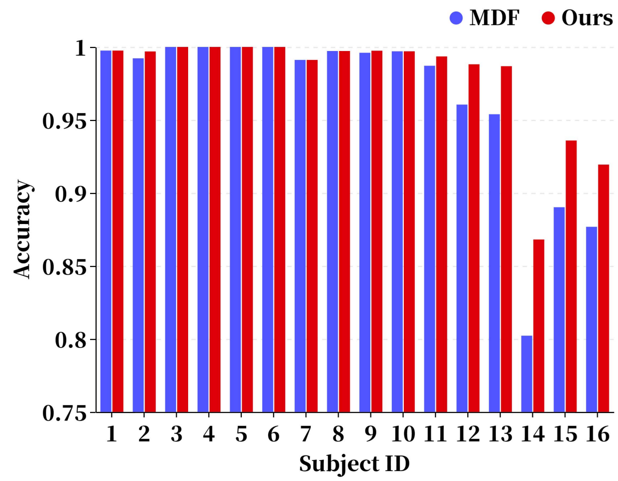


Fig. 10. Epileptiform discharge and eye blink clustering performance comparison between our methods and MDF [35].

Algorithm 2 The Proposed Eye Blink Detection Algorithm

Input: EEG $x(i)$, a, w_d, k, Δn, N, $label$

Output: Detection results by sensitivity, specificity, accuracy

1: Obtain $S(i)$ by (1) and $V(i)$ by algorithm in [30]
2: Calculate features in Table III
3: Unsupervised clustering based epileptiform discharge filtering
4: Obtain the threshold feature by IVME in Algorithm 1
5: Calculate the multi-dimensional EEG features in [35]
6: Perform feature selection by variance filtering algorithm
7: Machine learning based eye blink detector training

C. Performance Evaluation

To test the performance, the CHZU eye blink artifact database is used, where the training and testing data are randomly selected with the ratio of 7:3. For performance evaluation, the sensitivity, specificity and accuracy are calculated, which are defined as

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \tag{10}$$

$$\text{Sensitivity} = \frac{TP}{TP + FN},\ \text{Specificity} = \frac{TN}{TN + FP} \tag{11}$$

Here, TP, FP, TN, FN represent true positive, false positive, true negative, and false negative, respectively. For SVM in the our method, the regularization parameter C and sensitivity coefficient are optimized within a given region by the cross validation. Meanwhile, the detection performance is also compared with 5 state-of-the-art algorithms: 1) MDF in [35], 2) VME algorithm in [30], 3) ICA-SVM algorithm in [26], 4) WICA-SVM algorithm in [27], and 5) SVM-AE algorithm in [28]. All compared algorithms are carried out in the paper by the following settings.

For MDF, the thresholds in feature extraction are same to [35], and SVM with the RBF kernel is applied as the detector, where the parameter optimization is same to the proposed algorithm. For VME, same optimization to [30] is used. In ICA-SVM [26], the second-order statistics are adopted to separate the EEGs, and 4 representative features, 1) the ratio of peak value to variance, 2) the absolute value of Skewness, 3) the maximum value of the average cross-correlation between the source component and the channels (FP1, FP2, F3, F4, O1, O2), 4) the statistical distance between the source component and the PDF of EEGs in the channel known to contain eye blinks, are extracted for SVM detector training. The kernel parameters of SVM after optimization are $C = 72$, $\sigma = 7$. Wavelet multiresolution analysis (WMA) is employed in WICA-SVM to perform preliminary EEG filtering, as such the wavelet details corresponding to the frequency range of interests can be retained [27]. The EEG variance, Kurtosis, Shannon entropy, and amplitude range after ICA are taken to train SVM for eye blink detection. The 'db8' is used as the mother wavelet and a linear kernel is used in SVM, with the optimal $C = 10$. SVM-AE [28] first applies SVM to extract variance, Kurtosis, and peak-to-peak amplitude characteristics for EEG classification, and then uses a pre-trained AE to perform eye blink filtering.

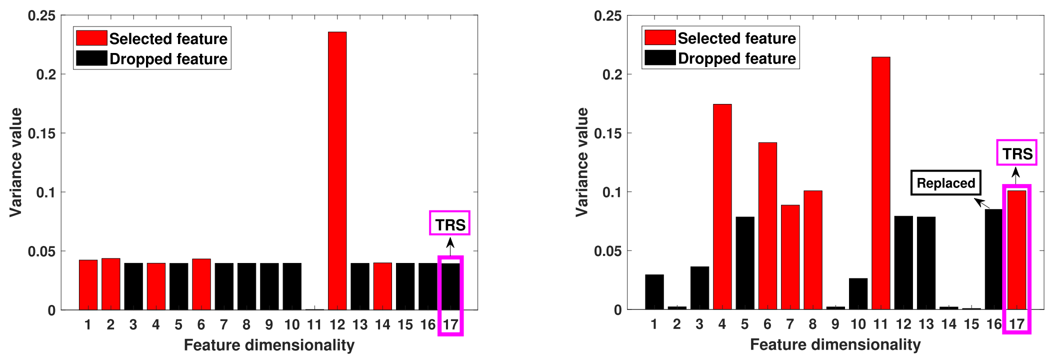
III. Results and Discussions

1) Epileptiform Discharge Filtering Performance: The first experiment shows the effectiveness of adding VME features in epileptiform discharge filtering by comparing with the clustering method in [35]. Fig. 10 plots the clustering accuracy on each subject. As observed, i) for subjects 2, 7, 9, 11∼16, obvious accuracy improvement can be achieved by our method. The enhancement is more significant on subjects 11∼16 as there exist many epileptiform discharges in these subjects, as shown in Table II. For all subjects, an average of 1.42% accuracy increment can be offered by our method over MDF. ii) Particularly, for subjects 14∼16, the average accuracy increment by our method is high to 5.14%. iii) For the rest subjects, the performance of our method is the same to or slight better than MDF as no epileptiform discharge samples existed in these data. VME filtering can smooth the original EEG, making it more suitable for filtering consecutive epileptiform discharges with high frequency fluctuations.

2) Performance on IVME Threshold Feature: Detection performance comparisons between our method by adding the IVME threshold feature with respect to MDF [35] and VME [30] are presented in Table IV. The detailed results on each subject and the average result on all subjects are obtained.

TABLE IV
PERFORMANCE COMPARISON WITH RESPECT TO VME AND MDF METHODS

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	S13	S14	S15	S16	**Average**
Sensitivity																	
VME [30]	68.49	49.22	63.25	73.82	69.26	79.6	89.56	58.55	71.06	75.79	72.27	75.21	72.06	71.62	72.53	69.63	70.75
MDF [35]	**97.33**	84.85	**100**	93.33	**97.22**	**95.9**	95.65	93.55	**94.44**	86.55	97.06	95.24	95.36	88.89	94.46	92.95	93.92
Ours	**97.33**	**86.42**	**100**	**95**	**97.22**	**95.9**	**96.22**	**93.98**	**94.44**	**88.93**	**98.08**	**96.53**	**97.51**	**93.16**	**95.89**	**94.09**	**95.04**
Specificity																	
VME [30]	78.26	63.49	82.88	78.3	82.08	72.22	60.98	63.89	71.7	49.23	57.41	35.71	43.86	48.72	30.65	50.68	60.63
MDF [35]	**100**	**89.22**	**100**	**95.56**	**100**	**85.42**	95.65	89.69	**98.7**	**94.59**	**96.88**	76.47	77.27	62.16	64.46	69.61	87.23
Ours	**100**	88.76	**100**	95.18	**100**	**85.42**	**95.74**	**90.37**	**98.7**	93.65	96.08	**83.75**	**86.41**	**66.15**	**75.86**	**76.19**	**89.52**
Accuracy																	
VME [30]	70.72	52.03	66.83	74.89	73	78.66	86.09	59.08	71.15	73.2	69.68	66.4	66.78	67.69	66.59	66.35	69.32
MDF [35]	**98.37**	87.5	**100**	94	**98.26**	**92.94**	95.65	91.86	**95.82**	87.98	96.67	88.93	88.82	78.02	83.07	82.94	91.30
Ours	**98.37**	**87.66**	**100**	**95.07**	**98.26**	**92.94**	**96.15**	**92.37**	**95.82**	**89.82**	**97.42**	**92.49**	**93.75**	**81.76**	**89.24**	**86.97**	**93.01**
Time (s)																	
VME [30]	7.4	11.7	11.2	8.3	6.7	10.2	6.2	13.4	14	12.2	5.7	4.6	5.6	8.3	8.1	7.7	8.83
MDF [35]	71.2	103.7	28.9	1.6	20.7	24	16.2	13.1	136.9	114.9	2.9	32.9	2.6	55.3	11.3	15.6	48.05
Ours	71.2	103.7	28.9	1.6	20.7	24	16.2	130.1	136.9	114.9	2.9	32.9	2.6	55.3	11.3	15.6	48.05



(a) Subject 1 (b) Subject 4

Fig. 11. (a) is the variance values of all features for subject 1, where the TRS feature is not selected due to the low variance value, and (b) variance values of all features for subject 4, where the TRS feature is selected for EEG representation.

The best result is highlighted using the bold font in the table. It is observed that: i) for all subjects, VME generally suffers poor detection performance with a relative low sensitivity, specificity and accuracy. The overall average accuracy by VME is only 69.32%. This is mainly because VME is not specifically designed for eye blink detection with EEGs containing epileptiform discharges. When encountering continuous eye blinks, data drifting, large overall signal amplitude or epileptiform discharges, the resulting threshold by VME may not be sensitive to eye blink artifacts. A large number of eye blink samples cannot be detected, resulting a poor performance. ii) Comparing with MDF, adding the threshold feature by the developed IVME into the proposed algorithm can enhance the detection performance for subjects 2, 4, 7, 8, 10∼16, where the improvement is obviously significant on the last 5 subjects. For the rest subjects, our method performs exactly the same to MDF. iii) Comparing with VME and MDF, our method achieves the highest average accuracy, specificity and sensitivity values. Particularly, the accuracy increment over VME is high to 23.69%. Similar enhancement on specificity and sensitivity can be achieved by our method over VME.

Due to individual differences, in the variance feature selection stage, the improved VME threshold feature may not be the optimal one been selected. For these subjects, that is the reason why the detection performance by our method is the same to MDF. Comparing with MDF, for subjects whose EEG containing plentiful epileptiform discharges and consecutive eye blinks, adding the improved threshold feature by IVME can effectively enhance the detection performance. Fig. 11 compares the variance value of each feature in variance filtering on subject 1 and subject 4. The 17-th feature is the threshold (TRS) by improved VME. As observed, for subject 1, TRS is not selected as the optimal feature for EEG representation, which explains the same performance between our method and MDF. While for subject 4, as there

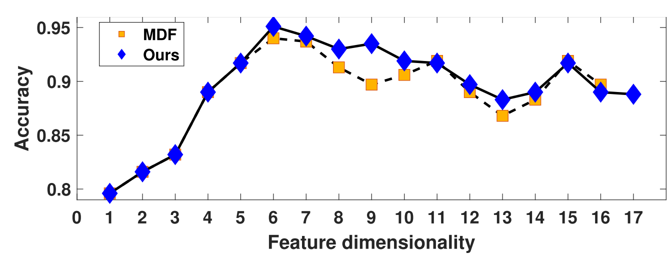
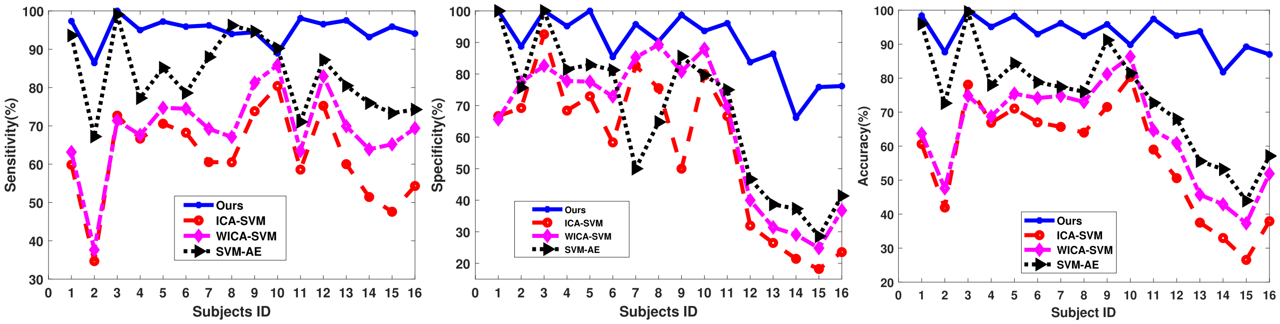


Fig. 12. Accuracy on variance filtering based feature selection between MDF [35] and our method on subject 4.



(a) Sensitivity (b) Specificity (c) Accuracy

Fig. 13. Detection performance comparison with respect to ICA-SVM [26], WICA-SVM [27] and SVM-AE [28].

contains epileptiform discharges and consecutive eye blinks in EEGs, TRS is selected as one of the optimal features, which can help to improve the detection performance. In particular, we show the feature selection step in variance filtering and the corresponding detection accuracy obtained by MDF and our method on the EEG data of subject 4 in Fig. 12. As clearly shown, the same first 5 features are selected by MDF and our method, resulting the same performance. When the 6-th feature is selected, both two methods reach the best performance, but MDF is worse than our method. The reason behind is that TRS is selected by our method as the 6-th feature.

3) Comparison With State-of-the-Art Methods: The proposed eye blink detection performance is also compared with 3 state-of-the-art (SOTA) methods, ICA-SVM [26], WICA-SVM [27] and SVM-AE [28] in this section. The sensitivity, specificity, and accuracy obtained on each subject are shown in Fig. 13. As depicted, for most of the subjects, our method performs overwhelmingly better than the 3 SOTA methods. Apparently, the improvement is more obvious on these subjects containing epileptiform discharges. For some subjects, the detection accuracy and specificity by the 3 SOTA methods are very poor, even less than 50%. The main reason of the poor performance is that all these 3 methods are not specifically designed for handling eye blink detection in EEGs containing rich epileptiform discharge waves. But our method not only can achieve a promising performance on eye blink detection in general EEG data, but is also robust to affections caused by epileptiform discharges.

IV. Conclusion

The paper presents a novel eye blink artifact detection algorithm that fully studies the affection of epileptiform discharges in EEGs. The clustering method by employing the improved VME filter based discriminate EEG features can effectively screen out epileptiform discharges. The improved VME threshold is capable in dealing with EEGs containing multiple consecutive eye blinks and large data drifting. Combining with MDF and variance filtering based multi-dimensional feature optimization, a robust eye blink detection method that can address subject EEG difference is developed. The most significant advantage of the proposed method is that it can efficiently filter out eye blinks existed in EEGs of complex neurological diseases, and has better applicability and portability to different subjects. Future work will focus on analyzing effective algorithms on more types of artifacts detection in EEGs of neurological diseases.

References

[1] I. Kaya, "A brief summary of EEG artifact handling," 2020, *arXiv:2001.00693*.

[2] M. Ogino, S. Kanoga, M. Muto, and Y. Mitsukura, "Analysis of prefrontal single-channel EEG data for portable auditory ERP-based brain–computer interfaces," *Frontiers Hum. Neurosci.*, vol. 13, p. 250, Jul. 2019.

[3] Z. Deng, P. Xu, L. Xie, K.-S. Choi, and S. Wang, "Transductive joint-knowledge-transfer TSK FS for recognition of epileptic EEG signals," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 8, pp. 1481–1494, Aug. 2018.

[4] J. Cao, J. Zhu, W. Hu, and A. Kummert, "Epileptic signal classification with deep EEG features by stacked CNNs," *IEEE Trans. Cognit. Develop. Syst.*, vol. 12, no. 4, pp. 709–722, Dec. 2020.

[5] J. Cao, D. Hu, Y. Wang, J. Wang, and B. Lei, "Epileptic classification with deep transfer learning based feature fusion algorithm," *IEEE Trans. Cognit. Develop. Syst.*, early access, Mar. 8, 2021, doi: 10.1109/TCDS.2021.3064228.

[6] D. Hu, J. Cao, X. Lai, J. Liu, S. Wang, and Y. Ding, "Epileptic signal classification based on synthetic minority oversampling and blending algorithm," *IEEE Trans. Cognit. Develop. Syst.*, vol. 13, no. 2, pp. 368–382, Jun. 2021.

[7] Z. Xu, T. Wang, J. Cao, Z. Bao, T. Jiang, and F. Gao, "BECT spike detection based on novel EEG sequence features and LSTM algorithms," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 1734–1743, 2021.

[8] G. Bartels, L.-C. Shi, and B.-L. Lu, "Automatic artifact removal from EEG—A mixed approach based on double blind source separation and support vector machine," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol.*, Aug. 2010, pp. 5383–5386.

[9] W.-D. Chang *et al.*, "Detection of eye blink artifacts from single prefrontal channel electroencephalogram," *Comput. Methods Programs Biomed.*, vol. 124, pp. 19–30, Feb. 2016.

[10] Q. Yang *et al.*, "Childhood epilepsy syndromes classification based on fused features of electroencephalogram and electrocardiogram," *Cognit. Comput. Syst.*, vol. 4, no. 1, pp. 1–10, Mar. 2022.

[11] A. K. Maddirala and K. C. Veluvolu, "Eye-blink artifact removal from single channel EEG with K-means and SSA," *Sci. Rep.*, vol. 11, no. 1, pp. 1–14, Dec. 2021.

[12] 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.

[13] M. Izzetoglu, A. Devaraj, S. Bunce, and B. Onaral, "Motion artifact cancellation in NIR spectroscopy using Wiener filtering," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 5, pp. 934–938, May 2005.

[14] R. Foodeh, A. Khorasani, V. Shalchyan, and M. R. Daliri, "Minimum noise estimate filter: A novel automated artifacts removal method for field potentials," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 8, pp. 1143–1152, Aug. 2017.

[15] M. Chavez, F. Grosselin, A. Bussalb, F. D. V. Fallani, and X. Navarro-Sune, "Surrogate-based artifact removal from single-channel EEG," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 3, pp. 540–550, Mar. 2018.

[16] M. Shahbakhti, M. Maugeon, M. Beiramvand, and V. Marozas, "Low complexity automatic stationary wavelet transform for elimination of eye blinks from EEG," *Brain Sci.*, vol. 9, no. 12, p. 352, Dec. 2019.

[17] C. J. James and C. W. Hesse, "Independent component analysis for biomedical signals," *Physiol. Meas.*, vol. 26, no. 1, pp. R15–R39, Feb. 2005.

[18] M. Pal, R. Roy, J. Basu, and M. S. Bepari, "Blind source separation: A review and analysis," in *Proc. Int. Conf. Oriental COCOSDA Held Jointly Conf. Asian Spoken Lang. Resea. Eval.*, 2013, pp. 1–5.

[19] P. Berg and M. Scherg, "Dipole modelling of eye activity and its application to the removal of eye artefacts from the EEG and MEG," *Clin. Phys. Physiol. Meas.*, vol. 12, pp. 49–54, Jan. 1991.

[20] R. N. Vigário, "Extraction of ocular artefacts from EEG using independent component analysis," *Electroencephalogr. Clin. Neurophysiol.*, vol. 103, no. 3, pp. 395–404, 1997.

[21] S. Romero, M. Mananas, S. Clos, S. Gimenez, and M. J. Barbanoj, "Reduction of EEG artifacts by ICA in different sleep stages," in *Proc. 25th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, vol. 3, Sep. 2003, pp. 2675–2678.

[22] R. Vigario, J. Särelä, V. Jousmiki, M. Hämäläinen, and E. Oja, "Independent component approach to the analysis of EEG and MEG recordings," *IEEE Trans. Biomed. Eng.*, vol. 47, no. 5, pp. 589–593, May 2000.

[23] 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. 2015.

[24] J. Fairley, G. Georgoulas, C. Stylios, and D. Rye, "A hybrid approach for artifact detection in eeg data," in *Proc. Int. Conf. Artif. Neural Netw.* Berlin, Germany: Springer-Verlag, 2010, pp. 436–441.

[25] S. Phadikar, N. Sinha, and R. Ghosh, "Automatic eyeblink artifact removal from EEG signal using wavelet transform with heuristically optimized threshold," *IEEE J. Biomed. Health Informat.*, vol. 25, no. 2, pp. 475–484, Feb. 2021.

[26] L. Shoker, S. Sanei, and J. Chambers, "Artifact removal from electroencephalograms using a hybrid BSS-SVM algorithm," *IEEE Signal Process. Lett.*, vol. 12, no. 10, pp. 721–724, Oct. 2005.

[27] 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 Inform.*, vol. 22, no. 3, pp. 664–670, May 2018.

[28] R. Ghosh, N. Sinha, and S. K. Biswas, "Automated eye blink artefact removal from EEG using support vector machine and autoencoder," *IET Signal Process.*, vol. 13, no. 2, pp. 141–148, Apr. 2019.

[29] B. Yang, K. Duan, and T. Zhang, "Removal of EOG artifacts from EEG using a cascade of sparse autoencoder and recursive least squares adaptive filter," *Neurocomputing*, vol. 214, pp. 1053–1060, Nov. 2016.

[30] M. Shahbakhti *et al.*, "VME-DWT: An efficient algorithm for detection and elimination of eye blink from short segments of single EEG channel," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 408–417, 2021.

[31] S. Narejo, E. Pasero, and F. Kulsoom, "Eeg based eye state classification using deep belief network and stacked autoencoder," *Int. J. Electr. Comput. Eng.*, vol. 6, no. 6, pp. 3131–3141, 2016.

[32] J. Cao *et al.*, "Unsupervised eye blink artifact detection from EEG with Gaussian mixture model," *IEEE J. Biomed. Health Informat.*, vol. 25, no. 8, pp. 2895–2905, Aug. 2021.

[33] M. Shahbakhti *et al.*, "Simultaneous eye blink characterization and elimination from low-channel prefrontal EEG signals enhances driver drowsiness detection," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 3, pp. 1001–1012, Mar. 2022.

[34] J. T. Valderrama, A. de la Torre, and B. Van Dun, "An automatic algorithm for blink-artifact suppression based on iterative template matching: Application to single channel recording of cortical auditory evoked potentials," *J. Neural Eng.*, vol. 15, no. 1, Feb. 2018, Art. no. 016008.

[35] J. Wang, J. Cao, D. Hu, T. Jiang, and F. Gao, "Eye blink artifact detection with novel optimized multi-dimensional electroencephalogram features," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 1494–1503, 2021.

[36] X. Cui *et al.*, "Deep feature fusion based childhood epilepsy syndrome classification from electroencephalogram," *Neural Netw.*, vol. 150, pp. 313–325, Jun. 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0893608022000922

[37] H. K. Garg and A. K. Kohli, "Nonstationary-epileptic-spike detection algorithm in EEG signal using SNEO," *Biomed. Eng. Lett.*, vol. 3, no. 2, pp. 80–86, Jun. 2013.