

EEG Reconstruction With a Dual-Scale CNN-LSTM Model for Deep Artifact Removal

Tengfei Gao, Dan Chen^{ID}, Yunbo Tang, Zhekai Ming, and Xiaoli Li^{ID}

Abstract—Artifact removal has been an open critical issue for decades in tasks centering on EEG analysis. Recent deep learning methods mark a leap forward from the conventional signal processing routines; however, those in general still suffer from insufficient capabilities 1) to capture potential temporal dependencies embedded in EEG and 2) to adapt to scenarios without a priori knowledge of artifacts. This study proposes an approach (namely *DuoCL*) to deep artifact removal with a dual-scale CNN (Convolutional Neural Network)-LSTM (Long Short-Term Memory) model, operating on the raw EEG in three phases: 1) *Morphological Feature Extraction*, a dual-branch CNN utilizes convolution kernels of two different scales to learn morphological features (individual sample); 2) *Feature Reinforcement*, the dual-scale features are then reinforced with temporal dependencies (inter-sample) captured by LSTM; and 3) *EEG Reconstruction*, the resulting feature vectors are finally aggregated to reconstruct the artifact-free EEG via a terminal fully connected layer. Extensive experiments have been performed to compare *DuoCL* to six state-of-the-art counterparts (e.g., 1D-ResCNN and NovelCNN). *DuoCL* can reconstruct more accurate waveforms and achieve the highest SNR & correlation (CC) as well as the lowest error (RRMSE_t & RRMSE_f). In particular, *DuoCL* holds potentials in providing a high-quality removal of unknown and hybrid artifacts.

Index Terms—Electroencephalogram (EEG), artifact removal, CNN, LSTM, end-to-end.

I. INTRODUCTION

LECTROENCEPHALOGRAPHY (EEG) is the most important non-invasive recording of brain electrical activities

Manuscript received 23 March 2022; revised 30 September 2022; accepted 2 December 2022. Date of publication 7 December 2022; date of current version 7 March 2023. This work was supported in part by Scientific and Technological Innovation 2030 under Grant 2021ZD0204300, in part by the National Natural Science Foundation of China, under Grants 62172304 and 61977027, in part by Science & Technology Major Project of Hubei Province, Next-Generation AI Technologies, under Grant 2021BEA159, and in part by the National Defence Basic Scientific Research Program of China. (*Corresponding author: Dan Chen*.)

Tengfei Gao, Dan Chen, and Zhekai Ming are with the National Engineering Research Center for Multimedia Software, School of Computer Science and Hubei Key Laboratory of Multimedia and Network Communication Engineering, Wuhan University, Wuhan 430072, China (e-mail: gaotengfei@whu.edu.cn; dan.chen@whu.edu.cn; zhekaiming@whu.edu.cn).

Yunbo Tang is with the College of Computer and Data Science, Fuzhou University, Fuzhou 350108, China (e-mail: yunbotang@whu.edu.cn).

Xiaoli Li is with the National Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing 100875, China (e-mail: xiaoli@bnu.edu.cn).

Digital Object Identifier 10.1109/JBHI.2022.3227320

in neuroscience research and neuro-engineering practice. Raw EEG is inevitably susceptible to intensive interference during recording from a variety of 1) electrophysiological activities of other sources *in vivo*, such as electrooculogram (EOG), electromyogram (EMG), electrocardiogram (ECG), and sweating, and 2) non-physiological signals *in vitro* from the acquisition equipment or the environment [1]. Such uncertain artifacts and noises impose severe biases on the characteristics of the genuine EEG, which often fails subsequent analyses. For decades, artifact removal has been an open critical issue in EEG-based tasks.

Artifact identification is challenging because: 1) artifacts are everywhere and 2) they are similar to artifact-free signals (“real EEG”) in many ways. Artifacts present a wide amplitude range and spectral distribution, perturbing all EEG waveforms and frequency bands of interest. Therefore, visual inspection is nearly impossible. Artifacts can even exhibit rhythmic properties that may be misinterpreted as important EEG characteristics, such as seizures. Differentiating the non-stationarity of artifacts from that of real (genuine, pure, or clean) EEG has long been placed at the core of EEG preprocessing [2].

The rejection of contaminated EEG portions appears to be a natural solution. However, it is subject to the risk of intolerable information loss, and the introduction of abrupt discontinuities will disrupt the temporal course of brain activities under examination [3]. Furthermore, it is critical to maintain the nonlinear characteristics of EEG recordings of the dynamics of the most complex chaotic system in the universe (the human brain) while removing artifacts. The principle here is to single out and preserve as much information about the underlying neural activities as possible. In other words, an artifact removal approach must identify artifacts and eliminate them under the ideal condition of leaving artifact-free recordings intact [4].

Traditional signal processing methods for artifact removal mainly fall into two categories: 1) straightforward regression and filtering (and similar) of the raw EEG and 2) decomposition and transformation: mapping the raw EEG into a new data space where artifact sources may be separated relatively easily, typically including empirical mode decomposition (EMD), wavelet transform (WT), blind source separation (BSS), and their variants [5], [6], [7]. All the aforementioned methods generally suffer from excessive residual artifacts and indefinite losses of useful information.

Furthermore, regression and filtering methods rely heavily on the available reference channels, limiting their applicability to certain types of artifacts [8] (the acquisition of signals from the applicable reference channel is also not trivial [9]). For the decomposition-based methods, permutation and selection of artifact components remain an open issue to minimize information loss and EEG degradation [10]. To remove artifacts that significantly overlap with real EEG in the frequency spectrum,

e.g., EMG, the aforementioned methods cannot be applied due to the difficulties of finding reference channels and distinguishing excessive artifact components [11].

Recently, booming methods based on deep learning have marked a leap forward from conventional signal processing routines, particularly in terms of generalization, benefiting from the ability to derive hierarchical and abstract representations from raw EEG [12]. However, existing deep learning methods, in general, still suffer from insufficient capabilities for capturing potential temporal dependencies embedded in EEG and adapting to scenarios without a priori knowledge of artifacts:

- 1) EEG may only be well characterized by joining a diversity of sophisticated features (e.g., duration, slope, waveform, and rhythm & non-rhythm) together to adapt to the morphological characteristics [13]. An end-to-end neural network model derives EEG features at a fixed scale of the data space, which apparently has limitations in describing the morphological characteristics.
- 2) The evolution of brain activities cannot be abrupt and exhibits temporal dependencies in different/continuous stages [14]. There is physiological regularity between the real EEG (in the form of a time series) and the underlying brain activities, not the origins of artifacts. Such temporal dependencies have not been properly addressed in existing studies.

There is still an urgent need for a “deeper” artifact removal method to meet the rapidly increasing data quality requirement in neuroscience/neuro-engineering applications such as online monitoring and neuro-feedback tasks [15], [16].

With features that better represent the morphological characteristics, more opportunities are available to track the real EEG. Artifacts cannot evolve in the same manner as non-artifacts, and their temporal dependencies should exhibit a significant distinction. It is plausible to single out artifacts by referencing morphological characteristics and temporal dependencies, that is, using these nonlinear features to jointly reconstruct the real EEG. Therefore, this study leverages deep neural networks to construct morphological features of different scales while autonomously learning the temporal dependencies from raw EEG.

Convolutional neural networks (CNN) excel in adaptive feature learning and extraction of distortion-invariant patterns. Long Short-Term Memory (LSTM) is also well-established for its ability to find the latent temporal dependencies in sequenced data. Given the above, this study designs a dual-scale CNN-LSTM model (namely, *DuoCL*) to support deep artifact removal (Section III) in an end-to-end manner, operating on the raw EEG in three phases as follows:

- 1) *Morphological Feature Extraction*. *DuoCL* employs a dual-branch network with different sizes of convolution kernel to extract dual-scale deep features. The number of convolution kernels of each layer is adjusted to determine the appropriate dimension of the extracted features (*each individual sample*) to optimally approximate the morphological characteristics of the EEG.
- 2) *Feature Reinforcement*. In each branch, LSTM then measures the temporal dependencies *between individual samples*, and the results are applied to reinforce the dual-scale

features highlighting more detailed differences between artifacts and real EEG.

- 3) *EEG Reconstruction*. The resulting hybrid features containing morphological characteristics (at both the sample and inter-sample levels) form the basis of the reconstruction of real EEG. Subsequently, *DuoCL* aggregates the feature vectors from both branches and reconstructs the “desired” artifact-free EEG via a terminal full-connected layer.

DuoCL is trained in a supervised fashion by minimizing the mean square error (MSE) between the reconstructed EEG and pure EEG (pre-set). After completing the training, *DuoCL* can eliminate the artifacts from raw input EEG and reconstruct the artifact-free EEG by passing through the aforementioned three phases.

Three semi-simulated EEG datasets have been constructed for training and evaluation of *DuoCL* (Section IV-B). The datasets contain EEG segments contaminated by five main types of artifacts (EOG or ocular activities, EMG or muscular activities, ECG or cardiac activities, electrode motion (EM), and baseline wander (BW)) and the corresponding pure EEG segments. Extensive experiments have been performed as follows:

- 1) A series of ablation experiments examines the effectiveness of the design (Section IV-C);
- 2) The performance of the *DuoCL* against that of its mainstream counterparts has been examined including 1) subjective evaluation, 2) objective evaluation, 3) the capability of coding nonlinear characteristics (Section IV-D1), and 4) removing unknown and hybrid artifacts (Section IV-D2).
- 3) The effectiveness of *DuoCL* in handling multi-channel EEG has been examined with the limitation explored (Sections IV-D3 and IV-E).

The main contributions of this study are as follows:

- 1) This study proposes a once-off approach to deep removal of various EEG artifacts, including unknown/hybrid artifacts, which supports automatic EEG reconstruction applicable to different scenarios. The proposed approach significantly outperformed its mainstream counterparts;
- 2) The proposed *DuoCL* model effectively represents the morphological characteristics of EEG. The model enables this by maintaining the EEG’s nonlinear features and highlighting temporal dependencies.

II. RELATED WORK

The removal of EEG artifacts has been an active research topic for several decades. Existing studies mainly consist of methods based on 1) traditional signal processing/machine learning and 2) recently emerging methods based on deep learning.

Classical filters, such as band-pass filtering, may only apply when the frequency bands of the artifact and the EEG of interest do not substantially overlap [17]. Subsequently, advanced alternatives such as adaptive filtering and Kalman filtering are developed. Bartoil et al. proposed using the Kalman filter to reduce EMG artifacts from EEG with the assumption that EMG artifacts conform to a Poisson distribution [18].

In regression-based approaches, the functions for fitting and smoothing the data are constructed to describe different sources.

The amplitude relation between the reference channel and EEG channel is used to estimate and subtract artifacts [19]. For example, Schlögl et al. proposed and validated a regression-based method for reducing EOG artifacts and the results suggest that the method can reduce EOG artifacts by 80% [20].

For another category of the traditional approach based on decomposition and transformation, a variety of BSS techniques (such as independent component analysis, principal component analysis, and canonical correlation analysis), WT, EMD, and their variants are applied to decompose raw EEG into multiple components. For example, Phadikar et al. propose an automatic eyeblink artifacts removal method from corrupted-EEG signals using discrete wavelet transform (DWT). The method is superior to the recently reported methods in terms of average correlation coefficient (*CC*) [21]. In addition, there is a series of hybrid methods to improve the efficiency of artifact removal, of which WT and ICA are the classic and efficient combinations [22], [23].

Owing to the powerful learning and generalization abilities of neural networks, DL-based methods have been booming. Yang proposes a DL networks EEG denoising method, which utilized powerful capabilities to obtain the structural features for eliminating EOG artifacts from EEG [24]. Sun et al. propose a 1D-ResCNN model to remove multiple types of artifacts [25]. This is the first application of CNN in EEG denoising, and the results show that the model can yield cleaner waveforms and achieve significant improvements in SNR and RMSE. In addition, Zhang et al. propose a novel CNN model (NovelCNN) to remove EMG artifacts from EEG signals, and the proposed network architecture improves the performance in the removal of EMG artifacts [26].

As a contrast to the above, this study aims 1) to enable the reconstruction of genuine EEG in scenarios with unknown/hybrid artifacts and 2) to appropriately capture the morphological characteristics to differentiate EEG from artifacts.

III. EEG RECONSTRUCTION BASED ON THE DUOCL MODEL

This section describes the proposed *DuoCL* model for automatic EEG reconstruction. This section consists of three parts: 1) an overview of the EEG reconstruction approach, 2) details of the dual-scale CNN-LSTM model, and 3) the learning process for EEG reconstruction/artifact removal.

A. The Overview of EEG Reconstruction Approach

Artifact removal can be regarded as the reconstruction of genuine EEG which is contaminated by the artifact components. Raw EEG evolves along the time dimension in a nonlinear and non-stationary manner, with artifacts complicating its structure even more. It is difficult to construct a mapping relation between raw and artifact-free EEG using linear methods without losing nonlinear information. Deep neural networks (DNNs) may ignore uncertainties and learn deep features representing the data distribution from nonlinear and non-stationary signals. Therefore, it is appropriate to use DNN to complete this reconstruction task.

Fig. 1 illustrates the DNN-based EEG reconstruction approach. The raw EEG over a long duration is split into short segments and fed into the training pipeline. Each layer in the model is regarded as a filter, learning to capture the core features of the initial inputs or outputs from the previous layer

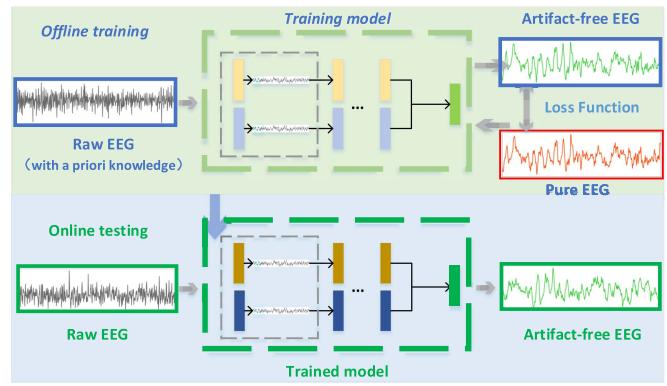


Fig. 1. Architecture of the EEG signal reconstruction model based on the deep neural network.

and producing filtered EEG waveforms for the next layer. The features extracted by the network can distinguish genuine EEG (non-artifacts) from artifacts. Subsequently, these features are then used to reconstruct an artifact-free EEG. By minimizing the loss function, the reconstructed EEG is as consistent as possible with the corresponding genuine EEG. In the online stage, the trained model directly operates on the raw EEG and generates artifact-free EEG as desired.

B. Dual-Scale CNN-LSTM Model

The *DuoCL* model consists of three parts for three phases: 1) *Morphological Feature Extraction*, 2) *Feature Reinforcement*, and 3) *EEG Reconstruction*. The raw EEG passes through Phases 1 and 2 in turn to derive the hybrid features. In Phase 3, the artifact-free EEG is reconstructed with the same length as the input by aggregating the features from the dual branches. The design aims to gain more opportunities to track real EEG.

The phase of *Morphological Feature Extraction* is at the core of the reconstruction model. Although the artifacts are similar to artifact-free signals in many ways, morphological characteristics can play a critical role in distinguishing the two. In this study, one-dimensional CNNs are applied to mine the local characteristics of raw EEG. When the EEG passes through the convolution block, the one-dimensional convolution operation is performed to obtain morphological features.

To support generalization, the network holds a vertical extension with five convolution blocks, with the number of convolution kernels progressively increasing by a factor of two, varying from 16 to 256, to obtain features with increasing dimensions. Furthermore, a 1D-Average pooling layer is added to each block to gradually downsample the EEG in the time domain. The last block is followed by a dropout layer to prevent overfitting.

EEG routinely manifests diverse morphological characteristics, showing different characteristics across different observation scales [27]. An end-to-end DNN model that derives features from dataspace with a fixed data scale is not sufficient to describe morphological characteristics [28]. As shown in Fig. 2, the features obtained by the same model are different at different observation scales (reception fields), and the resulting details of the reconstructed waveforms are also different.

The depth, width, and filter size are the most important factors in network construction. Deepening the network is a conventional way to enhance the feature processing capability of CNN.

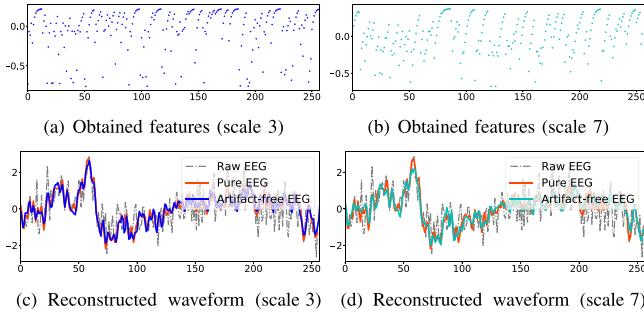


Fig. 2. Features obtained from the same network layer at different scales and the corresponding waveform results.

However, the features obtained by simply deepening the network cannot describe the diverse morphological characteristics at different scales. The essence of the one-dimensional convolution kernel is a filter for performing the convolution operation on a time series. The kernel size (reception field) determines the granularity of the extracted information, which has a significant impact on the level of the feature on which the consequent vector focuses, whether in a general or detailed view.

Inspired by the success of using a wide network architecture over narrow networks [14], [29], the proposed model adopts a multi-branches architecture with convolution kernels of different sizes in each branch to capture the different-scale morphological features of EEG and artifacts. Following an extensive investigation, it was found that the model performed well when employing the dual-branch structure with dual-scale convolution kernels (see Section IV), and adding more branches did not lead to more significant performance improvements. Considering the model's system complexity, execution efficiency, and performance, this study adopts a dual-branch network architecture. According to the preliminary experiments, the convolution kernel sizes of the convolution layers in the two branches are set to three and seven, respectively.

The phase of **Feature Reinforcement** aims to learn the temporal dependencies between samples, reinforcing the morphological features with the differences between the EEG and the artifacts highlighted. As shown in Fig. 3, the CNN blocks are followed by a flatten layer and two dense layers, and this design enhances the morphological features; the feature vectors are then reorganized for the use of LSTM. LSTM captures the temporal dynamics of the EEG to reinforce the feature representation.

The reorganized feature vector of each branch can be formalized as $A = (a_1, a_2, \dots, a_N)$. The process of feature reinforcement learning is defined as follows:

$$h_t, c_t = LSTM_\theta(h_{t-1}, c_{t-1}, a_t), 1 \leq t \leq T. \quad (1)$$

where h and c are vectors of hidden and cell states of the LSTM; $T = N$.

The LSTM generally conforms to the following functions:

$$f_t = \text{sigm}(W_f \cdot [h_{t-1}, a_t] + b_f), \quad (2)$$

$$i_t = \text{sigm}(W_i \cdot [h_{t-1}, a_t] + b_i), \quad (3)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, a_t] + b_c), \quad (4)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t, \quad (5)$$

$$h_t = o_t * \tanh(C_t), \quad (6)$$

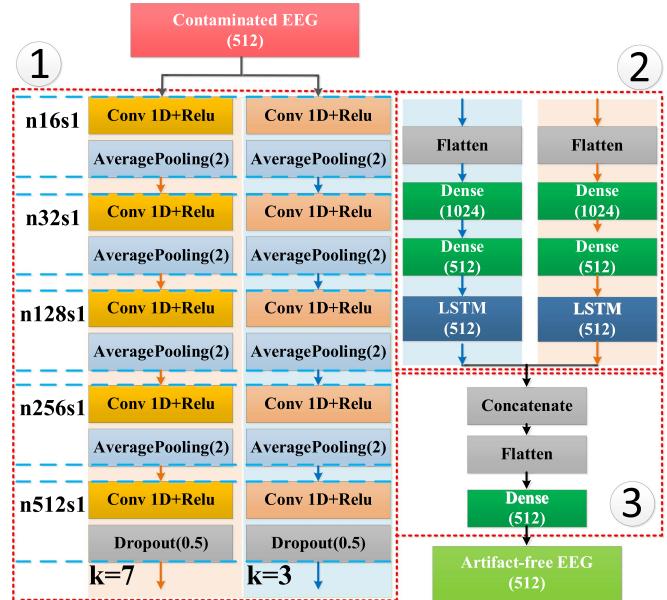


Fig. 3. Network structure of the proposed model.

where W denotes the weight matrices, and b denotes the biases. The f , \tilde{C} , and i variables represent the forget gate vector, the updated cell state vector candidate, and the input gate layer vector, respectively.

The LSTM then produces the sequence of output vectors $O = (o_1, o_2, \dots, o_T)$, where o_t is calculated as:

$$o_t = \text{sigm}(W_o \cdot [h_{t-1}, a_t] + b_o). \quad (7)$$

At the end of **Feature Reinforcement**, a set of dual-scale hybrid features can be obtained that maintains the morphological characteristics and temporal dependencies of the original EEG.

In the phase of **EEG Reconstruction**, the two feature vectors of size (512×1) with different observation scales are aggregated into one vector (1024×1) via the concatenate layer and the flatten layer. Then, the terminal fully connected layer outputs the reconstructed EEG with the same size as the inputs.

Fig. 3 provides an overview of the network architecture of *DuoCL*. The network consisted of two similar branch structures. Each branch comprises five convolution blocks with similar structures, two dense layers, and one LSTM unit. A one-dimensional raw EEG with a fixed length is regarded as the input of *DuoCL*. Through data conversion in each network layer, it is finally transformed into a series of dual-scale feature vectors of genuine EEG and artifacts. The module of **EEG Reconstruction** is relatively simple, with only three network layers, namely, concatenate, flatten, and dense layers.

The structure of *DuoCL* allows both reasonable horizontal and vertical expansions. Rather than simply stacking network layers together, a carefully designed structure can efficiently capture the deep hybrid features containing morphological characteristics and temporal dependencies from raw EEG. With the obtained features, *DuoCL* can better differentiate between genuine EEG and artifacts, allowing for a cleaner reconstruction of the raw EEG with lower information loss.

C. Learning Process

The essence of artifact removal is the tracking of genuine EEG from the raw EEG. Let $x \in R^{1 \times T}$ denote the genuine EEG conforming to the distribution $P(x)$. Let $y \in R^{1 \times T}$ denote the contaminated EEG (i.e., raw EEG by default in the remainder of this paper) conforming to another distribution $P(y)$. Furthermore, T denotes the length of a sample, that is, an individual segment of the raw EEG or genuine EEG. The goal here is to formulate a function $f(\Theta)$ that projects y to x_r :

$$f(\Theta) : y \longrightarrow x_r \quad (8)$$

The function maps raw EEG samples from $P(y)$ onto a certain distribution $P(x_r)$. The goal is to optimize the function $f(\Theta)$ to move $P(x_r)$ closer to $P(x)$. Here, Θ denotes the parameters to be learned. The reconstruction model is completed when the difference between $P(x_r)$ and $P(x)$ is minimized.

This study utilized data pairs with raw EEG and the corresponding pure EEG for training. The learning objective is pure EEG segments, to bring $P(x_r)$ closer to $P(x)$. The loss function of this study is MSE, which calculates the point-wise error between a pure EEG with a reconstructed EEG.

$$L(\Theta) = \frac{1}{n} \sum_{i=1}^n \|x_i - x_{ri}\|_2^2, \quad (9)$$

where n denotes the number of sampling points in each segment. In addition, this study uses the Adam optimizer to optimize *DuoCL* as well as the moments of $\beta_1 = 0.9$ and $\beta_2 = 0.999$. The initial learning rate is set to 10^{-3} . The batch size is set to 512, and the epoch of model training is set to 100. (See *Supplementary Materials*¹ for the formal flowchart for *DuoCL*).

IV. EXPERIMENTS AND RESULTS

Experimental studies were conducted to evaluate the performance of *DuoCL* using three semi-simulated EEG datasets including: 1) ablation experiments to examine the effectiveness of the design; 2) performance evaluation of *DuoCL* against mainstream counterparts; and 3) effectiveness examination of *DuoCL* in handling multi-channel EEG.

A. Performance Metrics

The performance of *DuoCL* was examined in terms of subjective evaluation, objective evaluation, and nonlinear characteristics. The subjective evaluation was performed through visual inspection. Several objective metrics were used to complete the quantitative performance analysis, including signal-to-noise ratio (SNR), relative root mean square error (RRMSE) in the temporal domain ($RRMSE_t$), and the spectral domain ($RRMSE_f$).

SNR (see (10)) reflected the signal quality after artifact removal and the larger values are expected. $RRMSE_t$ (see (11)) and $RRMSE_f$ (see (12)) defined the energy error of the signal in the temporal and spectral domains, respectively. Smaller values indicated better reconstruction for the raw EEG. In addition, PSD, average correlation coefficient (CC, see (13)), and autocorrelation function (ACF) were applied to measure the influence of these methods on the nonlinear characteristics of

the reconstructed EEG.

$$SNR = 10 \log \frac{\sum_{i=1}^n x_i^2}{\sum_{i=0}^{n-1} (x_i - x_{ri})^2}, \quad (10)$$

$$RRMSE_t = \frac{RMS(x_r - x)}{RMS(x)}, \quad (11)$$

$$RRMSE_f = \frac{RMS(PSD(x_r) - PSD(x))}{RMS(PSD(x))}, \quad (12)$$

where PSD was the power spectral density of the input signal.

$$CC = \frac{Cov(x_r, x)}{\sqrt{Var(x_r)Var(x)}}, \quad (13)$$

where functions Cov and Var denoted the covariance and variance of the signal, respectively.

B. Datasets for Performance Evaluation

In practice, the underlying artifact-free brain recordings were actually unavailable. How closely close the reconstructed EEG approximated the genuine EEG remained unclear. Simulated datasets might apply in order for a quantitative evaluation of artifact removal methods. In this study, three semi-simulated datasets were constructed for this purpose, including two single-channel EEG datasets ((Dataset: I & II)) and one multi-channel EEG dataset (Dataset: III).

Contaminated EEG (EEG_C) was synthesized by “pure” EEG (EEG_P) with dedicated artifacts (such as EOG) from real datasets commonly used by the neuroscience community (see *Supplementary Materials* for more details). The ground-truth EEG EEG_P blended in the three datasets were then used to assess the quality of the reconstructed EEG (EEG_R).

The overview of constructing single-channel semi-simulated EEG datasets ((Dataset: I & II)) was shown in Fig. 1 of *Supplementary Materials*. The pure EEG was contaminated by linearly mixing additional artifacts, according to (14).

$$EEG_C = EEG_P + \lambda \cdot A, \quad (14)$$

where A denoted artifacts, and λ denoted a hyperparameter that controlled the SNR of EEG_C . The desired SNR of EEG_C was obtained by changing λ according to (15).

$$SNR = 10 \log \frac{RMS(EEG_P)}{RMS(EEG_C)}, \quad (15)$$

where RMS was the root-mean-square energy of the signal.

According to (14) and (15), both the training and test sets were constructed by combining signals at multiple SNR levels. Four types of artifacts, EMG, EOG, ECG, and EM, were used to construct **Dateset I**. The synthesis process was illustrated by an example of EEG segments contaminated by EMG artifacts. In constructing the test set, 10% of the segments were first randomly selected from EMG artifacts, and the same number of the segments of pure EEG were also selected. Every EMG artifact segment was randomly matched one-to-one with a pure EEG segment, and mixing processes were performed once at each of the ten SNR levels ($-7, -6, -5, -4, -3, -2, -1, 0, 1$, and 2 dB). The training set was constructed from the remaining segments. The pure EEG segments and artifact segments were randomly mixed in pairs at a random SNR level within -7 to 2 dB. Finally, the aforementioned two steps were repeated ten times to increase the amount of data in the training set. All

¹Supplementary Materials can be referred to the website: https://figshare.com/articles/media/supplementary_materials/21670550.

TABLE I
AVERAGE PERFORMANCES OF ALL SNR LEVELS IN ABLATION EXPERIMENT

| Module | | EEG _C | S3 | S7 | LSTM | S3-LSTM | S7-LSTM | S3-S7 | TFA-MLP | TFA-LSTM-MLP | S3-S7-SC | DuoCL |
|--------|--------------------|------------------|---------|---------|--------|---------|---------|---------|---------|--------------|----------|---------|
| EMG | SNR | 1.1183 | 10.3129 | 10.9755 | 6.0691 | 11.5326 | 12.1110 | 12.6547 | 2.0737 | 2.4556 | 7.4210 | 14.3146 |
| | RRMSE _t | 0.9964 | 0.4752 | 0.4730 | 0.5632 | 0.3849 | 0.3766 | 0.3942 | 0.8811 | 0.9092 | 0.5138 | 0.3339 |
| | RRMSE _f | 0.7775 | 0.4774 | 0.4274 | 0.6122 | 0.3932 | 0.3725 | 0.3753 | 0.8035 | 0.6998 | 0.4548 | 0.3338 |
| | CC | 0.5023 | 0.8829 | 0.8671 | 0.8443 | 0.9231 | 0.9268 | 0.9196 | 0.5026 | 0.5034 | 0.8597 | 0.9426 |
| EOG | SNR | 1.8764 | 14.9161 | 15.4184 | 7.4661 | 15.8281 | 16.2903 | 18.2601 | 9.7010 | 11.1216 | 9.6139 | 19.7519 |
| | RRMSE _t | 0.9965 | 0.1952 | 0.1835 | 0.4702 | 0.1708 | 0.1617 | 0.1327 | 0.3983 | 0.3468 | 0.3598 | 0.1092 |
| | RRMSE _f | 1.6235 | 0.1488 | 0.1355 | 0.4855 | 0.1220 | 0.1126 | 0.1191 | 0.4150 | 0.2906 | 0.3342 | 0.0869 |
| | CC | 0.5039 | 0.9808 | 0.9830 | 0.8939 | 0.9853 | 0.9868 | 0.9913 | 0.9168 | 0.9378 | 0.9325 | 0.9941 |
| ECG | SNR | 3.6130 | 13.1335 | 15.0211 | 7.2736 | 14.6262 | 16.0865 | 18.0038 | 9.6542 | 10.9499 | 9.6871 | 19.4961 |
| | RRMSE _t | 1.0284 | 0.2478 | 0.1920 | 0.4790 | 0.2079 | 0.1699 | 0.1332 | 0.3983 | 0.3699 | 0.3564 | 0.1203 |
| | RRMSE _f | 1.2489 | 0.2025 | 0.1500 | 0.5143 | 0.1614 | 0.1259 | 0.1107 | 0.4219 | 0.3099 | 0.3306 | 0.0991 |
| | CC | 0.5151 | 0.9689 | 0.9793 | 0.8908 | 0.9781 | 0.9854 | 0.9892 | 0.9065 | 0.9267 | 0.9340 | 0.9928 |
| EM | SNR | -1.2107 | 11.3619 | 11.8233 | 6.7281 | 12.5755 | 12.7823 | 13.6179 | 6.1835 | 7.2786 | 8.0830 | 15.0860 |
| | RRMSE _t | 1.1641 | 0.3760 | 0.3724 | 0.5227 | 0.3085 | 0.3092 | 0.3141 | 0.6307 | 0.6075 | 0.4586 | 0.2672 |
| | RRMSE _f | 6.4541 | 0.3566 | 0.3405 | 0.5511 | 0.3147 | 0.3022 | 0.3169 | 0.6072 | 0.4961 | 0.4144 | 0.2764 |
| | CC | 0.4715 | 0.9269 | 0.9284 | 0.8626 | 0.9519 | 0.9514 | 0.9501 | 0.7778 | 0.7983 | 0.8993 | 0.9640 |

EEG_C: the input contaminated EEG segments; S3: Single-branch CNN with a convolution kernel size of 3; S7: Single-branch CNN with a convolution kernel size of 7; S3-S7: dual-branch CNN without LSTM units; SC: self-correlation; TFA: time-frequency analysis; and MLP: the network structure in *EEG Reconstruction* phase.

the simulated EEG_C segments synthesized by the four types of artifacts and the corresponding EEG_P segments together formed the training set.

An unknown artifact test dataset (**Dataset II**) was constructed to explore the robustness of the model when processing raw EEG signals contaminated by unknown artifacts. The BW artifact was used as an unknown artifact because it was completely different from the artifacts in Dataset I. The SNR levels of the BW artifact were set to -10 dB, -5 dB, 0 dB, 5 dB, and 10 dB, which were different from the SNR levels in Dataset I. Thus, the experimental conditions were more realistic. Furthermore, hybrid artifacts were constructed by mixing different types of artifacts (BW, EMG, EOG, and ECG) to fully evaluate the performance of removing more complex artifacts. The setting of the SNR levels was similar to that for unknown artifacts.

To evaluate the performance of the model in removing artifacts from multi-channel EEG, **Dataset III** was constructed based on the public dataset (KLADOS) [30]. As the ocular artifacts were prominent in the frontal region, the eight channels (FP1, FP2, F3, F4, F7, F8, T3, and T4) were chosen for examination.

C. Design Examination

In this study, the dual-branch structure with the dual-scale convolution kernels and the introduction of LSTM units were the two major designs for better reconstructions of EEG. To verify the contributions of these designs, *DuoCL* was separated to perform ablation experiments with Dataset I, whereas the complete model was regarded as the baseline. In addition, self-correlation and time-frequency analysis methodologies were also substituted into the model to explore the capability of *DuoCL* in enhancing temporal dependencies and capturing morphological features, respectively.

Note that the dual-branch structure (morphological analysis) was replaced by the time-frequency feature extractor [14]. It extracted time-frequency EEG features with the short-time Fourier transform (STFT). In addition, the temporal dependencies between feature points were referred to as the self-correlation coefficients of the morphological feature vectors for each branch, rather than the LSTM any more (see *Supplementary Materials* for comparison).

As shown in Table I, S and LSTM denoted the designs of dual-branch and LSTM units, respectively. S3, S7, and S3-S7

denoted single-scale (3 & 7) single-branch convolution feature extractors and *DuoCL* without LSTM, respectively. The main observations of the results were as follows:

- The best results were obtained for the complete network (*DuoCL*), whereas the worst results were obtained using LSTM alone. Compared with the single-branch structure (S3, S7, S3-LSTM, and S7-LSTM), the model with a dual-branch structure (S3-S7 and *DuoCL*) achieved higher artifact removal performance.
- The model (S3-LSTM, S7-LSTM, and *DuoCL*) obtained better results after the introduction of LSTM units. Moreover, the results of TFA-MLP and TFA-LSTM-MLP also showed this, although the morphological feature extraction was replaced by the time-frequency analysis method.
- S3-S7-SC did not get better results compared to *DuoCL*, suggesting that self-correlation was weaker than LSTM for the extraction of temporal dependencies in this study. In addition, a performance dip was observed after replacing the morphological feature extraction with the time-frequency analysis method, especially when removing EMG artifacts, which may be caused by the inability for severe frequency overlap.

The aforementioned results suggested that 1) the two-branch structural design facilitated the tracking of real EEG, and 2) the introduction of LSTM enhanced the outputs of the former phase.

D. Performance Evaluation

Experiments were conducted to compare *DuoCL* with its state-of-the-art counterparts and evaluate the capability in terms of removing artifacts. These methods were divided into two categories: 1) traditional methods: Wiener filtering (WF) and EMD; and 2) DL-based methods: fully connected network (FCNN), LSTM, 1-ResCNN, and NovelCNN. As reported in Table II, all of the DL-based methods shared the same attributes, and NovelCNN held the largest number of network parameters. (See *Supplementary Materials* for more details, e.g., the network structures of DL-methods).

1) Benchmark Performance Evaluation Via Dataset I: The benchmark experiments were conducted to test the capability of *DuoCL* to eliminate artifact components from the raw EEG. To provide a qualitative analysis, the results in the temporal

TABLE II
A COMPARISON OF MENTIONED METHODS

| Type | Method | Reference channel | Online | Single channel | Pure data for training | Network parameter (million) |
|--------------------|-----------|-------------------|--------|----------------|------------------------|-----------------------------|
| Traditional method | WF | N | Y | Y | Y | - |
| | EMD | N | Y | Y | N | - |
| DL-based method | FCNN | N | Y | Y | Y | 1.05 m |
| | LSTM | N | Y | Y | Y | 0.78 m |
| | 1D-ResCNN | N | Y | Y | Y | 8.46 m |
| | NovelCNN | N | Y | Y | Y | 33.56 m |
| | DuoCL | N | Y | Y | Y | 18.79 m |

Y: Yes, N: No. All the methods used can be applied for online artifact removal of single-channel EEG without reference channels.

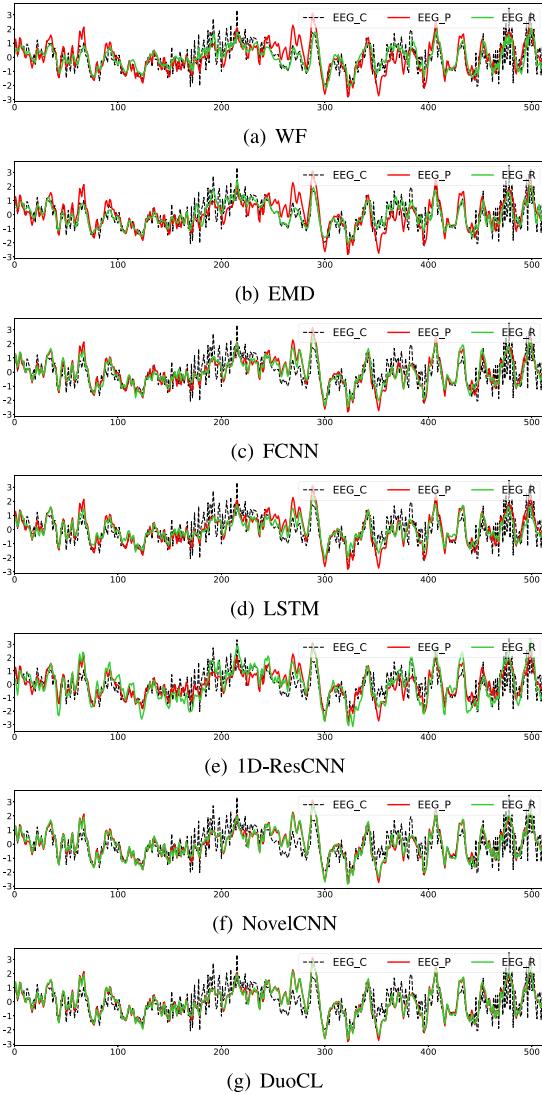


Fig. 4. Example of experimental waveform results for eliminating EMG artifacts.

and frequency domains were shown in followed figures. The waveform results of eliminating EMG, EOG, ECG, and EM artifacts from the contaminated EEG at SNR = 0 dB were shown in Fig. 4, *Supplementary Materials* Figs. 3, 4, and 5, respectively. The black line indicated the raw EEG with artifacts, whereas the red and green lines indicated the pure EEG and reconstructed EEG (EEG_R), respectively. A better fit of the red and green lines indicated less loss in the temporal domain and better artifact removal. The main observations of the results were as follows:

- Compared with the contaminated EEG, the EEG_R of all methods had a large change in the waveform. It could be observed that the reconstructions of the 1D-ResCNN, NovelCNN, and *DuoCL* were significantly better than the other five methods. However, it was noted that the reconstructed waveforms of 1D-ResCNN and NovelCNN had peak amplitude overflow and local waveform offset, which indicated that some details of the waveform were lost (see Fig. 4(e), 4(f), *Supplementary Materials* Fig 3(e)). On contrary, *DuoCL* eliminated almost all artifacts, and the EEG_R better restored the detailed features of low frequency and small amplitude.
- The EMD method eliminated most of the artifacts, however, the reconstructed waveforms had significant distortion compared to the EEG_P . All the waveforms were extremely smooth, and the amplitudes were reduced, which indicated that many high-frequency signal details were lost.
- The reconstructions of WF, FCNN, and LSTM were better than EMD visually; however, the peak characteristics of the reconstructed waveforms had not been restored, and there was a significant difference in amplitude from that of the EEG_P . In particular, more burrs existed in the waveforms of the reconstruction via WF, which indicated that high-frequency artifacts had not been completely eliminated.
- Observing all the reconstructed waveforms, it could be concluded that the removal effect of EMG artifacts was slightly worse than the other three types of artifacts.

Fig. 5(a), (b), (c) and (d) showed the PSD results of Dataset I. The black line indicated the PSD of EEG_P which was used as the ground truth. The closer the PSD line was to the black line, the less loss in the frequency domain. The main observations of the results were as follows:

- The fitting degrees of DL-based methods were higher than those of traditional methods. All the PSD lines of *DuoCL* had the highest degree of fitting with the black lines, which indicated that the removal of these four artifacts using *DuoCL* lost the least amount of frequency information. NovelCNN achieved the second-best PSD results, although its PSD lines were distorted in some frequency bands (e.g. [65–70] Hz in Fig. 5(c)). Furthermore, FCNN and LSTM lost the most frequency information, especially when handling EMG and EOG artifacts.
- EMD performed the worst in PSD results, and its PSD lines had the highest deviation from the ground truth. The PSD lines of Fig. 5(c) indicated that the EMD lost a significant

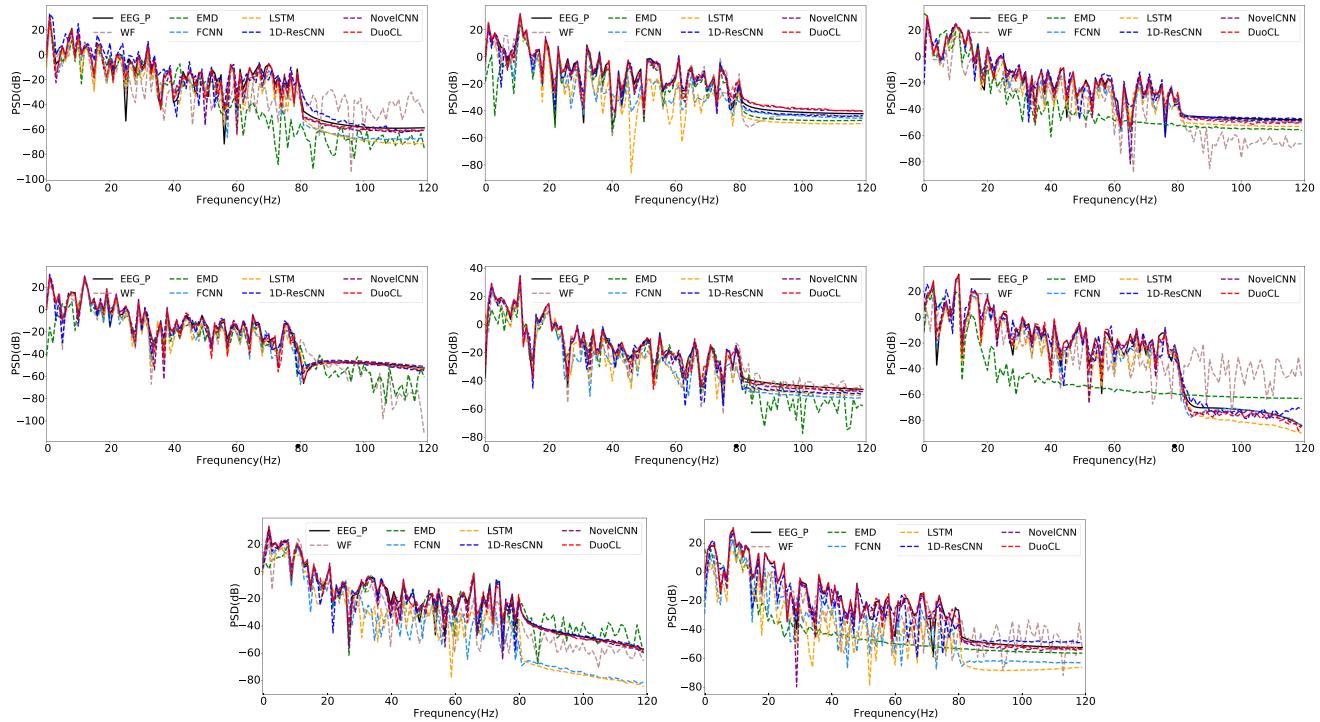


Fig. 5. PSD results after eliminating artifacts from the contaminated EEG.

amount of frequency information above 20 Hz. In addition, the results showed that WF could not eliminate the artifact components in high-frequency bands (e.g. [80-120]Hz in Fig. 5(c)).

The SNR level described the degree of contamination by artifacts. Fig. 6 illustrated the results for four metrics (SNR, RRMSE_t, RRMSE_f, and CC) at different SNR levels. The main observations of the results were as follows:

- *DuoCL* obtained the highest SNR and CC values and the lowest RRMSE_t and RRMSE_f values at each SNR level of all types of artifacts. The two methods, *DuoCL* and NovelCNN performed better, however *DuoCL* had a superior capability for improving signal SNR than NovelCNN. When the SNR level was -7 dB, the SNR values of *DuoCL* for removing EOG and ECG artifacts were close to the SNR values of NovelCNN when the SNR level was 2 dB.
- As the SNR level increased, the quality of the reconstructed signal improved for all methods. According to the CC results, the performance of the traditional methods (WF and EMD) was more sensitive to the degree of artifact contamination.
- The results obtained by *DuoCL* at different SNR levels of the EOG and ECG artifacts exhibited fewer variations and those of EMG exhibited larger variations, which indicated the degree of contamination of EMG artifacts had a greater influence on the performance of artifact removal methods than other three types artifacts.

The quantitative benchmarks obtained by averaging the values for all methods across all SNRs (-7 dB to 2 dB) were summarized in Table III. *DuoCL* had the lowest average RRMSE_t and RRMSE_f, as well as the highest average SNR. The values

of these three quantitative metrics indicated that the EEG_R via *DuoCL* maintained the smallest deviation in both the temporal and frequency domains and the highest data quality. *DuoCL* outperformed the comparison method in terms of removing these four types of artifacts. The larger the CC values and the smaller the difference in the ACF values of EEG_P and EEG_R were, the better the preservation of the nonlinear characteristics. As shown in the Tables III and IV, *DuoCL* obtained the minimum difference values of ACF and the maximum CC values, indicating that nearly all of the genuine EEG information was preserved and the artifacts were eliminated.

2) Exploration Performance Evaluation Via Dataset II: The experiments in this section evaluated the potential of *DuoCL* to reconstruct raw EEG contaminated by unknown or hybrid artifacts, which were similar to practical application scenarios. The model trained by dataset I was used directly to complete the artifact removal of dataset II.

As the results were shown in Table III and IV, when eliminating unknown and hybrid artifacts, *DuoCL* still obtained superior results in each metric: the highest SNR and CC values, and the lowest RRMSE_t, RRMSE_f and the differences in the ACF. To clearly demonstrate the artifact removal effect of *DuoCL*, Figs. 7 and 5 (e, f, g, and h) showed examples of resulting waveforms and PSD results. The waveforms of all EEG_R and EEG_P maintained good overlap and the PSD results of EEG_R segments remained consistent with the ground truth. The results indicated that the trained model could effectively remove the artifacts, maintain the nonlinearity of original signal, and improve the quality of raw EEG. In addition, this study also provided the results of each method for removing unknown & hybrid artifacts at different SNR levels. Data quality was improved at all levels after artifact removal, and *DuoCL* excelled. Details were available at *Supplementary Materials*.

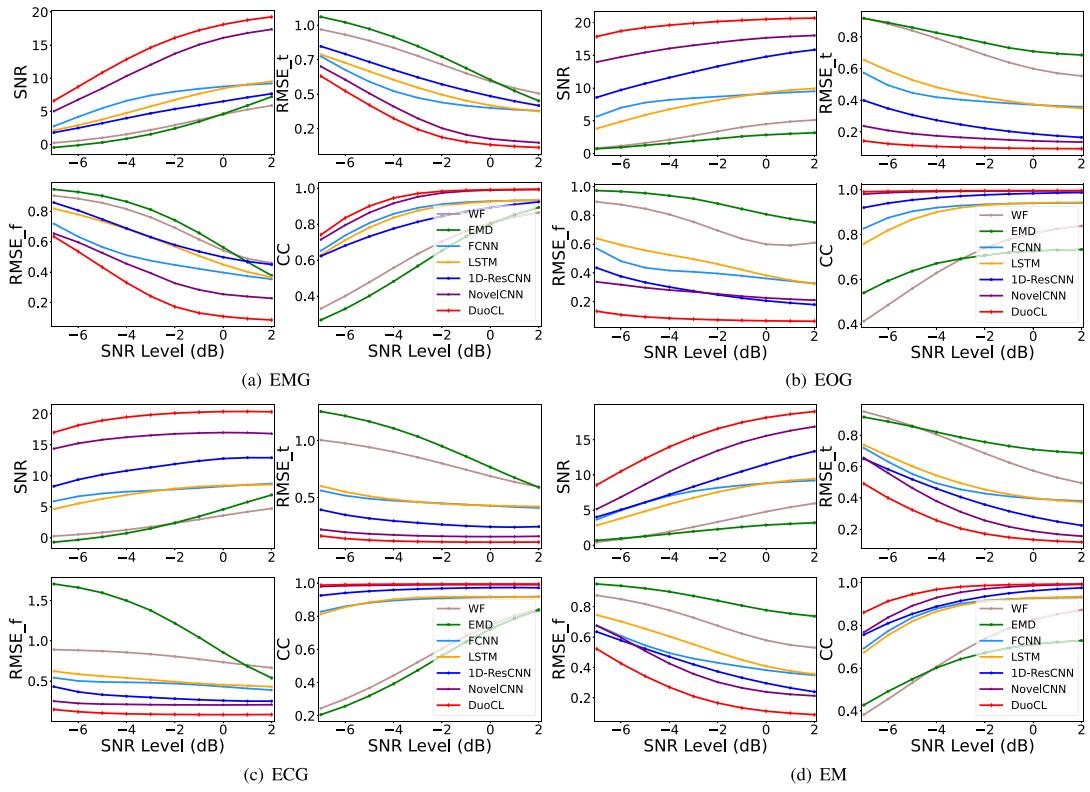


Fig. 6. Experimental results at different SNR levels after artifact removal.

TABLE III
AVERAGE PERFORMANCE OF MULTIPLE ARTIFACTS AT MULTIPLE SNR LEVELS

| Dataset | Artifact | Metric | EEG _C | WF | EMD | FCNN | LSTM | 1D-ResCNN | NovelCNN | DuoCL |
|-------------------|----------------|--------------------|------------------|--------|--------|--------|--------|-----------|----------|---------|
| Dataset I | EMG | SNR | 1.1183 | 2.7800 | 2.5614 | 6.9750 | 6.0691 | 4.8681 | 12.1682 | 14.3146 |
| | | RRMSE _t | 0.9964 | 0.7570 | 0.8113 | 0.5211 | 0.5632 | 0.6275 | 0.3958 | 0.3339 |
| | | RRMSE _f | 0.7775 | 0.7189 | 0.7501 | 0.5016 | 0.6122 | 0.6405 | 0.4217 | 0.3338 |
| | | CC | 0.5023 | 0.6529 | 0.6091 | 0.8555 | 0.8443 | 0.7907 | 0.9218 | 0.9426 |
| | EOG | SNR | 1.8764 | 3.0348 | 2.0390 | 8.2616 | 7.4661 | 12.6446 | 16.4442 | 19.7519 |
| | | RRMSE _t | 0.9965 | 0.7330 | 0.7911 | 0.4257 | 0.4702 | 0.2639 | 0.1738 | 0.1092 |
| | | RRMSE _f | 1.6235 | 0.7391 | 0.8836 | 0.4175 | 0.4855 | 0.2883 | 0.2689 | 0.0869 |
| | | CC | 0.5039 | 0.6805 | 0.6424 | 0.9122 | 0.8939 | 0.9646 | 0.9882 | 0.9941 |
| | ECG | SNR | 3.6130 | 2.6488 | 2.4587 | 7.6413 | 7.2736 | 11.3147 | 16.2696 | 19.4961 |
| | | RRMSE _t | 1.0284 | 0.8059 | 0.9889 | 0.4656 | 0.4790 | 0.2899 | 0.1722 | 0.1203 |
| | | RRMSE _f | 1.2489 | 0.8043 | 1.2835 | 0.4661 | 0.5143 | 0.3090 | 0.2136 | 0.0991 |
| | | CC | 0.5151 | 0.5650 | 0.5099 | 0.8917 | 0.8908 | 0.9574 | 0.9860 | 0.9928 |
| | EM | SNR | -1.2107 | 3.0650 | 2.0505 | 7.3106 | 6.7281 | 8.8103 | 12.0028 | 15.6860 |
| | | RRMSE _t | 1.1641 | 0.7343 | 0.7885 | 0.4982 | 0.5227 | 0.4274 | 0.3740 | 0.2672 |
| | | RRMSE _f | 6.4541 | 0.7105 | 0.8527 | 0.4829 | 0.5511 | 0.4329 | 0.4115 | 0.2764 |
| | | CC | 0.4715 | 0.6793 | 0.6207 | 0.8718 | 0.8626 | 0.9038 | 0.9299 | 0.9640 |
| Dataset II | BW | SNR | 1.4750 | 5.2056 | 2.4106 | 8.0311 | 8.1032 | 12.0469 | 16.3242 | 18.6923 |
| | | RRMSE _t | 1.5773 | 0.6257 | 0.7631 | 0.4916 | 0.4735 | 0.3871 | 0.2349 | 0.1914 |
| | | RRMSE _f | 38.0192 | 0.7442 | 0.8191 | 0.5003 | 0.4632 | 0.3864 | 0.2984 | 0.1979 |
| | | CC | 0.42194 | 0.7815 | 0.6481 | 0.8713 | 0.8823 | 0.9222 | 0.9747 | 0.9805 |
| | BW+EMG | SNR | 4.8758 | 3.7324 | 2.2231 | 7.1155 | 7.0309 | 6.9266 | 13.2524 | 15.2404 |
| | | RRMSE _t | 0.9730 | 0.7108 | 0.8491 | 0.5669 | 0.5521 | 0.5930 | 0.4434 | 0.3979 |
| | | RRMSE _f | 2.3654 | 0.6926 | 0.8658 | 0.5214 | 0.5456 | 0.5648 | 0.4354 | 0.3911 |
| | | CC | 0.5773 | 0.7070 | 0.5393 | 0.8229 | 0.8354 | 0.8151 | 0.9001 | 0.9173 |
| | EMG+EOG+BW | SNR | 3.8537 | 3.9818 | 2.3038 | 7.6452 | 7.6122 | 11.8098 | 15.68441 | 18.1581 |
| | | RRMSE _t | 1.1061 | 0.7034 | 0.7759 | 0.5166 | 0.5093 | 0.4017 | 0.2736 | 0.2269 |
| | | RRMSE _f | 8.0222 | 0.7231 | 0.8339 | 0.5025 | 0.4881 | 0.4019 | 0.3158 | 0.2324 |
| | | CC | 0.5545 | 0.7160 | 0.6334 | 0.8595 | 0.8626 | 0.9157 | 0.9636 | 0.9739 |
| | EMG+EOG+BW+ECG | SNR | 5.8063 | 3.4477 | 1.4126 | 6.5572 | 6.2126 | 6.3318 | 12.3770 | 15.6828 |
| | | RRMSE _t | 0.9237 | 0.7537 | 0.9316 | 0.5942 | 0.6053 | 0.6433 | 0.4732 | 0.3749 |
| | | RRMSE _f | 1.4614 | 0.7055 | 0.9316 | 0.5277 | 0.6131 | 0.5985 | 0.4420 | 0.3721 |
| | | CC | 0.5959 | 0.6565 | 0.4543 | 0.8033 | 0.7960 | 0.7886 | 0.8827 | 0.9270 |

TABLE IV
ACF RESULTS BEFORE AND AFTER ARTIFACT REMOVAL

| | EMG | EOG | ECG | EM | BW | BW+EMG | EMG+EOG+BW | EMG+EOG+BW+ECG |
|------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|
| EEG _P | 0.7696 | 0.7712 | 0.7699 | 0.7688 | 0.7683 | 0.7682 | 0.7689 | 0.7688 |
| EEG _C | 0.2590(± 0.5106) | 0.9229(± 0.1517) | 0.7262(± 0.0436) | 0.8852(± 0.1164) | 0.8744(± 0.0886) | 0.5351(± 0.2331) | 0.8891(± 0.0902) | 0.6893(± 0.0795) |
| WF | 0.6888(± 0.0808) | 0.7876(± 0.0164) | 0.7388(± 0.0311) | 0.7924(± 0.0236) | 0.7786(± 0.0106) | 0.7147(± 0.0536) | 0.8615(± 0.0927) | 0.7505(± 0.0183) |
| EMD | 0.6051(± 0.1645) | 0.6588(± 0.1124) | 0.8733(± 0.1034) | 0.6472(± 0.1216) | 0.5273(± 0.2410) | 0.9528(± 0.1846) | 0.5627(± 0.2061) | 0.9652(± 0.1964) |
| FCNN | 0.8431(± 0.0735) | 0.8453(± 0.0741) | 0.8432(± 0.0733) | 0.8428(± 0.0740) | 0.8425(± 0.0742) | 0.8419(± 0.0737) | 0.8456(± 0.0768) | 0.8454(± 0.0766) |
| LSTM | 0.8396(± 0.0700) | 0.8487(± 0.0775) | 0.8408(± 0.0709) | 0.8423(± 0.0735) | 0.8363(± 0.0680) | 0.8389(± 0.0707) | 0.8419(± 0.0730) | 0.8395(± 0.0707) |
| 1D-ResCNN | 0.7985(± 0.0289) | 0.7609(± 0.0103) | 0.7653(± 0.0045) | 0.7500(± 0.0188) | 0.7518(± 0.0165) | 0.7809(± 0.0127) | 0.7524(± 0.0164) | 0.7957(± 0.0269) |
| NovelCNN | 0.7966(± 0.0271) | 0.7864(± 0.0152) | 0.7908(± 0.0209) | 0.7958(± 0.0270) | 0.7860(± 0.0177) | 0.7939(± 0.0256) | 0.7890(± 0.0202) | 0.7994(± 0.0307) |
| DuoCL | 0.7649(± 0.0053) | 0.7770(± 0.0057) | 0.7767(± 0.0069) | 0.7751(± 0.0063) | 0.7732(± 0.0050) | 0.7605(± 0.0077) | 0.7735(± 0.0047) | 0.7570(± 0.0118) |

The lag coefficient (K) was set to 2. The values outside the brackets were the mean ACF values of reconstructed EEG, and the values in brackets were the difference from the pure EEG, and the smaller the difference was, the better the preservation of nonlinear features.

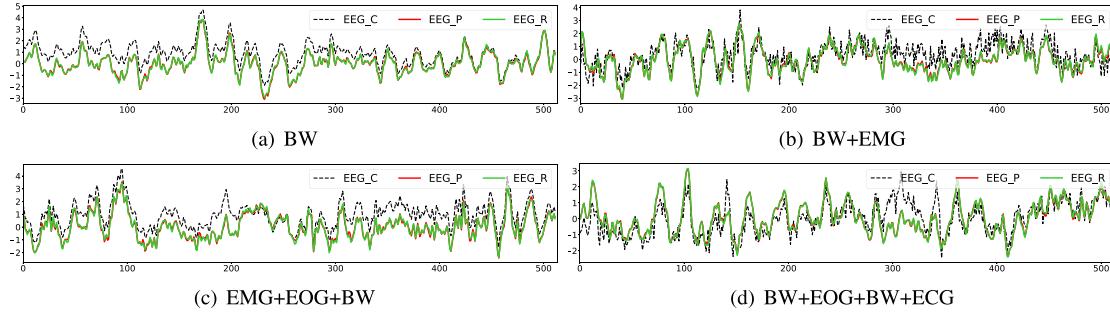


Fig. 7. The waveform results after eliminating unknown and hybrid artifacts by DuoCL from contaminated EEG signal when SNR = 0 dB.

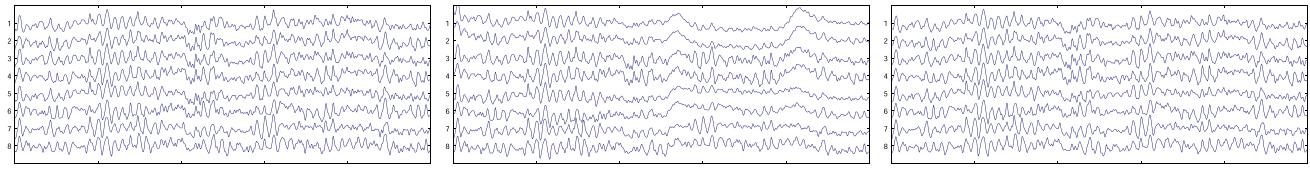


Fig. 8. The waveform results of DuoCL for the multi-channel contaminated EEG.

In summary, the above experimental results demonstrated that *DuoCL* was able to effectively eliminate unknown and hybrid artifacts in EEG.

3) Validation Performance Evaluation Via Dataset III: The experiments in this section examined the validity of *DuoCL* for processing multi-channel EEG. DL-based methods were checked against *DuoCL* for their good performance. Considering the limited number of samples in Dataset III, the models under evaluation were obtained by fine-tuning the trained model with EOG artifact contaminated data segments of Dataset I. Leave-one-out cross validation (LOOCV) was applied to evaluate the performances of the models.

Fig. 8 showed an example of processing multi-channel EEG with *DuoCL*. The artifacts highlighted in contaminated EEG were removed in artifact-free EEG, and the artifact-free EEG was almost the same as the pure EEG. Detailed results of the channels before and after artifact removal were listed in Table V. It could be observed that all four metrics of each channel were improved after the artifacts removal using *DuoCL*. Table VI showed the comparison of the average performance of artifact removal for multi-channel EEG. The results suggested that *DuoCL* still achieved better results in all four metrics. In general, *DuoCL* could effectively perform the task of artifact removal from multi-channel EEG.

TABLE V
THE RESULTS FOR EACH CHANNEL BEFORE AND AFTER ARTIFACT REMOVAL BY DUOCL

| Channel | | SNR | RRMSE _t | RRMSE _f | CC |
|---------|------------------|---------|--------------------|--------------------|--------|
| FP1 | EEG _C | 0.4487 | 1.0820 | 2.7492 | 0.5064 |
| | EEG _R | 13.9838 | 0.2205 | 0.2186 | 0.9756 |
| FP2 | EEG _C | 0.7399 | 1.0755 | 2.4893 | 0.5112 |
| | EEG _R | 13.5376 | 0.2697 | 0.1898 | 0.9621 |
| F3 | EEG _C | 5.6620 | 0.7009 | 1.2975 | 0.7686 |
| | EEG _R | 14.8757 | 0.1966 | 0.0994 | 0.9800 |
| F4 | EEG _C | 5.7197 | 0.7091 | 1.1539 | 0.7619 |
| | EEG _R | 14.7455 | 0.1975 | 0.1516 | 0.9807 |
| F7 | EEG _C | 1.4940 | 0.9779 | 3.9126 | 0.6129 |
| | EEG _R | 15.2137 | 0.1757 | 0.1083 | 0.9847 |
| F8 | EEG _C | 1.7888 | 0.9616 | 2.8363 | 0.6212 |
| | EEG _R | 13.9726 | 0.2201 | 0.2033 | 0.9757 |
| T3 | EEG _C | 5.1082 | 0.6688 | 1.9738 | 0.7982 |
| | EEG _R | 13.0602 | 0.2432 | 0.1586 | 0.9714 |
| T4 | EEG _C | 5.3625 | 0.6626 | 1.5586 | 0.8000 |
| | EEG _R | 14.2450 | 0.2140 | 0.1591 | 0.9773 |

TABLE VI
AVERAGE PERFORMANCES OF ARTIFACT REMOVAL FOR MULTI-CHANNEL EEG

| Method | EEG _C | FCNN | LSTM | ID-ResCNN | NovelCNN | DuoCL |
|--------------------|------------------|--------|--------|-----------|----------|---------|
| SNR | 3.2905 | 7.0579 | 7.8412 | 10.1223 | 13.0911 | 14.2043 |
| RRMSE _t | 0.8548 | 0.4594 | 0.4191 | 0.3363 | 0.2596 | 0.2172 |
| RRMSE _f | 2.2464 | 0.4137 | 0.4131 | 0.3710 | 0.2690 | 0.1611 |
| CC | 0.6726 | 0.8994 | 0.9197 | 0.9409 | 0.9648 | 0.9759 |

TABLE VII
AVERAGE INFERENCE TIMES FOR THE REMOVAL OF ARTIFACTS FROM
2-SECONDS EEG SIGNAL

| Methods | Tested | #1 | #2 | #3 | #4 | Average |
|----------|--------|------------|------------|------------|------------|------------|
| NovelCNN | | 3.3394(ms) | 2.1393(ms) | 2.5332(ms) | 2.6969(ms) | 2.6772(ms) |
| DuoCL | | 1.7906(ms) | 1.3369(ms) | 1.5292(ms) | 1.4245(ms) | 1.5203(ms) |

#1: Intel(R) Core(TM)i5-7500 @3.4GHz; #2: Intel(R) Core(TM)i5-8400 @2.8GHz;
#3: Intel(R) Core(TM)i7-10700 @2.9GHz; #4: Intel(R) Core(TM)i7-11390H @3.4GHz

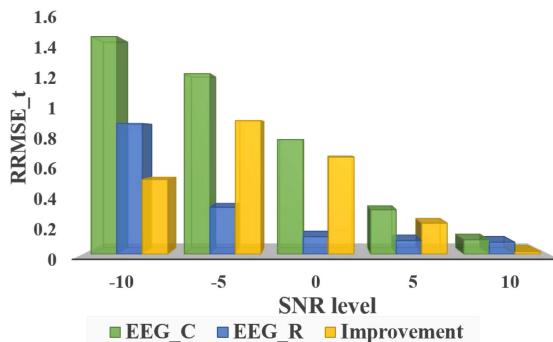


Fig. 9. Results and degree of improvements of DuoCL before and after removal of hybrid artifacts (EMG+EOG+BW+ECG) at different contamination levels.

4) Inference Time Analysis: Considering the scenarios of online EEG analysis, the current artifact removal session should have ended on the arrival of the next EEG segment. In other words, the inference time of the model should be short enough. This study also evaluated the models (NovelCNN and *DuoCL*) with the top performance in the previous experiments on the four different devices. All tests were performed using only the CPU, with no acceleration from the GPU.

As shown in Table VII, *DuoCL* took an average of 1.5203 milliseconds to process an EEG segment of a 2 s duration. Its inference time was shorter in three magnitudes than the recording time of a single-channel EEG. Even processing EEG of an ultra-high density (e.g., 256 channels) per each channel in turn (the most inefficient setting), the inference time was only about 20% of the recording time. Also, note that the online model was already trained operating in an end-to-end manner. Apparently, *DuoCL* held great potentials in providing high-quality EEG for online EEG analysis applications.

E. Discussions

This study also evaluated the ability of the proposed model to improve signal quality at various contamination levels. Fig. 9 showed the results of the RRMSE_t before and after the hybrid artifact (EMG+EOG+BW+ECG) removal via *DuoCL*. It could be observed that the signal quality improved after reconstruction regardless of the degree of contamination. When the SNR levels were 10 dB and -10 dB, the signal quality improvement caused by reconstruction was substantially smaller than with the other three levels of SNR. The former was because of the good quality of the raw EEG, whereas the latter was because of severe signal corruption. The RRMSE_t s of the reconstruction EEG at an SNR level of -10 dB were still larger than that of the raw EEG at the SNR level of 0 dB. Consistent conclusions could be obtained from the results of the other three metrics (see Supplementary

Materials). These results indicated that when raw EEG segments were severely damaged, even if their quality could be improved by reconstruction, their quality remained low, they should be considered to be rejected. Only then was the workflow for artifact removal sufficiently complete.

Moreover, for multi-channel EEG, the evolution patterns of artifact and EEG were different in each channel. *DuoCL* operated on each channel independently without concerning correlation amongst channels, and its operation would not be interfered even all channels were reorganized as a single one. For future work, how to incorporate correlation amongst channels, i.e., spatial information, in artifact removal was an interesting issue.

V. CONCLUSION

To address the need for deep artifact removal in EEG-based neuroscience/neuro-engineering applications, particularly those with intensive unknown artifacts, this study constructed a dual-scale CNN-LSTM model operating on raw EEG through *Morphological Feature Extraction, Feature Reinforcement, and EEG Reconstruction*.

DuoCL first employed a dual-branch network with different sizes of convolution kernel to extract dual-scale deep features, with the appropriate dimension of the extracted features explored to approximate the morphological characteristics of EEG. LSTM then measured the temporal dependencies between individual samples to reinforce the dual-scale features and highlight more detailed differences between artifacts and real EEG. The resulting hybrid features formed the basis for the reconstruction of the real EEG.

Extensive experiments were conducted for evaluating the proposed model using three EEG datasets constructed with real EEG. In comparison with six state-of-the-art counterparts (e.g., 1D-ResCNN and NovelCNN): 1) *DuoCL* could reconstruct more accurate waveforms and achieve the highest SNR, CC, and lowest error (RRMSE_t & RRMSE_f); 2) at various SNR levels with different artifacts, the reconstructed EEG could best preserve the nonlinear characteristics; and 3) the model achieved satisfactory performance in the removal of unknown and hybrid artifacts and processing multi-channel EEG.

Overall, this study suggested that using deep representations of the morphological characteristics might help in tracking down the real EEG from contaminated raw EEG. The key lied with maintaining EEG's nonlinear features and highlighting the temporal dependencies.

ACKNOWLEDGMENT

We would like to show our deepest gratitude to Mr. Bingcan Lin for providing the original ideas for this work.

REFERENCES

- [1] S. R. Sreeja, R. R. Sahay, D. Samanta, and P. Mitra, "Removal of eye blink artifacts from EEG signals using sparsity," *IEEE J. Biomed. Health Informat.*, vol. 22, no. 5, pp. 1362–1372, Sep. 2018.
- [2] D. Cosandier-Rimé, J.-M. Badier, P. Chauvel, and F. Wendling, "A physiologically plausible spatio-temporal model for EEG signals recorded with intracerebral electrodes in human partial epilepsy," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 3, pp. 380–388, Mar. 2007.
- [3] J. Li, C. Li, N. Thakor, A. Cichocki, and A. Bezerianos, "A new perspective of noise removal from EEG," in *Proc. 8th Int. IEEE/EMBS Conf. Neural Eng.*, 2017, pp. 501–504.

- [4] X. Jiang, G.-B. Bian, and Z. Tian, "Removal of artifacts from EEG signals: A review," *Sensors*, vol. 19, no. 5, 2019, Art. no. 987.
- [5] G. Wang, C. Teng, K. Li, Z. Zhang, and X. Yan, "The removal of EOG artifacts from EEG signals using independent component analysis and multivariate empirical mode decomposition," *IEEE J. Biomed. Health Informat.*, vol. 20, no. 5, pp. 1301–1308, Sep. 2016.
- [6] A. G. Correa, E. Lacić, H. Patiño, and M. Valentiniuzzi, "Artifact removal from EEG signals using adaptive filters in cascade," *J. Physics: Conf. Ser.*, vol. 90, no. 1, 2007, Art. no. 012081.
- [7] K. T. Sweeney, T. E. Ward, and S. F. McLoone, "Artifact removal in physiological signals practices and possibilities," *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 3, pp. 488–500, May 2012.
- [8] P. Sawangjai et al., "EEGANet: Removal of ocular artifact from the EEG signal using generative adversarial networks," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 10, pp. 4913–4924, Oct. 2022.
- [9] G. D. Flumeri, P. Aricò, G. Borghini, A. Colosimo, and F. Babiloni, "A new regression-based method for the eye blinks artifacts correction in the EEG signal, without using any EOG channel," in *Proc. 38th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2016, pp. 3187–3190.
- [10] M. Miao, W. Hu, B. Xu, J. Zhang, J. J. P. C. Rodrigues, and V. H. C. D. Albuquerque, "Automated CCA-MWF algorithm for unsupervised identification and removal of EOG artifacts from EEG," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 8, pp. 3607–3617, Aug. 2022.
- [11] X. Chen et al., "Removal of muscle artifacts from the EEG: A review and recommendations," *IEEE Sensors J.*, vol. 19, no. 14, pp. 5353–5368, Jul. 2019.
- [12] H. Dong, D. Chen, L. Zhang, H. Ke, and X. Li, "Subject sensitive EEG discrimination with fast reconstructable CNN driven by reinforcement learning: A case study of ASD evaluation," *Neurocomputing*, vol. 449, pp. 136–145, 2021.
- [13] X. Yong, R. K. Ward, and G. E. Birch, "Artifact removal in EEG using morphological component analysis," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2009, pp. 345–348.
- [14] L. Zhang, D. Chen, P. Chen, W. Li, and X. Li, "Dual-CNN based multi-modal sleep scoring with temporal correlation driven fine-tuning," *Neurocomputing*, vol. 420, pp. 317–328, 2021.
- [15] W. Mumtaz, S. Rasheed, and A. Irfan, "Review of challenges associated with the EEG artifact removal methods," *Biomed. Signal Process. Control*, vol. 68, 2021, Art. no. 102741.
- [16] B. Wei, X. Zhao, L. Shi, L. Xu, T. Liu, and J. Zhang, "A deep learning framework with multi-perspective fusion for interictal epileptiform discharges detection in scalp electroencephalogram," *J. Neural Eng.*, vol. 18, no. 4, 2021, Art. no. 0460b3.
- [17] S. Seifzadeh, K. Faez, and M. Amiri, "Comparison of different linear filter design methods for handling ocular artifacts in brain computer interface system," *J. Comput. Robot.*, vol. 7, no. 1, pp. 51–56, 2014.
- [18] F. Bartoli and S. Cerutti, "An optimal linear filter for the reduction of noise superimposed to the EEG signal," *J. Biomed. Eng.*, vol. 5, no. 4, pp. 274–280, 1983.
- [19] G. L. Wallstrom, R. E. Kass, A. Miller, J. F. Cohn, and N. A. Fox, "Automatic correction of ocular artifacts in the EEG: A comparison of regression-based and component-based methods," *Int. J. Psychophysiol.*, vol. 53, no. 2, pp. 105–119, 2004.
- [20] A. Schlogl, C. Keinrath, D. Zimmermann, R. Scherer, R. Leeb, and G. Pfurtscheller, "A fully automated correction method of EOG artifacts in EEG recordings," *Clin. Neurophysiol.*, vol. 118, no. 1, pp. 98–104, 2007.
- [21] 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.
- [22] N. P. Castellanos and V. A. Makarov, "Recovering EEG brain signals: Artifact suppression with wavelet enhanced independent component analysis," *J. Neurosci. Methods*, vol. 158, no. 2, pp. 300–312, 2006.
- [23] N. Mamnone, F. L. Foresta, and F. C. Morabito, "Automatic artifact rejection from multichannel scalp EEG by wavelet ICA," *IEEE Sensors J.*, vol. 12, no. 3, pp. 533–542, Mar. 2012.
- [24] B. Yang, K. Duan, C. Fan, C. Hu, and J. Wang, "Automatic ocular artifacts removal in EEG using deep learning," *Biomed. Signal Process. Control*, vol. 43, pp. 148–158, May 2018.
- [25] W. Sun, Y. Su, X. Wu, and X. Wu, "A novel end-to-end 1D-ResCNN model to remove artifact from EEG signals," *Neurocomputing*, vol. 404, pp. 108–121, 2020.
- [26] H. Zhang, C. Wei, M. Zhao, Q. Liu, and H. Wu, "A novel convolutional neural network model to remove muscle artifacts from EEG," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2021, pp. 1265–1269.
- [27] P. Chen, D. Chen, L. Zhang, Y. Tang, and X. Li, "Automated sleep spindle detection with mixed EEG features," *Biomed. Signal Process. Control*, vol. 70, 2021, Art. no. 103026.
- [28] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 9, pp. 1904–1916, Sep. 2015.
- [29] E. Zerhouni, D. Lányi, M. Viana, and M. Gabrani, "Wide residual networks for mitosis detection," in *Proc. IEEE 14th Int. Symp. Biomed. Imag.*, 2017, pp. 924–928.
- [30] M. A. Klados and P. D. Bamidis, "A semi-simulated EEG/EOG dataset for the comparison of EOG artifact rejection techniques," *Data Brief*, vol. 8, pp. 1004–1006, 2016.