

## Automatic ECG Artefact Removal from EEG Signals

Mohamed F. Issa<sup>1,2</sup>, Gergely Tuboly<sup>1</sup>, György Kozmann<sup>1</sup>, Zoltan Juhasz<sup>1</sup>

<sup>1</sup>Department of Electrical Engineering and Information Systems, Faculty of Information Technology, University of Pannonia, Egyetem u.10, 8200, Veszprém, Hungary, mohamed.issa@virt.uni-pannon.hu

<sup>2</sup>Department of Scientific Computing, Faculty of Computers and Informatics, Benha University, Fared Nada, 13511, Benha, Egypt

Electroencephalography (EEG) signals are frequently contaminated by ocular, muscle, and cardiac artefacts whose removal normally requires manual inspection or the use of reference channels (EOG, EMG, ECG). We present a novel, fully automatic method for the detection and removal of ECG artefacts that works without a reference ECG channel. Independent Component Analysis (ICA) is applied to the measured data and the independent components are examined for the presence of QRS waveforms using an adaptive threshold-based QRS detection algorithm. Detected peaks are subsequently classified by a rule-based classifier as ECG or non-ECG components. Components manifesting ECG activity are marked for removal, and then the artefact-free signal is reconstructed by removing these components before performing the inverse ICA. The performance of the proposed method is evaluated on a number of EEG datasets and compared to results reported in the literature. The average sensitivity of our ECG artefact removal method is above 99 %, which is better than known literature results.

Keywords: EEG, ECG artefact removal, Independent Component Analysis, QRS detection, cardiac cycle classification.

### 1. INTRODUCTION

Electroencephalography (EEG) is a non-invasive and cost-effective measurement technology capable of recording the bioelectrical activity of the brain with sub-millisecond resolution. This makes EEG essential in epilepsy diagnosis, cognitive or sensorimotor experiments, where rapid activity changes must be examined. Unfortunately, the measured EEG signals are regularly contaminated by equipment and environmental noise as well as artefacts caused by extracerebral physiological sources. Among the latter types, ocular, muscle, and cardiac artefacts are especially problematic due to their high amplitude and non-periodic (ocular, muscle) or quasi-periodic (cardiac) nature; they can easily turn valuable EEG measurements unusable. Consequently, artefact removal is a key step in every EEG processing pipeline.

Traditional artefact removal requires visual inspection and manual rejection of artefact contaminated data segments or epochs. This approach is laborious, requires trained personnel, can largely reduce the number of usable epochs, and prevents the automatic and high-speed analysis of large-scale EEG experiments. In this paper, we propose a fully automatic method for removing ECG artefacts from EEG signals. Our approach is based on using Independent Component Analysis (ICA) that is able to separate the observed signals into statistically independent source

components, some of which may be attributed to artefact sources. ECG-related independent components are then classified using an ECG detection method, and these identified components are removed from the component set to reconstruct the EEG data with ECG artefacts removed. A sophisticated classification method ensures that only components that reflect real ECG activity are rejected. The main advantage of our method is that (i) it is fully automatic, (ii) it does not require a reference ECG channel, thus can be used in situations where ECG data is not available, and (iii) it can also detect and remove ECG artefacts generated by pathological cardiac activities which can make the method more robust when analysing EEGs of elderly patients.

#### A. Related work

Unlike ocular and muscle artefacts, which can be eliminated or at least reduced with careful experiment control – subjects asked to minimize movement, refrain from blinking or eye movements – involuntary cardiac activity and its effects are always present in EEG signals. From the EEG analysis point of view, ECG is seen as a quasi-periodic noise, whose frequency spectrum overlaps with the EEG spectrum and has a relatively high power that causes distortions in the original EEG signal [1]. While cardiac artefacts might present only a moderate contamination when we are interested in averaged event related potentials (ERP), in single-trial analysis,

epilepsy related inter-ictal spike analysis or in cardiac cycle-related EEG averaging, the effect of ECG can distort the results. Consequently, ECG artefact removal is crucial in producing clean EEG signals, yet, it is a relatively underrepresented area in artefact removal research.

Early ECG removal attempts included subtraction and ensemble average subtraction (EAS) [2] methods. Current mainstream methods are based on adaptive filtering [3], [4], blind source separation (such as ICA) [5], [6] or wavelet decomposition [7], [8] methods, although wavelets are increasingly more often used in combination with ICA-based methods [9], [10]. Since visual inspection is slow, tiring and requires an expert assistant, several authors proposed methods for semi or fully automatic component detection.

One semi-automated method using ICA for identifying artefact components is presented by Delorme *et al.* in [11]. Various statistical measures (entropy, kurtosis, spatial kurtosis) are calculated for each independent component to label them automatically as artefact but validation and rejection is performed manually. Since the method is based on statistical analysis of the components, which does not take into account the physiological model of artefacts, the performance of the method is not perfect. A similar statistical approach is followed in the FASTER [12] and ADJUST [13] artefact removal toolboxes.

The works of Dora and Biswal [7], and Jiang *et al.* [8] use wavelet decomposition-based ECG detection methods. In both cases, the Continuous Wavelet Transform is used to detect QRS waves in the EEG signal. In the first case, a reference ECG channel and linear regression are used to remove the detected QRS waves. Reported sensitivity varies between 91 and 100 % depending on the input dataset, providing lower values in more difficult cases, such as sinus arrhythmia. In the second case, no reference channel is used; the detected artefact signal (wavelet coefficients) is subtracted from the original signal to obtain the clean one. Although the reported detection performance of this method is above 97.5 %, the method ignores the removal of the P and T waves that may also contaminate EEG data.

Hamaneh *et al.* [9] use an automatic ICA-based approach. A reference spatial distribution template of the ECG artefact [14] and a Continuous Wavelet Transform-based periodicity test are used in combination to identify ECG independent components. If a component shows correlation with the spatial template (threshold > 0.6) and passes the wavelet periodicity test, it is marked for removal. While the method provides good true detection rate (95-99 % depending on the ECG contamination rate), the required spatial ECG component template has to be created by averaging manually-selected ECG components of several subjects, which reduces the level of possible automation. Mak *et al.* [10] propose an automatic ECG removal method for EMG (Electromyography) signal cleaning. Similarly to others, the wavelet transform is used to detect R peaks, after which a set of decision rules are applied to the candidate component (checking heart rate, RR interval, variance of RR interval) to detect the ECG component. Although the method is developed for cleaning trunk muscle signals instead of EEG, the reported excellent ECG detection sensitivity (100 %)

makes it worth mentioning. Unfortunately, there are no testing results for detecting pathological ECG artefacts.



Fig.1. Annotated measurement from the MIT-BIH Polysomnographic database showing ECG, Blood Pressure and EEG signals. Note the pronounced ECG artefact contamination on the EEG channel (data record: [https://physionet.org – slp/slp32](https://physionet.org/~slp/slp32)).

### B. Independent Component Analysis

Independent Component Analysis (ICA) was originally developed to solve the Blind Source Separation (BSS) problem [5] and normally refers to a class of algorithms that can recover statistically independent signals (components) from a linear mixture, based on higher-order statistics as a measure of independence. ICA is considered a robust method for identifying and removing artefacts normally found in EEG signals.

A brief formal introduction is as follows. Let us assume that we have  $N$  statistically independent sources,  $s_i(t)$ ,  $i = 1, \dots, N$ . Suppose that the sources cannot be observed directly, only via  $N$  sensors that obtain  $N$  observation signals,  $x(t)$ . The observed signals are mixtures of the original sources. Sensors must be spatially separated (e.g. as the electrodes on the scalp), as each sensor must measure a mixture different from the others. The mixing process than can be described as

$$\mathbf{x}_t = \mathbf{A}\mathbf{s}_t \quad (1)$$

where  $\mathbf{A}$  is the unknown square mixing matrix. An “unmixing matrix”  $\mathbf{W} = \mathbf{A}^{-1}$  must be obtained in order to obtain an estimate  $\hat{\mathbf{s}}_t$  of the original source as

$$\hat{\mathbf{s}}_t = \mathbf{W}\mathbf{x}_t. \quad (2)$$

The following restrictions apply to ICA in order to produce a solution: (i) the sources must be statistically independent, (ii) the sources cannot have Gaussian distribution, and (iii) the mixing matrix must be invertible. The estimation of  $\hat{\mathbf{s}}_t$  requires pre-processing steps (dimensionality reduction, centering, and uncorrelation).

The ICA method has few important limitations, however. The order and the energy of the components can be arbitrary; these are referred to as permutation and scale ambiguity, respectively. A consequence of these ambiguities is that most often ICA components are examined manually, and

component types (EOG, ECG, EMG, etc.) are identified visually.

Several ICA algorithms (e.g. Infomax ICA, Fast-ICA, JADE, SOBI) exist that differ, e.g. in the way independence is computed, in their convergence properties, the quality of source separation or their runtime complexity [15]. For a more detailed description of the theoretical foundations of ICA and the various ICA algorithm variants, the interested reader is referred to the literature [5], [6], [15]-[19].

## 2. SUBJECT & METHODS

### A. Materials

Multiple EEG datasets were selected for testing our method. Public EEG datasets used in similar studies [7], [8] were included for performance comparison purposes. We have also used resting state EEG data measured by our group on healthy volunteers who all had given their written consent in participating in the experiments.

#### PhysioNet EEG datasets

a) The MIT-BIH Arrhythmia Database [20] contains 48 half-hour excerpts of two-channel ambulatory ECG recordings recorded at the Beth Israel Hospital in Boston. Sampling frequency is 360 Hz.

b) The MIT-BIH Polysomnographic Database [20]-[22] contains sleep measurements of varying duration (ranging from 1:17 to 6:30 hours) from 16 patients monitored in the Beth Israel Hospital Sleep Laboratory in Boston. The datasets contain one EEG channel. The sampling rate of the measurements is 250 Hz.

c) The CAP Sleep Database is a collection of 108 polysomnographic recordings measured at the Sleep Disorders Center of the Ospedale Maggiore of Parma, Italy [20], [21]. Each dataset contains at least three EEG channels as well as ECG, EOG, respiration, etc. physiological signals. The sampling rate of the measurements is 250 Hz.

#### Resting state EEG dataset

Closed and open eye resting state EEG data were recorded from 61 adult volunteers (males and females, between ages 17 and 35) of 2-3 minute's duration. During the experiment, subjects had to sit and relax in a silent room. Data was recorded using a Biosemi ActiveTwo EEG system ( $f_s = 2048$  Hz) using the 128-channel ABC radial electrode layout. The volunteers gave their written consent to participating in the experiments.

#### Analysis software

All analyses were carried out in the Fieldtrip toolbox [22].

### B. Methods

The flow chart of the proposed method that includes signal pre-processing, independent component analysis and subsequently, component checks for ECG presence and artefact removal is shown in Fig.2. Each step of the method is described in detail in the following subsections.

#### Pre-processing

Signals of each dataset were filtered with a 1 – 47 Hz 4<sup>th</sup>-order zero-phase Butterworth bandpass filter to remove the DC component, slow drifts, line noise and unwanted high-frequency components. The resting state measurements were then downsampled to  $f_s = 256$  Hz. This optional step was chosen in order to reduce the execution time of the subsequent Independent Component Analysis. Average reference was used for the resting state dataset.

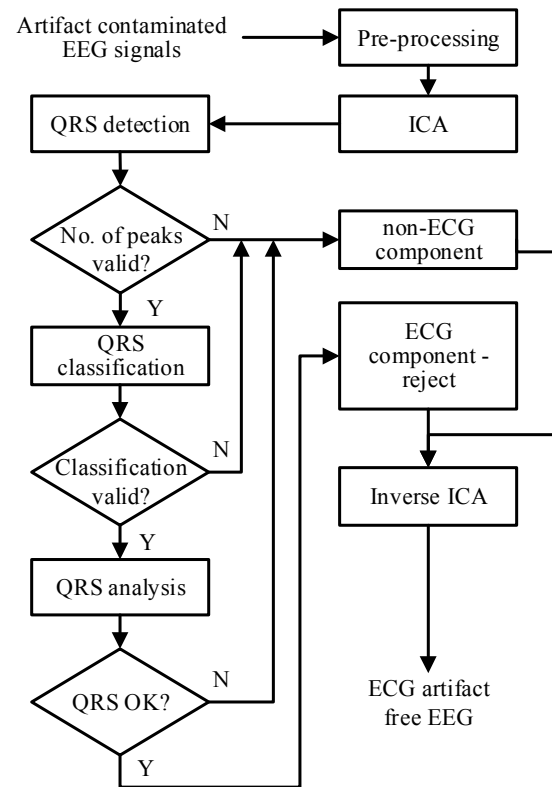


Fig.2. The flow chart of the ECG artefact removal method.

#### Independent component analysis

The pre-processed signals were partitioned into 20-second long non-overlapping segments. The Infomax ICA algorithm was performed on each segment to generate components

#### ECG component detection

The output of the ICA algorithm is a set of independent components,  $c_i$ ,  $i = 1, \dots, N$ , where  $N$  is the number of components that is also equal to the number of EEG channels. Since the order of the components produced by the ICA algorithm is arbitrary, we cannot pre-select components based on a-priori information; each component has to be examined for ECG-like activity. The underlying assumption is that an ECG independent component is similar to a real ECG signal in terms of its QRS interval, general waveform morphology and periodicity (see Fig.3. as an example). Since the most characteristic feature of an ECG signal is the QRS complex, the presence of this is used for identifying an ECG artefact component.

The complete ECG component detection hence involves the following two main stages: first, an automatic QRS detection is performed on an independent component,  $c_i$ , and then an ECG-cycle classifier decides whether the given component is in fact an ECG artefact component. These stages include several internal processing steps which we describe in detail in the following paragraphs.

#### Step 1 – Amplitude range transformation

Since the output components of the ICA algorithm may have arbitrary scale and polarity, the first step is to scale the component signal into the typical amplitude range of the ECG signals.

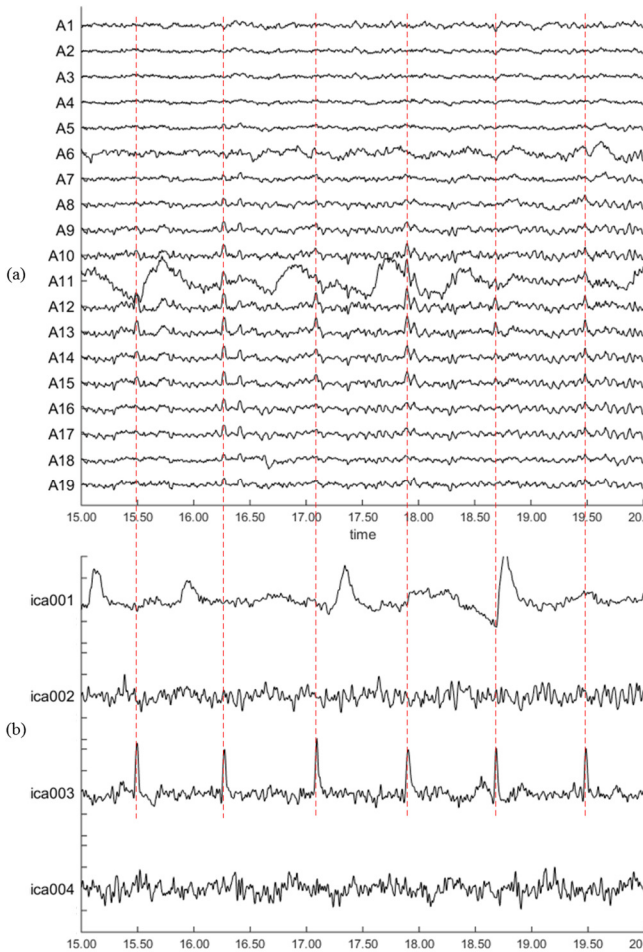


Fig.3. Input signals of a 128-channel EEG signal data segment contaminated with artefacts (a), and a subset of the resulting independent components representing EEG signals (ica002, ica004), ocular (ica001), and cardiac artefacts (ica003) (b).

The component signal,  $c_i$ , is segmented into  $K$  consecutive, non-overlapping, two-second data segments,  $s_j$ ,  $j = 1, \dots, K$ . Then, the local maximum,  $m_j = \max_{y_k \in s_j} (y_k)$ , is calculated in each segment, where  $y_k$  is the vector of samples of segment  $s_j$ ,  $k = 1, \dots, L$ , and  $L = 2fs$ . Once the local maxima of all segments are determined, their median is calculated,  $med = \text{median}(m_j)$ , and finally, the component signal is

transformed into the  $\pm 700 \mu\text{V}$  range that is required for the QRS detection algorithm:

$$\hat{y}_k = \frac{700}{med} y_k.$$

#### Step 2 – QRS detection

The next step of the method is scanning the independent components for the presence of ECG waveforms, i.e. QRS complexes. In our example in Fig. 3, component ica003 shows ECG features. The QRS detection step uses an adaptive threshold-based R-peak detection algorithm developed by Christov [23]. The algorithm operates on the derivative of the given component signal  $c(t)$ . Let  $y(t)$  denote the absolute value of the derivate,

$$y(t) = |c'(t)| \approx \left| \frac{c(t+h) - c(t-h)}{2h} \right| = \frac{|c_{k+1} - c_{k-1}|}{2\Delta t}, \quad (3)$$

where  $c_{k+1}$  and  $c_{k-1}$  are the  $(k+1)^{\text{th}}$  and  $(k-1)^{\text{th}}$  samples of the given data segment  $s_j$  of the current component.

A combined adaptive threshold function

$$MFR = M + F + R, \quad (4)$$

is calculated for each time instant using: (i)  $M$  (the steep slope threshold), (ii)  $F$  (the integrative threshold for high frequency signal components, and (iii)  $R$  (the beat expectation threshold). The exact rules for calculating the adaptive threshold can be found in [23].

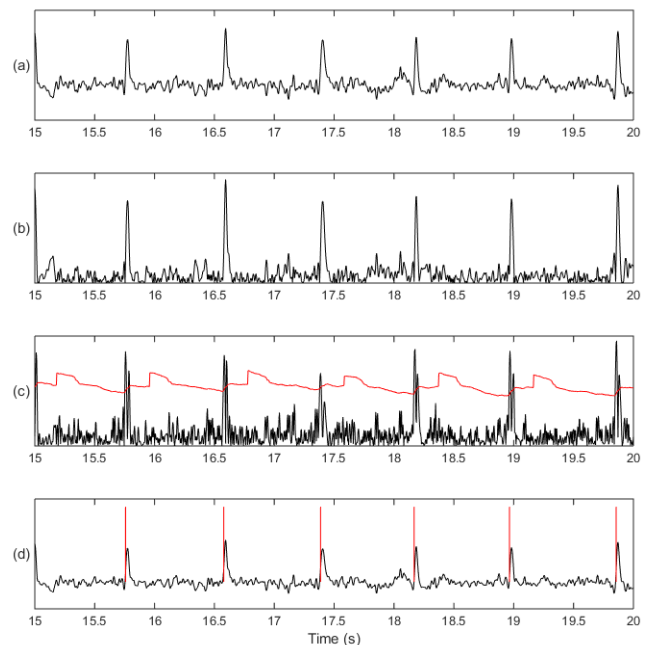


Fig.4. Successive steps of the ECG artefact detection process; a) ECG independent component, b) absolute value of the component, c) absolute value of the component derivative and the MFR adaptive threshold (solid red line), d) QRS peaks detected.



Each derivate sample  $y_k$  is compared to the *MFR* threshold, and the position of the first sample for which the condition  $y_k > MFR$  holds is stored as an R-peak position. Since the algorithm may not always pick the true position of the R-peak, a local peak search is performed subsequently within the neighbourhood of each detected position to find the global maximum in the window centred on the initial R-peak position. The positions of the final detected peaks are then stored for the next (classification) stage of our method.

### Step 3 – Cardiac cycle classification

The goal of the classification stage is to verify whether the detected peaks can be attributed to cardiac activity. If the peaks do not show characteristic ECG properties (periodicity, QRS distance and amplitude), the component is labelled as a non-ECG component. The details of the classification process are described below.

1) The first ECG detection criterion is related to the number of detected peaks; if it is outside the normal human heart rate range (<30 or >250 beats/minute), the algorithm skips the component (i.e. marks it as non-ECG).

2) The next step is the classification of the cardiac cycles,  $\mathbf{h}_k$ , into a majority heart cycle class and possible extra classes (e.g. low-quality majority cycles, extreme-amplitude artefacts). Using the detected R-peak positions as synchronization points, an average ECG waveform,  $\mathbf{h}_{avg}$ , is generated by defining a -300 ms to 400 ms window around each candidate peak, and the corresponding samples are averaged point-by-point. The generated averaged ECG will serve as a reference waveform in the classification of each cardiac cycle.

After this step, the Pearson-correlation is calculated between the average baseline ECG,  $\mathbf{h}_{avg}^{QRS}$ , and each interval waveform,  $\mathbf{h}_k^{QRS}$ , using a narrower QRS [-60ms, 80ms] window. If the correlation and the amplitude of the waveform are above the pre-determined threshold values, the interval is assigned to the majority ECG class  $C_{ECG}$ . The following formula defines the rules more formally:

$$C_{ECG} = \{\mathbf{h}_k | k \leq H, \text{corr}(\mathbf{h}_{avg}^{QRS}, \mathbf{h}_k^{QRS}) \geq 0.7 \text{ and } |\max(|\mathbf{h}_{avg}^{QRS}|) - \max(|\mathbf{h}_k^{QRS}|)| < 0.5 * \max(|\mathbf{h}_{avg}^{QRS}|)\} \quad (5)$$

where  $\mathbf{h}_k$  is the sample vector of the  $k^{th}$  detected cardiac cycle in the ICA component segment under test,  $H$  is the number of detected cycles,  $\mathbf{h}_{avg}^{QRS}$  is the vector of samples of the averaged QRS cycles [-60 ms, 80 ms], and  $\mathbf{h}_k^{QRS}$  is the vector corresponding to the same window of cardiac cycle  $k$ ,  $k = 1, \dots, H$ .

3) Next, the majority class is examined for consistency and beat periodicity. If there are too few ECG cycles in the majority class (less than 10 % of the total number of detected cycles) or the detected heart rate in the class is outside the valid human heart rate (<30 or >250 beats/minute), the component is not considered as ECG artefact.

4) If the majority class test succeeds, the final verification step is based on the average QRS interval. ECG cycles in the majority class are averaged, then the QRS interval is

calculated after locating the QRS onset and offset of the averaged cycle. If the QRS interval is too narrow or too wide (<30 or >200 ms), the majority class – and consequently the current component – is not classified as an ECG artefact. The component is marked as ECG if and only if it was not rejected in any of the preceding steps.

Since our method relies heavily on cycle averaging, it is crucial to use a sufficient number of cycles in order to obtain a high-quality averaged cycle. We recommend the use of segments with duration of 20 seconds or more.

### Component removal and inverse ICA

The final step of the method is the reconstruction of the signal from its components. The rejected independent components are removed from the component set,  $\hat{\mathbf{s}}$ , creating an artefact free set,  $\hat{\mathbf{s}}^{af}$ , by zeroing out the rejected component samples,  $\hat{\mathbf{s}} \rightarrow \hat{\mathbf{s}}^{af}$ , then the estimate of the cleaned observed EEG signal can be computed as

$$\hat{\mathbf{x}}_t = W^{-1} \hat{\mathbf{s}}_t^{af} \quad (6)$$

Once the ECG classifier identifies an ECG independent component, the entire ECG component waveform (not just the QRS complex) is rejected from the set of components. Fig.5. illustrates the result of the cleaning method, whereas Fig.6. compares the original, contaminated channel A13 of Fig.3.a) with the one after artefact removal. Note how the ECG peaks are removed from the signal without introducing any additional distortion. A different view of the cleaning effect is shown in Fig.7., which shows the scalp potential map of a QRS-peak interval before and after artefact removal. The original contaminated map clearly shows the typical spatial distribution pattern of an ECG artefact. The artefact free map illustrates to what extent the ECG artefact concealed the underlying resting state activity.

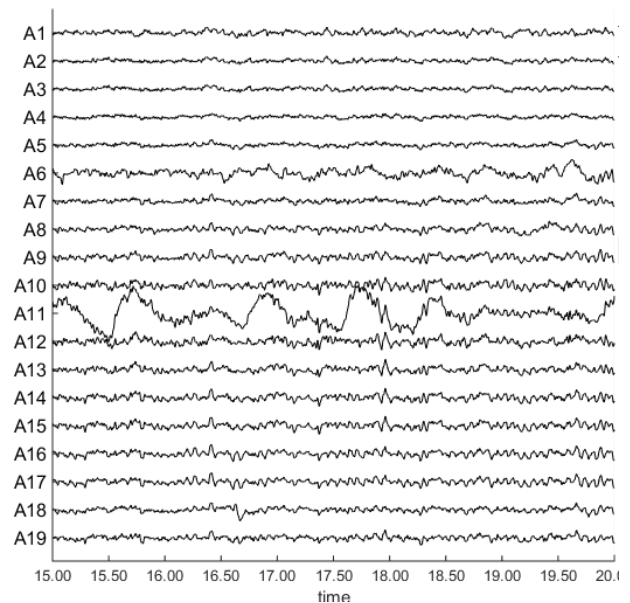


Fig.5. EEG signals with the ECG artefact removed. Compare channels A12-A15 with the contaminated originals in Fig.3.a).

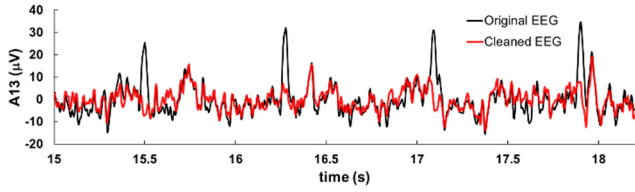


Fig.6. The original (black) and cleaned (red) samples of channel A13 of Fig.3.a). Note the four removed QRS peaks.

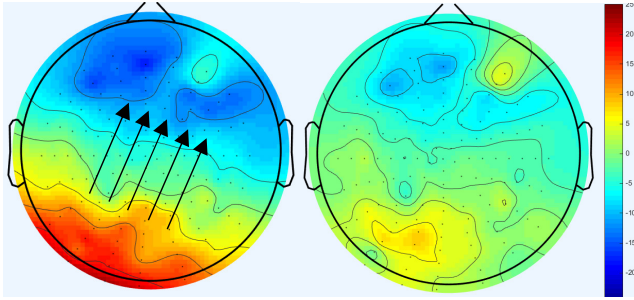


Fig.7. The scalp potential distribution of the averaged QRS peak before (left) and after (right) ECG artefact removal. Note the superimposed left-occipital-right-frontal ECG potential field (marked by the black arrows) disappearing after cleaning.

### 3. RESULTS

Our proposed method was tested on publicly available datasets and on resting state EEG data obtained in our laboratory. The public datasets we selected are the MIT-BIH Arrhythmia, the MIT-BIH Polysomnographic, and the CAP Sleep datasets. These allow our method to be compared with results reported in the literature [7], [8]. The outcome of these comparisons can be found in Table 1. to Table 3.

The overall performance of our proposed method depends on the performance of the QRS detector and the ECG component classifier. Both are examined in terms of their sensitivity, Sen, and specificity, Spe.

#### A. Artefact detection performance metrics

The sensitivity of the QRS detector is calculated as  $Sen = TP / (TP + FN)$ , where  $TP$  is the number of true positive (accurately detected), whereas  $FN$  is the number of false negative (missed) QRS peaks. We cannot use the standard formula for the Specificity,  $Spe = TN / (TN + FP)$  for the QRS detection tests, as each input signal contained QRS complexes, consequently, the true negative case  $TN$  is undefined. Instead, we use the alternative formulation  $Spe^* = TP / (TP + FP)$  as suggested in [23].

The performance of the classifier is also measured by the sensitivity and specificity. In this case,  $TP$  is the number of true ECG independent component segments,  $TN$  is the number of non-ECG component segments,  $FN$  is the number of true ECG component segments rejected by the classifier rules (amplitude, periodicity, QRS interval, number of cycles

in the majority beat class), and  $FP$  is the number of false positive segments (falsely detected QRS segments). The specificity of the component classifier is calculated using the traditional formula,  $Spe = TN / (TN + FP)$ .

#### B. QRS detector performance

Table 1. shows the results of our QRS detection method performed on ECG signals of the MIT-BIH Arrhythmia database using one and five-minute long data segments.

Table 1. QRS detection sensitivity and specificity, proposed method (pm, 1- and 5-minute segments) vs literature: MIT-BIH Arrhythmia dataset – ECG signal.

Data set	Sen (%)			Spe* (%)	
	Dora [7]	pm 1m	pm 5m	pm 1m	pm 5m
100m	97.1	100.0	100.0	100.0	100.0
101m	100.0	100.0	99.7	100.0	99.7
103m	100.0	100.0	100.0	100.0	100.0
106m	91.5	100.0	96.6	100.0	96.7
107m	100.0	100.0	100.0	100.0	100.0
117m	100.0	100.0	100.0	100.0	100.0
118m	100.0	100.0	100.0	100.0	100.0
208m	89.4	98.9	97.5	99.0	97.5
223m	92.2	100.0	100.0	100.0	100.0
231m	100.0	100.0	99.3	100.0	99.3
<b>avg.</b>	<b>97.0</b>	<b>99.9</b>	<b>99.3</b>	<b>99.8</b>	<b>99.0</b>

Table 2. and Table 3. show the results of our QRS detection method applied to the MIT-BIH Polysomnographic and the CAP sleep datasets. For both datasets, we ran the Infomax ICA to calculate the independent components. From these, we manually selected the ECG component and used this as the input to the QRS detector.

Table 2. QRS detection sensitivity and specificity, proposed method (pm, 1-minute segments) vs literature: MIT-BIH Polysomnographic Dataset – EEG signal.

Dataset	Sen (%)			Spe* (%)
	Dora [7]	Jiang [8]	pm	pm
slp01a	98.5	98.9	100.0	100.0
slp02b	95.2	98.1	100.0	100.0
slp03	97.1	96.3	100.0	100.0
slp16	100.0	-	100.0	100.0
slp32	100.0	-	100.0	100.0
slp41	100.0	-	100.0	100.0
slp59	100.0	-	100.0	100.0
slp60	100.0	-	100.0	100.0
slp66	100.0	-	100.0	100.0
slp67x	97.9	-	100.0	100.0
<b>average</b>	<b>98.7</b>	<b>97.8</b>	<b>100.0</b>	<b>100.0</b>

Table 3. QRS detection sensitivity and specificity, proposed method (**pm**, 1-minute segments) vs literature: CAP Sleep Dataset – EEG signal.

Dataset	Sen (%)		Spe* (%)
	Dora [7]	pm	pm
ins 2	100.0	100.0	100.0
ins 3	100.0	98.6	100.0
ins 5	98.9	100.0	100.0
n2	98.7	100.0	100.0
n8	98.4	100.0	100.0
nfle15	98.6	100.0	98.6
nfle35	96.5	100.0	100.0
plm3	96.7	100.0	100.0
plm4	97.0	100.0	100.0
plm9	100.0	100.0	100.0
<b>average</b>	<b>98.5</b>	<b>99.9</b>	<b>99.9</b>

### C. ECG component classifier performance

This section shows the results of our proposed ECG component classification method on an arbitrary EEG dataset in automatic mode. For the tests, seven subjects were selected at random from a 61-subject 128-channel closed-eye resting state EEG dataset.

Table 4. lists the performance results obtained on the recordings. For subjects s1 and s2, no sensitivity results could be calculated, since no ECG contamination was detectable in the datasets. Correctly, our method did not find any QRS complexes in the components, and consequently none of the independent components were classified as ECG, meaning that  $TP = FP = 0$ .

Table 4. The ECG artefact detection performance of our proposed method on 128-channel resting state EEG.

Dataset	Proposed method		
	QRS detection Sen (%)	Classifier Sen (%)	Classifier Spe (%)
s1	N/A	N/A	100.00
s2	N/A	N/A	100.00
s3	99.11	100.00	100.00
s4	100.00	100.00	100.00
s11	99.61	100.00	100.00
s24	99.40	100.00	100.00
s25	92.81	100.00	99.61
<b>average</b>	<b>98.19</b>	<b>100.00</b>	<b>99.94</b>

## 4. CONCLUSIONS

ECG artefacts present in EEG signals can lead to analysis errors in single-trial experiments, epileptiform signal or cardiac cycle-related brain activity analysis. This paper proposed a fully automatic ECG artefact removal method working without human assistance or reference ECG channel, which can be used in high-throughput, high-speed EEG analysis, continuous monitoring or clinical diagnostic systems.

The acquired EEG signals are subjected to independent component analysis and the resulting independent components are examined for cardiac activity characteristics. The applied adaptive threshold-based QRS detector and subsequent rule-based cardiac cycle classifier identify ECG activity and mark component segments for rejection with high reliability.

In QRS detection, the proposed method achieves sensitivity above 99.3 % on the PhysioNet datasets (specificity > 99 %), higher than all known automatic methods reported in literature. For our own high-density resting state EEG data, the QRS detection sensitivity is above 98.1 %, however, the sensitivity of the ECG component classifier (the complete process) is 100 %. This is due to the fact that the classifier does not need all the component QRS peaks to identify a component segment as ECG.

The significance of our method is that due to its excellent sensitivity and specificity, it can be used reliably for automatic, unsupervised artefact removal, where similar reported methods might incorrectly remove non-artefacts or leave contaminating components in the dataset.

We hope that our method advances the current practice of ECG artefact removal, and due to its clear advantages, i.e. the fully automatic operation, better sensitivity than previous approaches, and the capability of detecting pathological ECG waveforms, such as frequent ventricular ectopic beats or bundle branch blocks, it will help practitioners in producing more accurate analysis results.

## ACKNOWLEDGMENT

This work has been supported jointly by the Hungarian Government and the European Social Fund [grant number EFOP-3.6.1-16-2016-00015].

## REFERENCES

- [1] Dirlich, G., Vogl, L., Plaschke, M., Strian, F. (1997). Cardiac field effects on the EEG. *Electroencephalography and Clinical Neurophysiology*, 102 (4), 307-315.
- [2] Park, H.J., Jeong, D.U., Park, K.S. (2002). Automated detection and elimination of periodic ECG artefacts in EEG using the energy interval histogram method. *IEEE Transactions on Biomedical Engineering*, 49 (12), 1526-1533.
- [3] Suresh, H.N., Puttamadappa, C. (2008). Removal of EMG and ECG artifacts from EEG based on real time recurrent learning algorithm. *International Journal of Physical Sciences*, 3 (5), 120-125.
- [4] Correa, A.G., Laciari, E., Patiño, H.D., Valentinuzzi, M.E. (2007). Artefact removal from EEG signals using adaptive filters in cascade. *Journal of Physics*, 90 (1), 1-10.
- [5] Jutten, C., Herault, J. (1991). Blind separation of sources, Part I: An adaptive algorithm based on neuromimetic architecture. *Signal Processing*, 24 (1), 1-10.

- [6] Vigário, R., Sarela, J., Jousmiki, V., Hamalainen, M., Oja, E. (2000). Independent component approach to the analysis of EEG and MEG recordings. *IEEE Transactions on Biomedical Engineering*, 47 (5), 589-593.
- [7] Dora, C., Biswal, P.K. (2016). Robust ECG artefact removal from EEG using continuous wavelet transformation and linear regression. In *International Conference on Signal Processing and Communications*, Bangalore, India. IEEE, 1-5.
- [8] Jiang, J., Chao, C., Chiu, M., Lee, R., Tseng, C., Lin, R. (2007). An automatic analysis method for detecting and eliminating ECG artifacts in EEG. *Computers in Biology and Medicine*, 37 (11), 1660-1671.
- [9] Hamaneh, M.B., Chitravas, N., Kaiboriboon, K., Lhatoo, S.D., Loparo, K.A. (2014). Automated removal of EKG artefact from EEG data using independent component analysis and continuous wavelet transformation. *IEEE Transactions on Biomedical Engineering*, 61 (6), 1634-1641.
- [10] Mak, J.N.F., Hu, Y., Luk, K.D.K. (2010). An automated ECG-artefact removal method for trunk muscle surface EMG recordings. *Medical Engineering & Physics*, 32, 840-848.
- [11] Delorme, A., Makeig, S., Sejnowski, T.J. (2001). Automatic artefact rejection for EEG data using high-order statistics and independent component analysis. In *3rd International Independent Component Analysis and Blind Source Decomposition Conference*, San Diego, 9-12.
- [12] Nolan, H., Whelan, R., Reilly, R.B. (2010). FASTER: Fully automated statistical thresholding for EEG artifact rejection. *Journal of Neuroscience Methods*, 192 (1), 152-162.
- [13] Mognon, A., Jovicich, J., Bruzzone, L., Buiatti, M. (2011). ADJUST: An automatic EEG artefact detector based on the joint use of spatial and temporal features. *Psychophysiology*, 48 (2), 229-240.
- [14] Sahonero-Alvarez, G. (2017). A comparison of SOBI, FastICA, JADE and Infomax algorithms. In *8th International Multi-Conference on Complexity, Informatics and Cybernetics (IMCIC 2017)*, Orlando, Florida, 17-22.
- [15] Rutledge, D.N., Jouan-Rimbaud Bouveresse, D. (2013). Independent components analysis with the JADE algorithm. *TrAC Trends in Analytical Chemistry*, 50, 22-32.
- [16] Shen, M., Zhang, X., Li, X. (2002). Independent component analysis of electroencephalographic signals. In *6th International Conference on Signal Processing*, Beijing, China, 1548-1551.
- [17] Winkler, I., Haufe, S., Tangermann, M. (2011). Automatic classification of artifactual ICA-components for artifact removal in EEG signals. *Behavioral and Brain Functions*, 7, 30.
- [18] Burger, C., van den Heever, D.J. (2015). Removal of EOG artefacts by combining wavelet neural network and independent component analysis. *Biomedical Signal Processing and Control*, 15, 67-79.
- [19] Ichimaru, Y., Moody, G.B. (1999). Development of the polysomnographic database on CD-ROM. *Psychiatry and Clinical Neurosciences*, 53, 175-177.
- [20] Terzano, M.G., Parrino, L., Sherieri, A., et al. (2001). Atlas, rules, and recording techniques for the scoring of cyclic alternating pattern (CAP) in human sleep. *Sleep Medicine*, 2 (6), 537-553.
- [21] Goldberger, A.L., Amaral, L.A., Glass, L., Hausdorff, J.M., Ivanov, P.C., Mark, R.G., Stanley, H.E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation*, 101 (23), 215-220.
- [22] Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, 2011, 1-9.
- [23] Christov, I.I. (2004). Real time electrocardiogram QRS detection using combined adaptive threshold. *Biomedical Engineering Online*, 3 (1), 1-9.

Received December 17, 2018  
Accepted May 30, 2019