

Automated EOG Removal from EEG Signal Using Independent Component Analysis and Machine Learning Algorithms



Tu Thanh Do, Thuong Hoai Nguyen, Tho Anh Le, Sinh Anh Thi Nguyen, Quyen Thao Ngoc Nguyen, Thinh Quang Vu Tran, Khoi Nguyen Khac Nguyen, Khiết Thu Thi Dang, Thien Hoai Luong, Chau Minh Ngoc Trinh, Hanh Hong Cao, Hieu Quang Le Truong, Khoa Anh Tran, Nhu Huynh Vo Nguyen, Phuong Nguyen Nguyen, Quyen Hoang Quoc Vo, and Huong Thanh Thi Ha

Abstract Contamination of EEG by ocular artifacts (EOG) is the major artifact that reduces the accuracy of applications using Electroencephalogram (EEG) signal. To resolve this issue, Independent Component Analysis (ICA) is a common method to remove EOG artifacts from EEG recordings, by decomposing multichannel EEG signals into maximally Independent Components (ICs). ICs representing ocular activities can be identified visually, then be eliminated to reconstruct EOG-free EEG signals. However, this approach requires prior domain knowledge, and hence, undermine reliability and reproducibility. To address this, our study proposed a method to remove EOG contamination by applying machine learning techniques. We acquired an EEG database of 20 healthy subjects using Alice 5 Polysomnography system to record signals from 12 electrodes. Randomly selected 15-s data segments from EEG channels were used to run ICA, which resulted in 10 ICs. For each IC, we plotted its topography map and labelled whether this IC is “EOG” or “non-EOG”. A total of 612 labelled data points of ICs, topography maps and labels were collected. After applying several classifiers for model training and evaluation using cross-validation, the best classifier, Extremely Randomized Tree, achieved an average accuracy of 92%, precision of 83%, recall of 71%, and F1 score of 76%. In conclusion, the proposed method showed promising results in identifying EOG components and attenuating ocular activity on reconstructed EEG signals. Compared with existing

Tu Thanh Do, Thuong Hoai Nguyen and Tho Anh Le: These authors contributed equally to this paper.

T. T. Do · T. H. Nguyen · T. A. Le · S. A. T. Nguyen · Q. T. N. Nguyen · T. Q. V. Tran · K. N. K. Nguyen · K. T. T. Dang · T. H. Luong · C. M. N. Trinh · H. H. Cao · H. Q. Le Truong · K. A. Tran · N. H. V. Nguyen · P. N. Nguyen · Q. H. Q. Vo · H. T. T. Ha (✉)
School of Biomedical Engineering, International University, Ho Chi Minh City, Vietnam
e-mail: htthuong@hcmiu.edu.vn

Vietnam National University, Ho Chi Minh City, Vietnam

automated solutions, our proposed method only used a small number of channels and had the potential to be applied in real-time applications due to its fast computation.

Keywords Independent component analysis · ICA · Ocular artifact · EOG · Classification models · Extra Trees Classifiers

1 Introduction

EEG signal (Electroencephalogram) has represented its unique role in neuroscience, clinical engineering, psychiatry studies as well as rehabilitation engineering with its non-invasive, inexpensive high temporal resolution technique [1–3]. Compared to other neuroimaging methods such as fMRI and PET, brain electrical signals have higher temporal resolution. This advantage of EEG signal enables various studies of cognitive processes. Different from conventional diagnostic tools used in mental and psychiatric studies such as questionnaires, EEG signals are more quantitative. Specifically, robust features from EEG signals that can be used for classification of different mental states or cognitive processes mainly fall in higher frequencies ranging from 30 to 80 Hz (gamma band) [4–6].

Due to its low amplitude, the EEG signal is sensitive to various noise sources coming from biological artifacts and the environment. This problem hinders doctors and researchers from obtaining good diagnostic information without excluding valuable EEG signals. For instance, power-line artifacts, as well as biological artifacts stemming from the subject including electrical signals from muscle tension, contractions of the heart and respiration, can also contaminate EEG signals [7]. EOG artifact is the most common source of artifacts that is affecting EEG signals and overlapping frequency spectrum. More specifically, Freeman and his colleagues have demonstrated that higher frequencies, e.g. gamma band or higher, are most easily overshadowed by EOG artifacts [8]. Hence, apart from rigorous experimental design for data collection, an algorithm for EOG artifact removal is imperative in most parts of EEG studies to eliminate undesired artifacts.

Many techniques have been proposed for EOG artifacts removal. These methods can be primarily separated into two categories: either by estimation of the artifact signals using reference channels or by decomposing the EEG signal into other domains [7]. Linear Regression is a method using reference channels that assumes that each EEG channel is the sum of the non-noisy source signal and a fraction of the source artifact that is available through a reference channel(s) [9]. While regression methods are simple and reduce computational demands, they still need good regression reference channels [9]. On the other hand, the Wavelet Transform algorithm decomposes the signal into a set of coefficients, for various scales, which represent the similarity of the signal with the wavelet at that scale. Nevertheless, it fails to identify EOG signals completely that overlap with the spectral properties [9]. Another

decomposing method, Empirical mode Decomposition (EMD) is a fully data-driven method for decomposing multicomponent signals into a set of amplitude & frequency modulated (AM/FM) components known as intrinsic mode functions (IMFs) [10]. This method is sensitive to noise because it could not work effectively with multi-dimensional signals [11]. On the other hand, because mutually independent sources generate EOG artifacts such as eye movements, eye blinks, Blind Source Separation (BSS) methods, especially Independent Component Analysis (ICA) can remove EOG with great accuracy [7].

In our study, we used the MNE library implementation of ICA to decompose EEG signals to their independent components (ICs). From said components, MNE provides us with a scalp topography map of each IC. With the topography map, we were able to identify which IC represents ocular activity. However, ICA requires the need for visual inspection by experts to classify EOG and EEG. Hence, in this paper, we propose a new automatic EOG removal technique that uses ICA to decompose EEG signals into ICs then apply Machine Learning algorithms to detect EOG from these ICs. The algorithms will take the topography map of each IC as an input vector to predict whether the IC should be rejected. This new technique gives us the advantage of removing EOG artifacts without a reference channel, while requiring a low number of electrodes, and short computing time.

2 Materials and Methods

2.1 Experiment and Database

The full database of EEG signal was obtained from the Alice 5 Polysomnography system using 10 EEG channels including Fp1, Fp2, F3, F4, F7, F8, T5, T6, O1 and O2 with the ground electrodes Fpz at the forehead, M1 and M2 channels on the mastoid bones (Fig. 1).

20 Subjects were undergraduate students between the ages of 18–22 years at the time of the research study. All subjects were chosen based on exclusion criteria include (i) smokers, (ii) left-handers, (iii) native English speakers, (iv) those with a vision that was not corrected to normal, (v) antihistamine, glucocorticoid or asthma medication users, (vi) those with exposure to general anaesthesia in the last year, (vii) those with a personal or first degree family diagnosis of a DSM-IV, axis I disorder (a list of these disorders was given at the time of initial inquiry), and (viii) those with endocrine abnormalities. These exclusion criteria were self-affirmed by the prospective participants.

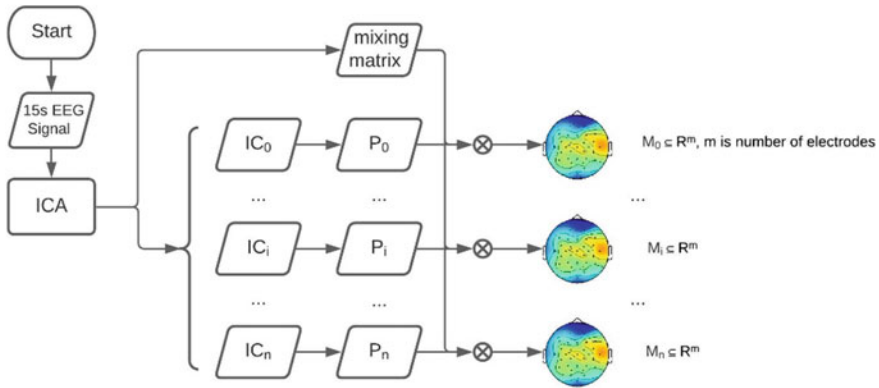


Fig. 2 Calculating the independent component’s topography map. P_i is the signal power of component i th; M_i is topography map of component i th, representing activity level of the component on the scalp

After applying ICA for each 15-s chunk of preprocessed EEG signal, we had a matrix consisting of 10 ICs time series (s) and mixing matrix (A). From the matrix of ICs, we calculated power spectrum density (PSD) for each component. From the mixing matrix, we extracted the topography map for each component, as followed Fig. 2.

2.4 Classification Models

Four supervised learning models were used for comparison: support vector machine (SVM), random forest (RF), extremely randomized trees (ExTrees) and extreme gradient boosting (XGBoost). For the former three, the implementations from the Python library sci-kit-learn [14] are used, and for the latter, there is a dedicated library called xgboost [2].

In classification tasks, true positive refers to the number of correctly classified positive points (in this case, the number of correctly classified EOGs), false positive is the number of incorrectly classified EOGs. Similarly, true negative means the number of correctly classified non-EOGs and false-negative represents the samples that are incorrectly classified as non-EOGs. The metrics used in the experiment are solely based on these four elements.

Accuracy is calculated as the fraction of the labels that exactly match the ground truth.

$$accuracy = \frac{TruePositives + TrueNegatives}{TruePositives + TrueNegatives + FalsePositives + FalseNegatives}$$

Precision is the fraction of correctly classified positives (among the samples classified as positive)

$$precision = \frac{TruePositives}{TruePositives + FalsePositives}$$

Recall (or sensitivity) is the fraction of correctly classified positives (among all the true positive samples).

$$recall = \frac{TruePositives}{TruePositives + FalseNegatives}$$

F1-score is the weighted average of precision and recall

$$F1 = 2 \times \frac{precision \times recall}{precision + recall}$$

Accuracy can be a good metric for balanced datasets (where the number of positives and negatives are roughly equal). However, the metric suffers from imbalanced classes (where the class distribution is not uniform). If a dataset has 100 samples with only 10 out of them are EOGs then a model that only predicts non-EOG for all samples would still have an accuracy of 90%, but such model would not be considered 'good' since it fails to recognize any of the EOG samples (recall = 0). In our study, there were only a few EOG samples compared to the large number of non-EOGs. This is an example of an imbalanced dataset, where precision, recall and F1-score can be particularly useful for model evaluation.

These subsections give a brief overview of the methods we used.

Support vector machine

Support vector machine (SVM) [15] is one of the commonly-used machine learning algorithms in EEG classification. In classification context, SVM tries to find a hyperplane, which can be a line in 2-dimensional space or a plane in 3-dimensional space, that maximizes the margins—the distances between the hyperplane and the closest points to such hyperplane in each class. Since the dataset we needed to classify is not linearly separable, SVM with a non-linear kernel is used to map the data into a high-dimensional space where linear separability can be obtained. In our experiment, the radial basis function (RBF) kernel was used.

Random forest

Random forest (RF) is a type of ensemble learning model. The main idea of the method is to take advantage of many decision trees, where each tree is built from a bootstrap sample (random sample drawn with replacement) taken from the data, and to build each tree with all or a random subset of variables. The randomness introduced above will help decrease variance and thus prevent model overfitting, which is one of the main drawbacks of vanilla decision trees. The random forest implementation in scikit-learn calculates the predicted output by averaging the probabilistic predictions.

Since decision trees are non-linear as there is no formal equation to express the relationship between the features and the target, the random forest is expected to be able to solve the problem of non-linearly separability of the dataset.

Extremely randomized trees

Extremely randomized trees (ExTree) was first introduced in 2006 by Pierre Geurts, Damien Ernst and Louis Wehenkel [16]. Though the algorithm is similar to the random forest, the difference between these two ensemble learning models lies in the level of randomness. In node splitting, while the random forest model tries to find the best split, ExTrees chooses the variable splitting value randomly. This can normally reduce the variance of the model even more, but at the cost of increased bias, according to the authors. Like random forest and other tree-based models, ExTrees is non-linear and is expected to solve the problem of non-linearly separability.

Extreme gradient boosting

Extreme gradient boosting (XGBoost), is a scalable implementation of the gradient boosting algorithm [17]. Gradient boosting is, like random forest and extremely randomized trees, an ensemble learning method in a sense that the predicted output will be based on an ensemble of many models. The difference between boosting and bagging, which is the technique used in random forest and extremely randomized trees, is that the bootstrap samples are weighted so that the samples with which the model incorrectly predicted get higher weights and thus be sampled more often. The idea behind weighing samples is that the model would focus more on 'difficult' samples. The gradient is used when optimizing the training loss. Hence the name gradient boosting. XGBoost further improves the original boosting method by introducing second-order gradients and regularization that help prevent overfitting.

3 Results

3.1 Preprocessing of EEG signal

Bandpass-filter with cut-off frequencies at 0.5 and 45 Hz and baseline correction were applied for each chunk of 15 s original EEG signals.

To understand the changes in raw EEG signals after our preprocessing, we compared raw EEG data (Fig. 3a) and preprocessed EEG data (Fig. 3b). The noise was reduced by the bandpass filter as indicated by the reduced thickness of the data line, especially at channels Fp1, Fp2 (Fig. 3a, b). Baseline drifts were removed in data lines after the baseline correction (Fig. 3c).

Nevertheless, the general waveforms of processed EEG recordings still kept their origins, which proved that EEG signals do not lose their representative information after the preprocessing step.

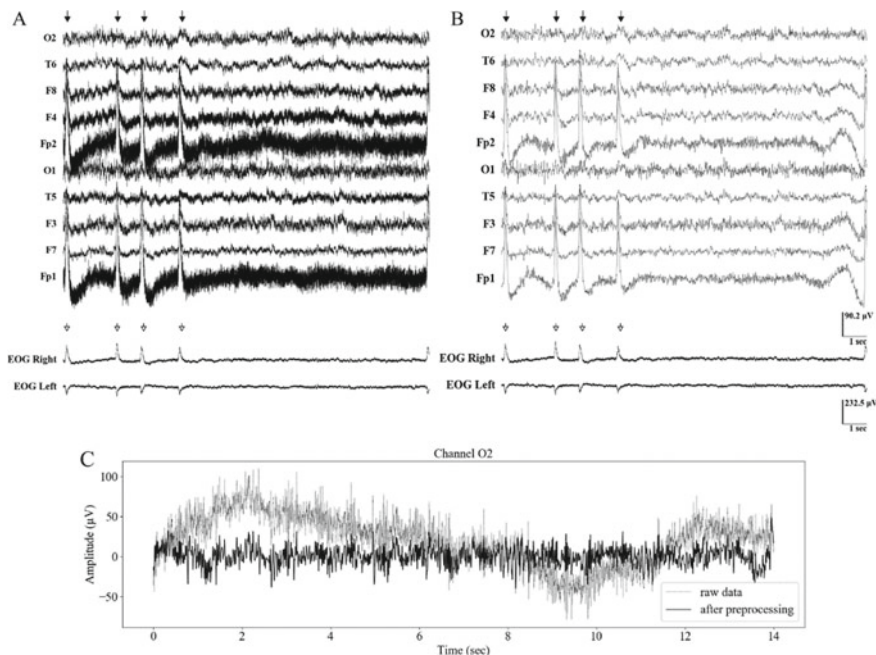


Fig. 3 Preprocessing of EEG signals. **a** A 15 s segment of raw EEG data (top) and EOG data (bottom). **b** the same segment of EEG data of figure **a** after band-pass filtering and baseline correction. Black arrows mark EOG peaks in EEG signal (above), white arrows mark EOG peaks in EOG reference’s channel. **c** Another 15 s segment of raw EEG data—channel O2 after bandpass filtering and baseline correction

3.2 Independent Component Analysis of EEG Signal

To acquire the training dataset of ICs signal, its topography map, and its label, we divided our preprocessed EEG signal into chunks of 15 s. For each chunk of 15 s preprocessed EEG signal, ICA was used to calculate a matrix consisting of 10 ICs time series (s) and mixing matrix (A). From the matrix of ICs, the power spectrum density (PSD) for each component is calculated. From the mixing matrix, we extracted the topography map for each component. Upon visual inspection of the topography map and the IC itself, ICs that represent ocular activity are labelled 1, and other ICs were labelled 0. We observed that ICA did not always successfully isolate EOG artifacts from EEG signals. For successful cases, ICs were very distinguishable from each other (Fig. 4a, c, e). In this successful case, there was one IC (ICA000) with waveform resembling EOG artifacts when comparing with EOG reference channels (Fig. 4a). Each EOG peak was marked with a black arrow for the IC and white arrow for the EOG reference channels. The topography of this IC represents activity exclusively in the frontal lobe area (Fig. 4c), which is expected for eye-derived electrical activity. From the PSD (Fig. 4e), we could see these ICs carry very little

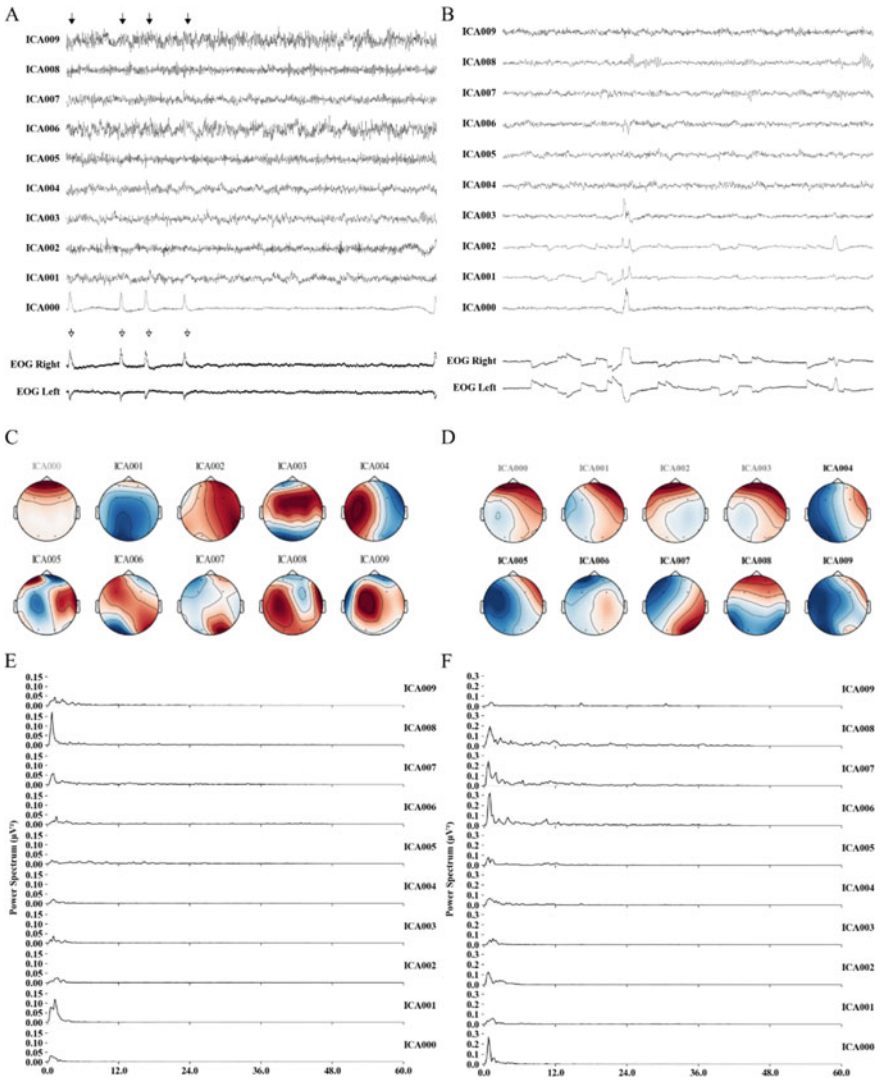


Fig. 4 Comparison of different ICA-related features. **a, b** Signal of independent components (ICs) from an example in which ICA successfully separated EOG artifact from EEG signal (**a**) or not (**b**). EOG reference channels are shown at the bottom. Black arrows mark EOG peaks in IC signal; white arrows mark EOG peak in EOG channels. **c, d** Topography maps corresponding to the ICs in A and B. **e, f** Power spectrum density plots of each ICs in A and B

bio-signal in the range 0–40 Hz. For unsuccessful cases, ICs were indistinguishable from each other. More specifically, EOG artifacts were not separated from the EEG signal and existed in several ICs (Fig. 4b).

Additionally, none of the topography exclusively represents activity in the frontal lobe area (Fig. 4d). For our training dataset, we only included cases in which ICA successfully separates EOG artifacts from the EEG signal. This training dataset was utilized for training several classifiers to detect EOG components in our ICs.

3.3 Applying Machine Learning for Automatic Removal of EOG Artifact

Once the topography map data has been successfully extracted from ICA, we obtained a dataset of 612 data points, each of which is a feature vector of raw IC features plus the map components we chose. Visually, one could notice a clear distinction between EOG and non-EOG components by looking at the topography maps of the samples. Still, we would like to find out how the learning models will perform with this particular dataset.

Figure 5 shows the data points in a 2-dimensional space. The map features were

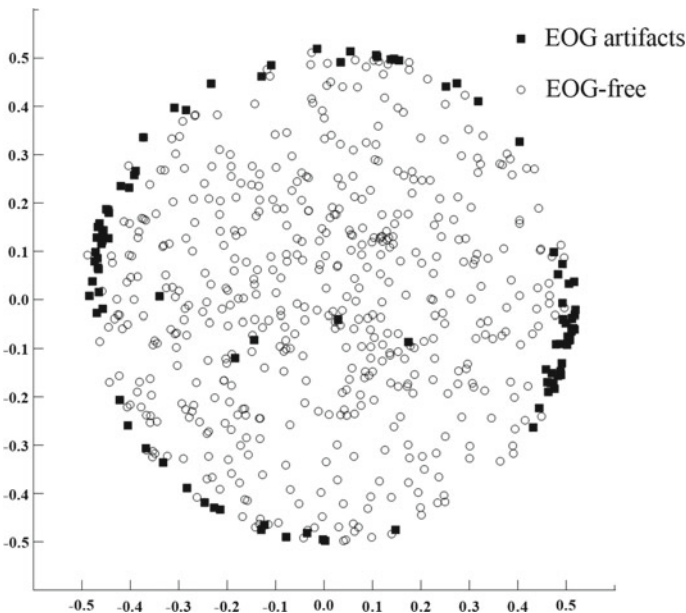


Fig. 5 Topographic map data visualized in 2-dimensional space after Principal Component Analysis (PCA). EOG components are presented as dark rectangles, and non-EOG components are presented as white circles

transformed from 10 dimensions into two dimensions using Principal Component Analysis (PCA). PCA is a widely used linear dimensionality reduction technique that aims to project multi-dimensional data into a lower-dimensional space and to retain maximum variance between data points [18]. Since Fig. 5 suggested our data is not linearly separable, we were tempted to use non-linear models for the dataset.

The data was standardized so that each component to have a mean of 0 and a standard deviation of 1 before being trained by the models. For each model, 3-fold cross-validation was used. The experiment on each model was repeated ten times with different random number generators for cross-validation splitting in order. The metrics were averaged across ten runs. The comparison boxplots for different metrics of the models with raw ICA features included along with topo map features used in training are shown in Fig. 5. The results in Fig. 5 suggested that all the models do not perform well when the raw ICA features are included along with topo map features in the training step. The best-performing model in this experiment was XGBoost with the top score in all metrics. While all models still managed to have accuracy above 0.8, only XGBoost had F1-score higher than 0.5 (0.59 ± 0.01). The rest failed to detect most EOGs, with the most extreme cases being ExTrees and SVM, which had precision, recall and F1-score of 0. We hypothesised that too many predictors as in the case of ICA features with 7500 dimensions would create the problem of high dimensionality, where the predictive power can at first increase along with more features, but then decreases when the number of observations is fixed [19].

To enhance the performance of the models, we selected another approach, which only included map features in training. From the results in Fig. 7, we observed that all models managed to have a high accuracy of over 0.9. ExTrees significantly outperformed other models in terms of F1-score and recall ($p = 0.001$ and $p = 0.022$) with the average F1-score of 0.77 ± 0.009 and the average recall of 0.71 ± 0.01 . In terms of precision, Random Forest produced the results with the highest score (0.85 ± 0.01), but the score was not significantly better than that of ExTrees (0.84 ± 0.008) ($p > 0.05$). Both SVM and XGBoost fell behind RF and ExTrees with clearer differences in the precision score.

To reconstruct EOG free signal from preprocessed EEG signals, we used ICA to decompose ten channels of EEG signal into a matrix of 10 ICs. With the trained classifier mentioned earlier, we were able to detect ICs representing EOG activity. We then proceeded to set the value of this IC in the matrix to zero. With this new matrix, we were able to inverse transform to EOG free signal [20]. Figure 8 demonstrated the result of the algorithm successfully removing EOG peaks from the signal while preserving other bio-signals. The black arrows on Fig. 8a marked EOG peaks that were removed by the algorithm. EOG-free EEG signals were shown in Fig. 8b.

In addition to evaluating the performance of EOG classification, we were also interested in investigating the computation time, which is an important factor for a scalable pipeline. We executed the pipeline from initial processing to EOG removal of a segment ten times and took the average computation time. The pipeline script was run on a laptop with 16 Gb of memory and a Core i5 processor. As Table 1 suggested, the total pipeline takes around 5 s on average, with most of the computation time being from the ICA processing step.

Table 1 The computation time of the steps in the preprocessing pipeline. Computations are executed ten times on a single machine with 16 Gb of RAM and an Intel Core i5 processor

Task	Time (s) (mean \pm SEM)
Bandpass filter	0.017 \pm 0.002
Baseline correction	0.004 \pm 0.0004
ICA	4.941 \pm 0.251
EOG classification	0.115 \pm 0.003
EOG removal	0.005 \pm 0.001
Total	5.083

4 Discussion

To summarize, the proposed approach to EOG artifact removal consists of three steps: preprocessing signal, decomposing preprocessed signals into components, and using a classifier to detect components that represent EOG activity. Firstly, baseline correction and bandpass filtering were proven to be an effective preprocessing method to remove powerline noise while preserving EEG waveforms. Secondly, the independent component analysis showed the capacity to isolate EOG artifacts from EEG signals. However, for certain cases, EOG artifact and EEG signals were still mixed in one or many ICs. And finally, several machine learning classifiers were applied to detect components representing ocular activities. However, the classifier was not yet able to detect IC with mixed signals from EOG artifact and EEG signal, which left room for improvement in the future.

With the proposed method, we could automatically remove EOG artifacts from EEG signals without the need for reference channel and domain expertise. Also, by removing the manual step of determining EOG artifacts, it was more convenient to implement an online artifact removal implementation using ICA.

In our EEG signal, the numbers of sources were larger than the number of recordings, and the EOG artifacts had high magnitude. Therefore, ICA could be applied successfully to isolate EOG artifacts from EEG signals. However, there were several shortcomings in the proposed approach. First, our classifier could not determine components with mixed EOG artifacts and EEG signals from components that include purely EOG artifacts. This results from our training process in which we only included two classes: EOG components—consisting only EOG artifact and non-EOG components—consisting only EEG signals. We excluded components with mixed EOG artifacts and EEG signals from the training dataset. Second, our approach did not offer to remove EOG artifacts from a signal channel recording of EEG and required a large resource of computing power. [21] Finally, we would like to discuss the classification techniques used to determine components representing EOG artifacts. From the results shown in Figs. 6 and 7, ExTrees gave a significantly better performance in terms of F1-score and recall. Interestingly, raw ICA features made the models fail to recognize EOG samples, hypothetically due to the problem of high dimensionality. Compared to a previous study [22] which used SVM for eye-blink artifact detection and a fourfold CV, our best classification accuracy was lower (99.3% vs. 93%). One

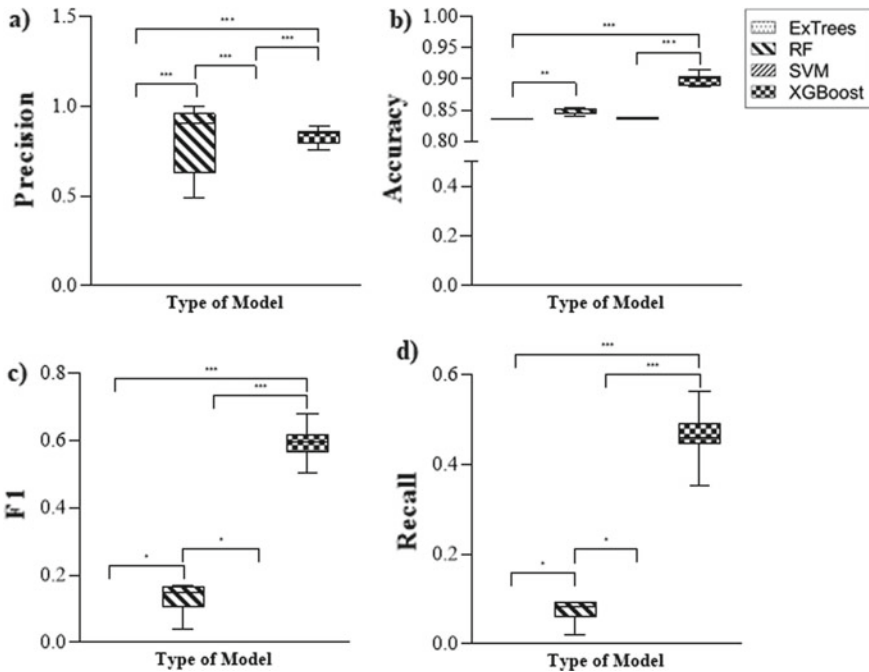


Fig. 6 Performance of different machine learning models using both signals from topography map and raw IC signals as input for training). **a** Precision. **b** Accuracy. **c** F1-score. **D**) Recall. One-way ANOVA followed by Kruskal–Wallis multiple comparisons (* $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$)

potential difference was that our study utilized an imbalance dataset while the dataset in [22] was perfectly balanced with 100 samples of each class. In another study [23] that used a similar classification approach, they managed to get high accuracy scores for eye-blink artifacts with a balanced dataset and more samples (99.39% for eye blink and 99.62% for eye movement). Given the limited number of samples and the imbalanced nature of the dataset we have, these results were encouraging.

Future Works

There exists certainly room for improvement in the aspect of F1-score by proper feature extraction for ICA data, using either statistical features (mean, median, kurtosis) or some sorts of signal transformations like discrete Fourier transform, or wavelet transform that might be able to capture the inner nature of the ICA components and the difference between EOGs and non-EOGs. Another topic that we would like to improve in the future is including mixed classes in our training dataset and curating a balanced dataset for the training. These approaches would help the classifier to determine which components consist of pure EOG artifacts and which components consist of both EOG artifacts and EEG signals and improve the accuracy of the models.

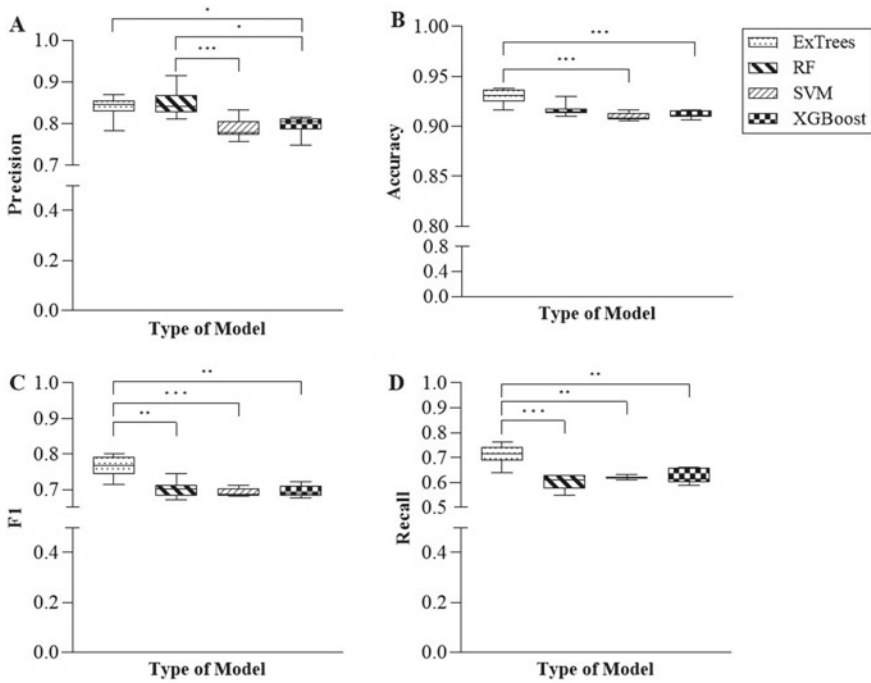


Fig. 7 Performance of different machine learning models using only the signal from topography maps as input for training. **a** Precision. **b** Accuracy. **c** F1-score. **d** Recall. One-way ANOVA followed by Kruskal–Wallis multiple comparisons (* $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$)

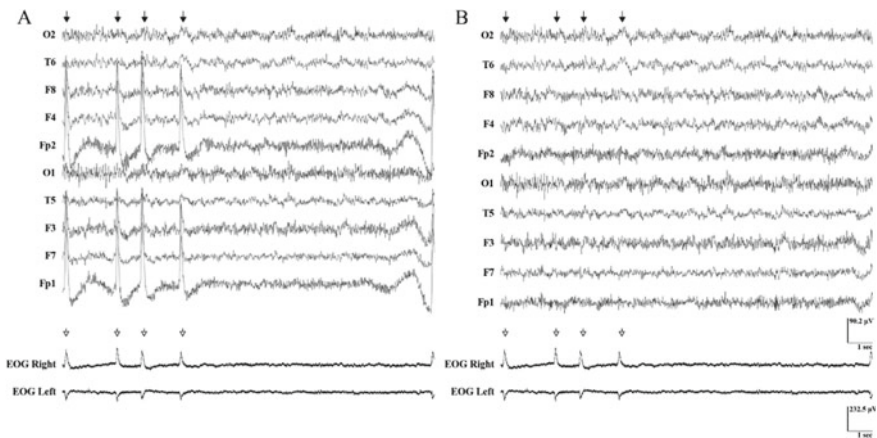


Fig. 8 Comparison of preprocessed EEG signal and EOG-free signal after using the algorithm. **a** 15 s of preprocessed EOG channels (top) and EOG channels (bottom). **b** The Same segment of EEG data after the EOG artifact was removed by the algorithm. Black arrows mark EOG peaks in the EEG channels. White arrows mark EOG peaks in the reference channels

Code deposit: https://github.com/Young1906/ica_paper

Acknowledgements This research is funded by Vietnam National University Ho Chi Minh City (VNU-HCM) under grant number C2020-28-06.

Disclosure of Potential The authors declare that they have no conflict of interest.

References

1. Light GA et al (2010) Electroencephalography (EEG) and event-related potentials (ERPs) with human participants. *Curr Protocols Neurosci* 52(1):6.25.1–6.25.24: <https://doi.org/10.1002/0471142301.ns0625s52>
2. Hughes JR, John ER (1999) Conventional and quantitative electroencephalography in psychiatry. *J Neuropsychiatry Clin Neurosci* 11(2):190–208
3. Loo SK, Makeig S (2012) Clinical utility of EEG in attention-deficit/hyperactivity disorder: a research update. *Neurotherapeutics* 9(3):569–587
4. Bucci P, Mucci A, Galderisi S (2011) Normal EEG patterns and waveforms. *Standard Electroencephalog Clin Psychiatry*, 33–57. <https://doi.org/10.1002/9780470974612.ch4>
5. Moffett SX, O'Malley SM, Man S, Hong D, Martin JV (2017) Dynamics of high frequency brain activity. *Sci Rep* 7(1). <https://doi.org/10.1038/s41598-017-15966-6>
6. Muthukumaraswamy SD (2013) High-frequency brain activity and muscle artifacts in MEG/EEG: a review and recommendations. *Front Hum Neurosci* 7:138
7. Jiang X, Bian G-B, Tian Z (2019) Removal of artifacts from EEG signals: a review. *Sensors* 19(5):987. <https://doi.org/10.3390/s19050987>
8. Freeman WJ, Burke BC, Holmes MD (2003) Aperiodic phase re-setting in scalp EEG of beta-gamma oscillations by state transitions at alpha-theta rates. *Hum Brain Mapp* 19(4):248–272
9. Urigüen JA, Garcia-Zapirain B (2015) EEG artifact removal—state-of-the-art and guidelines. *J Neural Eng* 12(3):031001. <https://doi.org/10.1088/1741-2560/12/3/031001>
10. Looney D, Li L, Rutkowski TM, Mandic DP, Cichocki A (2007) Ocular artifacts removal from EEG using EMD. In: *Advances in Cognitive Neurodynamics ICCN 2007*, pp 831–835. https://doi.org/10.1007/978-1-4020-8387-7_145
11. Xu X, Chen X, Zhang Y (2018) Removal of muscle artefacts from few-channel EEG recordings based on multivariate empirical mode decomposition and independent vector analysis. *Electron Lett* 54(14):866–868. <https://doi.org/10.1049/el.2018.0191>
12. Virtanen P et al (2020) SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods* 17(3):261–272
13. Gramfort A, Luessi M, Larson E, Engemann D, Strohmeier D, Brodbeck C, Goj R, Jas M, Brooks T, Parkkonen L, Hämäläinen M (2020) MEG and EEG data analysis with MNE. *Front Neurosci*. <https://mne.tools/dev/generated/mne.preprocessing.ICA.html>. Accessed 26 Apr 2020
14. Varoquaux G, Buitinck L, Louppe G, Grisel O, Pedregosa F, Mueller A (2015) Scikitlearn. *GetMobile: Mobile Comput Commun* 19(1):29–33. <https://doi.org/10.1145/2786984.2786995>. <http://paperpile.com/b/SPfWWe/pOxf>
15. Cortes C, Vapnik V (1995) Support-vector networks. *Mach Learn* 20(3):273–297. <https://doi.org/10.1007/bf00994018>
16. Geurts P, Ernst D, Wehenkel L (2006) Extremely randomized trees. *Mach Learn* 63(1):3–42. <https://doi.org/10.1007/s10994-006-6226-1>
17. Chen T, Guestrin C (2016) XGBoost. In: *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining—KDD '16*. <https://doi.org/10.1145/2939672.2939785>

18. Jolliffe IT (2013) *Principal component analysis*. Springer Science & Business Media
19. Trunk GV (1979) A problem of dimensionality: a simple example. *IEEE Trans Pattern Anal Mach Intell* 1(3):306–307
20. Djuwari D, Kant Kumar D, Palaniswami M (2005) Limitations of ICA for artefact removal. *Conf Proc IEEE Eng Med Biol Soc* 2005:4685–4688
21. Nguyen H-AT et al (2012) EOG artifact removal using a wavelet neural network. *Neurocomputing* 97:374–389. <https://doi.org/10.1016/j.neucom.2012.04.016>
22. Shoker L, Sanei S, Chambers J (2005) Artifact removal from electroencephalograms using a hybrid BSS-SVM algorithm. *IEEE Signal Process Lett* 12(10):721–724. <https://doi.org/10.1109/lsp.2005.855539>
23. Halder S et al (2007) Online artifact removal for brain-computer interfaces using support vector machines and blind source separation. *Comput Intell Neurosci*, 82069