Proceedings of CONF-SEML 2026 Symposium: Multimodal Data Acquisition: Applications in Physiological and Behavioral Research
DOI: 10.54254/2755-2721/2026.AD33083

Application of Independent Component Analysis (ICA) Methods in EEG Signal Processing

Leyan Chen

College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, China
3240104595@zju.edu.cn

Abstract. EEG is a non-invasive technique for recording brain activity, valued for its high temporal resolution and low cost. Independent Component Analysis (ICA), a blind source separation method, effectively decomposes EEG signals into independent components. Widely used in both medical research and brain-computer interfaces (BCIs), ICA has become essential for improving EEG data quality. As demand grows for real-time and robust neural signal processing, advanced methods like ICA remain crucial for advancing EEG applications. This paper investigated the application of ICA in EEG signal processing. The research first reviewed the core principles of ICA, including its mathematical model and algorithms such as Infomax and FastICA. The results clearly demonstrated that ICA can effectively decompose raw EEG signals into statistically independent components. This enables the identification and removal of artifacts while preserving neural information. It also showed that while ICA is a mature method for EEG preprocessing, challenges such as source number determination and real-time processing bottlenecks still exist. At the end of the research, future developing trends of ICA and its integration with other methods are discussed.

Keywords: Independent Component Analysis (ICA), Electroencephalogram (EEG), Signal Processing, Artifact Removal, Blind Source Separation

1. Introduction

Electroencephalography, or EEG, captures time-series electrical signals from the brain using electrodes placed on the scalp. Raw EEG recordings, though, do not show individual neural sources directly. Instead, they are mixed signals formed by the superposition of electrical fields from multiple brain regions. What's more, EEG recordings are often disturbed by artifacts—both physiological and technical ones—that have nothing to do with actual brain activity.

By using Independent Component Analysis (ICA), researchers can break EEG signals down into separate components. From there, artifact components can then be identified either manually or automatically, based on their time courses, scalp topographies, and spectral characteristics. Compared with conventional filtering or regression methods, ICA removes artifacts more precisely while causing less damage to useful EEG information.

This paper focuses on ICA's use in EEG artifact removal, a fundamental and widely studied task. We will summarize existing research, review recent advances, and discuss about current limitations

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2026.AD33083

and challenges. We will also look into future directions of ICA development. These include for example, combining ICA with deep learning methods and exploring its use in real-time processing for brain-computer interface systems.

2. Introduction to ICA methods

2.1. Core principles and fundamental concepts of ICA

Independent Component Analysis (ICA) is a blind source separation method based on higher-order statistical properties. Its core principle involves decomposing observed signals into statistically independent, non-Gaussian source signals through linear transformations. Unlike signal processing techniques such as Principal Component Analysis (PCA), which rely solely on second-order statistics, ICA optimizes separation by maximizing the non-Gaussianity of source signals, thereby leveraging higher-order statistical characteristics.

Assume there exist N independent source signals: $s\left(t\right) = \left[s_1\left(t\right), s_2\left(t\right), s_3\left(t\right), \ldots, s_n\left(t\right)\right]^T$ and N observed signals: $x\left(t\right) = \left[x_1\left(t\right), x_2\left(t\right), x_3\left(t\right), \ldots, x_n\left(t\right)\right]^T$. These observed signals result from the linear mixing of the source signals, i.e.,

$$x\left(t\right) = \sum_{i=1}^{N} a_i s_i(t) \tag{1}$$

$$x\left(t\right) = As(t) \tag{2}$$

a_i is a column vector of A . Thus, the observed signals can be regarded as a linear combination of the matrix's column vectors, with the weights by the respective independent sources. To some extent, a_i reflects the spatial patterns of the independent sources $s_i(t)$,such as their approximate locations within the brain. Consequently, the matrix A is also referred to as scalp topography.

After linear combination, the components of the observed signal are no longer independent. The ICA method seeks to find a linear transformation W , known as the unmixing matrix, which is applied to the observed signal to maximize the independence of the resulting output signals.

$$y\left(t\right) = Wx(t) \tag{3}$$

The resulting output $y\left(t\right)$ then serves as an estimate of the source signal $s\left(t\right)$.

It is crucial to note that ICA methods inherently involve ambiguities, manifested in two aspects: 1. Permutation indeterminacy: The order of components in the recovered signal $y\left(t\right)$ may differ from that in $s\left(t\right)$; 2. Scaling indeterminacy, where the amplitude of the recovered signal is uncertain because the demixing matrix may contain amplitude information. For the second point, in practical applications, one can assume the variance of the centered source signal is 1, transferring the amplitude information to the mixing matrix without affecting the analysis of the separated waveforms.

$$y\left(t\right) = Wx\left(t\right) = WAs\left(t\right) \tag{4}$$

Let $P = WA$. Under ideal conditions (accounting for the aforementioned indeterminacy), $P = RS$, where R is the permutation matrix and S is the diagonal matrix.

DOI: 10.54254/2755-2721/2026.AD33083

Beyond the fundamental assumption of independent source signals, the analysis of the above model imposes several additional constraints that are prerequisites for its validity. These primarily include: 1. At most one source signal follows a normal distribution; 2. The number of observed signals must be greater than or equal to the number of source signals (for computational convenience, algorithms typically treat both as equal); 3. The matrix A has full column rank; 4. Observed signals contain no noise or only additive low-level noise [1].

In electroencephalogram (EEG) analysis and processing, ICA possesses inherent applicability. During EEG signal recording, non-EEG-related activities often introduce spurious signals, known as artifacts. Pseudo-disturbances in EEG manifest in various forms, with common types including electrocardiographic (ECG), electromyographic (EMG), blink, eye movement, sweat, and power-frequency interference. Since EEG signals originate from cortical neuronal activity, whereas eye movements, muscle activity, power-frequency interference, and ECG signals are typically independent of brain activity—i.e., mutually independent; For EEG signals, conduction can be approximated as instantaneous and linear. Furthermore, both EEG signals and artifact signals can be strictly considered non-Gaussian distributions. Although the number of effective sources generating scalp EEG activity is unknown, extensive experiments have demonstrated that the ICA algorithm can accurately distinguish time processes with relatively large instantaneous independent components from EEG signals [2].

2.2. Implementation of the ICA algorithm

Research on different ICA algorithms primarily focuses on the selection of independence metrics and optimization algorithms. Given the nature of ICA, determining the objective function to measure independence is a fundamental challenge all algorithms must address. From an information-theoretic perspective, the most natural criterion for achieving statistical independence among output components is requiring their mutual information to be zero.

The first major ICA algorithm is Infomax, a self-organizing neural network algorithm proposed by Bell and Sejnowski in 1995. Its objective function is the joint differential entropy of the output components. When the network selects an appropriate nonlinear function, maximizing this function directly minimizes mutual information(as shown in Figure 1) [3].

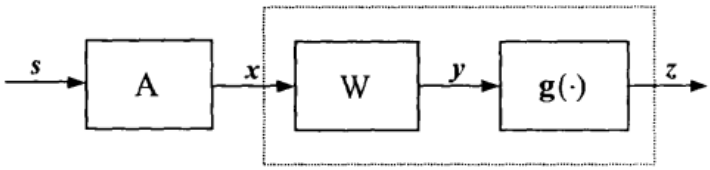

Figure 1. Schematic diagram of the infomax algorithm

Based on the properties of mutual information, it is known that the nonlinear mapping from one component to component has no impact on mutual information. Minimizing mutual information of z implies minimizing the mutual information of y.

$$H(z) = \sum_{i=1}^{n} H(z_i) - I(z) \tag{5}$$

In the equations, $H(z_i)$ denotes the output edge entropy, and $I(z)$ represents their mutual information. Derivation yields:

DOI: 10.54254/2755-2721/2026.AD33083

$$H(z) = \sum_{i=1}^{n} -E\left\{\log \frac{p(y_i)}{|g_i'(y_i)|}\right\} - I\left(z\right) \tag{6}$$

$$\frac{\partial}{\partial W} H(z) = \frac{\partial}{\partial W} \sum_{i=1}^{n} -E\left\{\log \frac{p(y_i)}{|g_i'(y_i)|}\right\} - \frac{\partial}{\partial W}\left(I(z)\right) \tag{7}$$

When $p(y_i) = |g_i'(y_i)|$ holds, maximizing the joint entropy $H(z)$ of the outputs minimizes the mutual information. In the Infomax algorithm, the choice of $g_i(y_i)$ is relatively fixed, such as approximating the probability density function of the source signal using a logistic function. Based on actual separation results, as long as this approximation falls within a certain range, the algorithm ultimately converges to obtain the demixing matrix W.

For the weight iteration process in W, an adjustment formula exists:

$$\Delta W \propto \frac{\partial}{\partial W} H(z) = \left(W^T\right)^{-1} - \varphi(y)x^T \tag{8}$$

$$\varphi(y) = \left[-\frac{g_1''(y_1)}{g_1'(y_1)}, \ldots, -\frac{g_n''(y_1)}{g_n'(y_1)}\right]^T \tag{9}$$

For this iterative process, Amari et al. proposed the natural gradient method in 1996 to optimize as follows, $\Delta W \propto \frac{\partial}{\partial W} H(z) W^T W$, simplifying the computation while also preventing W from becoming a singular matrix [4].

Furthermore, considering that the logistic function assumes a hyper-Gaussian distribution of source signals as a nonlinear function, when both hyper-Gaussian and sub-Gaussian distributions coexist in the source signals, the nonlinear function may not match well with the cumulative density of the source signals, leading to the algorithm's inability to perform correct separation. Therefore, Lee et al. proposed an extended ICA algorithm. This approach adjusts the separation matrix W while simultaneously modifying $g_i(y_i)$ based on the fourth-order moment of y_i. This adaptation enables the algorithm to perform blind separation for most signal source distributions, including those with both super-Gaussian and sub-Gaussian components [5].

The second major ICA algorithm is FastICA, also known as the fixed-point algorithm. It adopts a batch - processing strategy, and takes the direction of maximum negative entropy as the search criterion to sequentially extract independent sources. It is known that among distributions with identical covariance matrices, the Gaussian distribution possesses maximum entropy. The so-called negative entropy quantitatively measures the deviation of a distribution from its corresponding Gaussian distribution. Negative entropy $J(x)$ is defined as follows:

$$J(x) = H_G(x) - H(x) \tag{10}$$

In 1994, Comon proposed approximating negative entropy using higher-order cumulants in his work, introducing a cost function based on fourth-order cumulants [6]. Finnish scholar Aapo further developed approximation methods for negative entropy, proposing a fixed-point algorithm with enhanced robustness and efficiency [7].

Overall, fixed-point algorithms have the following key characteristics: 1. Fast convergence; 2. Integration of projection pursuit, a classic linear transformation method for sequentially extracting distinct independent components; 3. No requirement for learning rate parameter tuning; 4. Flexible selection of nonlinear functions [8].

DOI: 10.54254/2755-2721/2026.AD33083

The two algorithms mentioned above are among the most influential ICA methods currently. Others include the Mutual Information Minimization (MMI) algorithm and the Maximum Likelihood (ML) algorithm. Although their specific implementations differ, they are unified from an information-theoretic perspective.

2.3. Application of ICA in EEG signal processing

ICA methods were first used in blind source separation for speech signals. Later, as neuroscientists needed cleaner EEG data, ICA was first attempted for removing electrooculographic (EOG) artifacts. Currently, ICA methods are maturely applied in EEG signal denoising, primarily involving the separation of various artifact signals and the exploration of their spatial patterns. In addition, since ICA can extract independent source signals within the brain, it is also frequently used for feature extraction and source localization in EEG signals.

ICA has clear advantages over methods that use reference template signals, wavelet transforms, or specific frequency notch filters. Unlike traditional time-frequency domain analysis methods, ICA preserves signal details during decomposition better. Its core strength lies in automatically identifying and removing artifacts without requiring prior knowledge of noise shapes—merely leveraging the assumption of statistical independence. This makes ICA particularly suitable for artifacts with fixed spatial distributions or specific statistical characteristics. Some of the typical artifact types are listed below in Table 1. The general ICA artifact removal process goes like this: separate independent components; identify and determine the artifact components to be eliminated; set the artifact components to zero and apply the inverse matrix to obtain the corresponding artifact-free EEG signal. Moreover, for a specific independent component obtained, an inverse transform yields its characteristic distribution across all detection sources, thereby indicating the spatial distribution of that component.

Table 1. Typical artifact types and their characteristics

Artifact Types	Characteristics
Electrooculographic (EOG) artifacts Arise from eye movements, blinking, or eyelid tremors	Extremely high amplitude (typically several times that of the EEG signal), low frequency (approximately 3-16Hz), most prominent at frontal electrodes (e.g., Fp1, Fp2, F7, F8)
Electromyographic artifacts (EMG) are generated by facial muscles, neck muscles, or head movements	Higher frequency (typically above 30 Hz), appearing near specific muscle locations
Electrocardiographic artifacts (ECG) are caused by electrical fluctuations from heartbeats	Manifest as regular spikes near the neck, jaw, or chest, with frequencies synchronized to heart rate

Currently, ICA serves as a fundamental method in EEG signal processing, supporting research in functional connectivity (extracting independent oscillatory sources to construct source-space brain networks), epileptic spike localization (separating lesion-specific discharge components), BCI system performance enhancement (making the motor imagery or P300 signals more classifiable), psychiatric disease biomarker discovery (e.g. temporal abnormalities in default mode network components of depressed patients), and multimodal fusion (achieving spatiotemporal super-resolution by combining fMRI and EEG via ICA).

DOI: 10.54254/2755-2721/2026.AD33083

3. Challenges of ICA in EEG signal processing

In practical EEG signal processing, ICA still faces a number of limitations. For one thing, it is not easy to decide how many source signals should be extracted. Noise identification can also be uncertain and unreliable. ICA relies on strict assumptions about signal distribution, which limits its performance. It may not work well when sources are missing or the signal-to-noise ratio is low. In addition, ICA is computationally demanding and hard to use in real-time tasks.

3.1. Empirical estimation of source number

ICA cannot automatically identify how many independent sources exist in an EEG signal. Researchers usually have to set this number manually based on experience. If the value is set either too high or too low, separation quality may drop significantly, and the algorithm may even fail to converge. In practice, researchers often choose the number of sources based on experience, usually between 20 and 30. However, this empirical approach lacks solid theoretical support and may lead to inconsistent results under different experimental conditions. Problems also appear when the number of sources is close to or larger than the number of electrodes. For example, setting 128 sources for a 128-channel EEG often produces many similar or redundant components.

3.2. Uncertainty in artifact identification

ICA results are not fully consistent across different runs. Automatic artifact detection is often unstable, so manual checking is still needed in many situations. One artifact can be spread over several components. At the same time, neural signals and noise may mix into a single component. Take motor imagery tasks for example, μ/β rhythms are often combined with EMG artifacts in the same component, which makes clean separation difficult. In some situations, important neural components may even be wrongly recognized as noise and removed.

3.3. Limitations of distribution and stationarity assumptions

Since ICA is based on the assumption that source signals are non-Gaussian and statistically independent. This means it struggles to handle Gaussian noise and may even fail at all. For instance, in clinical settings, this may result in poor removal of common interference such as 50 or 60 Hz power line noise.

When the statistical properties of source signals (e.g., mean, variance) change over time, traditional ICA assumptions of statistical independence and distribution invariance are no longer hold. The algorithm may then converge to wrong results or fail to converge [9]. This is especially obvious in dynamic conditions such as epileptic seizures. Traditional ICA only works with linear transient mixing models, whereas actual EEG signals may involve nonlinear mixing processes. For instance, the nonlinear propagation of certain electromyographic (EMG) interferences can also cause distorted separation results [8].

3.4. Real-time performance

The computational complexity of ICA mostly comes from matrix operations and iterative convergence processes. As data dimensions and sample sizes get bigger, the required computing resources grow rapidly. At present, ICA relies heavily on offline processing, leaving room for improvement in real-time applications.

DOI: 10.54254/2755-2721/2026.AD33083

4. ICA as a development direction for EEG signal processing technology

Future development of ICA can be pursued in several ways: Integration with deep learning to build end-to-end automated processing frameworks; Combination with multi-dimensional analysis methods such as tensor decomposition to enhance performance for non-stationary and underdetermined problems (where the number of observed signals is fewer than the original signals); Lightweight models and hardware acceleration to achieve real-time processing; Multimodal data fusion to further expand its application scenarios.

To improve real-time performance, researchers have proposed online ICA or neural network-based variants (e.g., OICNet). These methods use online learning, updating model parameters with minimal data to reduce computation time, which makes them suitable for real-time brain-computer interface (BCI) systems and real-time EEG source separation [10].

5. Conclusion

Independent Component Analysis (ICA) has proven to be a powerful and effective tool for EEG signal processing, particularly in the separation of neural sources and the removal of artifacts such as eye movements and muscle activity. Although widely applied, ICA still faces unresolved challenges. With continued improvements to these limitations, ICA is expected to become increasingly valuable not only in routine EEG signal processing but also in cutting-edge cognitive neuroscience research and clinical applications.

References

[1] LI Ting. Independent Component Analysis Method for EEG Noise Reduction and Its Application Research [D]. Anhui University, 2002.
[2] LI Ting, QIU Tianshuang, NIU Jie. The Development of Independent Component Analysis and its Application in Brain Signals [J]. School of Electronic and Information Engineering, Dalian University of Technology, Dalian 116024, China.
[3] A. J. Bell and T. J. Sejnowski. An information-maximization approach to blind separation and blind deconvolution [J]. Neural Computation, vol. 7, pp. 1129-59, 1995.
[4] S. Amari. Natural gradient works efficiently in learning [J]. Neural Computation, vol. 10, pp. 251-276,
[5] T. W. Lee, M. Girolami, and T. J. Sejnowski. Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources [J]. Neural Computation, vol. 11, pp. 417-41, 1999.
[6] P. Comon. Independent component analysis, a new concept [J]. Signal Processing, vol. 36, pp. 287-314, 1994.
[7] A. Hyvarinen. New approximations of differential entropy for independent component analysis and project pursuit [C]. Advances in Neural Information Processing Systems, vol. 10, pp. 273-279, MIT Press, 1998.
[8] Long Fei. Independent Component Analysis Method for EEG Noise Reduction and Its Application Research [D]. Anhui University, 2002.
[9] Wei Lu; Jagath C. Rajapakse. Eliminating Indeterminacy in ICA [J]. Neurocomputing*, 2003.
[10] YEH P T, TSAI A C, HSIEH C Y, et al. OICNet: A Neural Network for Online EEG Source Separation using Independent Component Analysis [J]. 2023-06-01.