Volume 9, Issue 1, 092501 (1-8)

Signal Processing and Renewable Energy (SPRE)

https://doi.org/10.57647/j.spre.2025.0901.01

Detection of ADHD from electroencephalogram using independent component analysis

Fereshteh Ghanavati, Mohammad Adeli*

Department of Biomedical Engineering, Dezful Branch, Islamic Azad University, Dezful, Iran.

*Corresponding author: mohammad.adeli@iau.ac.ir

Original Research

Received:
13 January 2025
Revised:
7 February 2025
Accepted:
8 February 2025
Published online:
1 March 2025

Abstract:

Attention deficit hyperactivity disorder (ADHD) is one of the most common mental disorders affecting children. Delayed diagnosis and treatment of ADHD might result in poor school performance. In this paper, an intelligent system is proposed for detecting ADHD from electroencephalograms. The system first preprocesses the electroencephalograms to remove the powerline and blinking interferences. Then, independent component analysis is applied to separate the signals of the sources that generate the preprocessed electroencephalograms. Later, 3 features of Shannon entropy, kurtosis, and skewness are extracted from the independent components. A total of 57 features are obtained for 19 electroencephalogram channels. Finally, two classifiers (i.e., the k-nearest neighbors with k = 3, and an artificial neural network) were used to detect ADHD from the extracted features. We used a public database containing signals from 61 children with ADHD and 60 healthy children to evaluate the performance of the proposed system. The accuracy, sensitivity, and specificity of the artificial neural network were 99.92%, 99.89%, and 99.95%, respectively, while these metrics were all 100% for the k nearest neighbors. The results imply that the features extracted from independent components of electroencephalograms are more powerful than other linear/nonlinear features directly extracted from the signals.

Keywords: Attention deficit hyperactivity disorder; Electroencephalogram; Independent component analysis; Signal processing; Feature extraction; Artificial neural network; *K*-nearest neighbors

1. Introduction

The symptoms of attention-deficit hyperactivity disorder begin in childhood. ADHD can affect social relationships and school performance. This disorder is common among children, but many adults also have the disorder.
People with ADHD experience the following symptoms [1]:

- Inattention: difficulty in fulfilling tasks, paying attention, or organizing.
- Hyperactivity: restlessness and constant movement, or excessive talking.
- Impulsivity: acting without thinking, interrupting others, or having trouble controlling oneself.

ADHD treatment usually includes a combination of medications and therapy. Typically, ADHD is diagnosed by a specialist through a conversation with the parents and sometimes by performing some psychological tests [1]. Given the qualitative nature of these methods, the diagnosis and treatment of ADHD may be delayed. Therefore, the purpose of this research is to provide a quantitative, efficient, and fast method based on the processing of electroencephalogram (EEG) signals to detect ADHD disorder.
An electroencephalogram is a recording of brain activities. There are many applications that EEG signal analysis can help with, including disease diagnosis [2]. Hereafter, some studies on ADHD detection from EEG are reviewed.
Ghassemi et al. used three nonlinear features of wavelet entropy, correlation dimension, Lyapunov power, and the k-nearest neighbor (KNN) algorithm to detect ADHD. This study achieved a classification accuracy of $\sim 96\%$ [3]. Mohammadi et al. (2016) [4] presented a method for classifying EEG from nonlinear features of Higuchi, Katz, and Petrosian fractal dimensions and approximate entropy. Finally, they obtained an accuracy of 93.65% using a multilayer perceptron (MLP) neural network [4]. Allahverdy et al. (2016) [5] also used EEG nonlinear features to diagnose ADHD. The nonlinear features of the Lyapunov exponent, Higuchi fractal dimension, Katz fractal dimension, and Sevcik fractal dimension were extracted from EEG data. Their MLP neural network classifier was able

to achieve 96.7% accuracy [5]. Markovska-Simoska and Pop-Jordanova (2017) [6] evaluated quantitative EEG in adults and children with ADHD. They compared the absolute and relative EEG power, especially the theta/beta ratio, for children with ADHD and normal subjects. They suggested that the theta/beta ratio should be used to identify ADHD children. Their results also showed that the absolute power of slow waves (delta and theta) was higher in ADHD children than in normal subjects [6]. Khoshnoud et al. (2018) [7] analyzed the brain function of ADHD children with nonlinear dynamics present in EEG signals. They reported that the average entropy values obtained using the approximate entropy in the frontal regions of the brains of ADHD children were significantly different from the control group. A neural network classifier provided an accuracy of 83.33% based on entropy [7]. Alchalabi et al. (2018) [8] used a supervised classifier to identify ADHD patients. In this study, participants had to play a "focus" game to assess their attention level. The accuracy of this method for detecting the correct attention state during the game was 96% and 98% for classifying the patients' EEG data [8]. Catherine Joy et al. (2022) [9] used five features of fuzzy entropy, logarithmic energy entropy, permutation entropy, Sure entropy (SURE), and Shannon entropy. Finally, using a multilayer perceptron neural network, they classified the signals into two categories healthy and ADHD. The accuracy obtained was 99.82% [9]. It is worth noting that they used an EEG database containing signals from 10 people (5 healthy and 5 patients). Taghi Beyglou et al. (2022) [10] used a convolutional neural network for feature extraction. The signals were first divided into 30-second segments. Then, the obtained image was applied as input to the convolutional neural network. Finally, a support vector machine (SVM) obtained an accuracy of 91.16% using the features extracted by the convolutional neural network [10]. Maniruzzaman et al. (2022) [11] first extracted 342 morphological (such as peak-to-peak amplitude, area under the graph, etc.) or temporal (such as standard deviation, kurtosis, skewness, etc.) features from the signals. Then, using the LASSO method, they reduced the number of features to 47. They achieved 94.2% accuracy using the SVM classifier [11]. Alim and Imtiaz (2023) [12] extracted 836 features from the 4 bands of delta, theta, alpha, and beta. In the feature reduction stage, features that made up 90% of the data variability were selected using PCA. The test results of this method showed that the classification accuracy was 93.2% using the SVM classifier [12].

In previous studies, various linear and nonlinear features of EEG signals have been used to detect ADHD. These features are usually extracted directly from the EEG signal or its delta, theta, alpha, beta, and gamma components. However, in this study, feature extraction is not performed directly from the EEG signal, but rather from components extracted using independent component analysis. For that purpose, EEG signals are first separated into source signals using independent component analysis (ICA). Then, various features of these sources are extracted and finally used to classify people into healthy and ADHD groups.

The rest of this paper is organized as follows. In section 2, the ADHD dataset and the proposed method for ADHD detection are introduced, the results of this study are presented in section 3, the findings are discussed and compared with similar studies in section 4, and the final conclusion of the study is presented in section 5.

2. Materials and methods

In this section, a system for detecting attention deficit hyperactivity disorder is proposed. The components of the proposed method, including preprocessing, independent component analysis (ICA), feature ex-traction, classification, and the required dataset, are explained in detail.

2.1 ADHD dataset

The EEG dataset used in this study is freely available to the public at IEEEDataPort (https://ieee-dataport.org) [13]. All participants in this study were school-aged and righthanded. The participants included 60 healthy children and 61 children with ADHD who were evaluated by an experienced child psychiatrist. The diagnosis of ADHD was made based on DSM-IV criteria [14]. Children with ADHD had been taking Ritalin for up to 6 months [13]. The EEG signals were recorded by 19 electrodes according to the $10-20$ standard EEG system (Fig. 1). The accuracy of the participants' responses during the designed scenario was not considered [15].

It should be noted that the sampling rate of the signals was 128 Hz, but in this study, we increased it to 256 Hz. Also, due to the small number of subjects (121 subjects), in order to increase the number of signals, we divided the signals of each individual into 10-second segments. Consecutive segments did not overlap. Five 10-second EEG segments were used for each individual. Thus, the data used in this study consist of 605 (121 subjects $\times$ 5 segments) 19-channel EEG signals, each with a duration of 10 seconds.

2.2 The proposed method

The block diagram of the proposed method for automatic ADHD detection from EEG signals is shown in Fig. 2. The

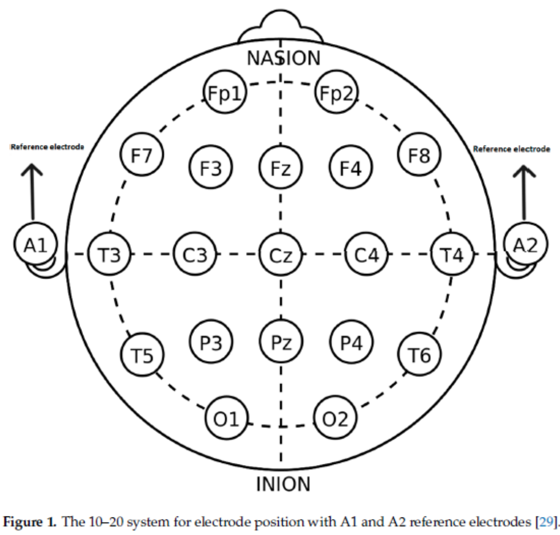

Figure 1. Location of electrodes in the EEG $10-20$ system.

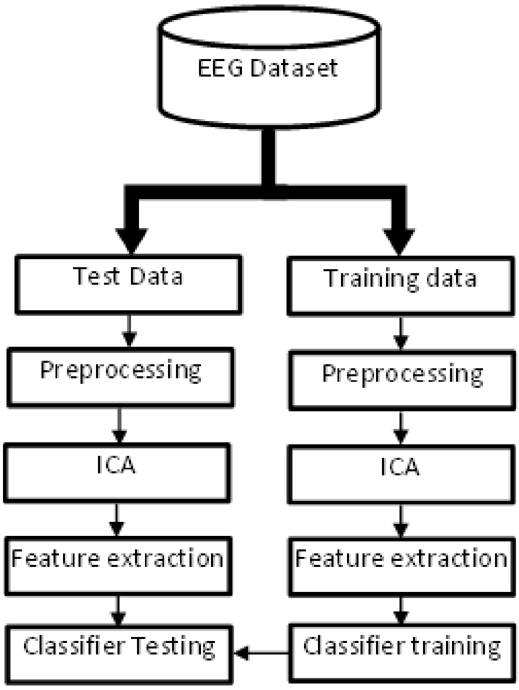

Figure 2. Block diagram of the proposed system for detecting ADHD from EEG signals.

proposed method includes partitioning the dataset into training and test sets, preprocessing the EEG signals, and ICA, extracting features from them, training a classifier using the extracted features, and testing the classifier using the test data. The components of the proposed method are described in detail below.

2.2.1 Preprocessing

EEG signals are preprocessed to remove powerline and blinking interferences. To remove powerline interference, we used a band-pass filter with finite impulse response (FIR). We designed this filter using a window method with a rejection range from 47 to 53 Hz.

Blinking interference is removed using discrete wavelet transform [16]. To achieve this goal, we first decompose the EEG signals up to 5 levels using discrete wavelet transform. Here, we used the Symlet 4 mother wavelet. As a result, the EEG signals are separated into two categories, namely approximation coefficients and detail coefficients [17]. Next, we apply a threshold to the detail coefficients. Thresholding means that if the detail coefficients in a wavelet transform level are less than a certain threshold value, they are set to 0, otherwise they remain unchanged. After applying thresholding to the detail coefficients of all wavelet decomposition levels, we reconstruct the EEG signal. The reconstructed signal is an estimate of the blinking interference [16]. If we assume that $x[n]$ is the EEG signal contaminated with blinking interference and $x_r[n]$ is the reconstructed signal after applying wavelet transform and thresholding, the signal without blinking interference, namely $x_c[n]$, is obtained by the equation (1):

$$x_c[n] = x[n] - x_r[n] \tag{1}$$

The signal $x_c[n]$ is the preprocessed signal.

2.2.2 Independent component analysis

The proposed method is based on the assumption that the EEG signals are recorded using the standard $10-20$ EEG system [2]. The $10-20$ system provides 19 EEG channels. We denote the preprocessed signals with $x_{c,1}[n]$, $x_{c,2}[n]$, ..., and $x_{c,19}[n]$, respectively. ICA is commonly used to solve the cocktail party problem [18], but we use it to separate EEG sources. In ICA [19], these signals are called mixed signals because we assume that each of them is a mixture of the components generated by 19 unknown sources. Assuming that the signals of these unknown sources are $s_1[n]$, $s_2[n]$, *ldots*, and $s_{19}[n]$, respectively, each EEG signal such as $x_{c,i}[n]$ is a linear combination of these components:

$$x_{c,i}[n] = a_{i,1}s_1[n] + a_{i,2}s_2[n] + \cdots + a_{i,19}s_{19}[n] \tag{2}$$

For convenience, we assume that the random vectors $\boldsymbol{x}$ and $\boldsymbol{s}$ contain the mixed signals $\{x_{c,i}[n], 1 \leq i \leq 19\}$, and the source signals $\{s_i[n], 1 \leq i \leq 19\}$. We also assume that the matrix $\boldsymbol{A}$ is the mixing matrix. Therefore, Eq. (2) can be rewritten as:

$$\boldsymbol{x} = \boldsymbol{A}\boldsymbol{s} \tag{3}$$

2.2.3 Signal centering

The first step of ICA is to remove the mean of the mixed signals [19]. To do this, if we assume that the vector $\boldsymbol{m} = E\{\boldsymbol{x}\}$ is the mean of the mixed signals, it is sufficient to subtract the mean of each signal from it as follows:

$$\ddot{\boldsymbol{x}} = \boldsymbol{x} - \boldsymbol{m} \tag{4}$$

where vector $\ddot{\mathbf{x}}$ represents the mixed signals after removing their mean.

2.2.4 Whitening

The second step of ICA is to whiten the zeromean signals $\ddot{\boldsymbol{x}}$ [19]. The goal is to obtain the vector $\tilde{\boldsymbol{x}}$ such that its components are white, that is, they are uncorrelated and their variance is equal to 1. Mathematically, the covariance matrix of $\tilde{\boldsymbol{x}}$ must be an identity matrix:

$$E\{\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}^{\boldsymbol{T}}\} = \boldsymbol{I} \tag{5}$$

where the operator $E\{\cdot\}$ represents expectation of a matrix of random variables. A common method for whitening is the eigenvalue decomposition of the covariance matrix [20]:

$$E\{\ddot{\boldsymbol{x}}\ddot{\boldsymbol{x}}^{\boldsymbol{T}}\} = \boldsymbol{E}\boldsymbol{D}\boldsymbol{E}^{\boldsymbol{T}} \tag{6}$$

where $\boldsymbol{E}$ is an orthogonal matrix containing the eigenvectors of the covariance matrix $E\{\ddot{\boldsymbol{x}}\ddot{\boldsymbol{x}}^{\boldsymbol{T}}\}$ and $\boldsymbol{D}$ is a diagonal matrix whose entries on the main diagonal are equal to the eigenvalues of the covariance matrix $E\{\ddot{\boldsymbol{x}}\ddot{\boldsymbol{x}}^{\boldsymbol{T}}\}$. It is worth noting that the covariance matrix $E\{\ddot{\boldsymbol{x}}\ddot{\boldsymbol{x}}^{\boldsymbol{T}}\}$ is estimated from the zeromean signals $\ddot{\boldsymbol{x}}$ By obtaining $\boldsymbol{D}$ and $\boldsymbol{E}$, we can estimate the whitened data $\tilde{\boldsymbol{x}}$ as:

$$\tilde{\boldsymbol{x}} = \boldsymbol{E}\boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{E}^{\boldsymbol{T}}\ddot{\boldsymbol{x}} \tag{7}$$

2.2.5 FastICA method

In the FastICA [19] method, the goal is to find the separation matrix $\boldsymbol{W}$ such that by multiplying it by the whitened vector $\tilde{\boldsymbol{x}}$ the unknown source signals are estimated:

$$\hat{\boldsymbol{s}} = \boldsymbol{W}\tilde{\boldsymbol{x}} \tag{8}$$

In Eq. (8), $\hat{\boldsymbol{s}}$ are the estimates obtained for the source signal-$\boldsymbol{s}$. Assuming that the vector $\boldsymbol{w}_i^T$ represents the i^{th} row of the matrix $\boldsymbol{W}$, we must find $\boldsymbol{w}_i^T$ such that $\boldsymbol{w}_i^T \tilde{\boldsymbol{x}}$, which is actually the estimate of $\boldsymbol{s}_i$, is non-Gaussian. The non-Gaussianity of $\boldsymbol{s}_i$ is equivalent to the independence of the sources [19]. The following pseudocode is used to obtain the $\boldsymbol{w}_i$ and source signals:

```
for i in 1 to C:
w_i ← random initialization
while w_i changes
w_i ← (1/N) X̃ g(w_i^T X̃)^T - (1/N) g'(w_i^T X̃) 1_N w_i
w_i = w_i - Σ_{j=1}^{i-1} (w_i^T w_j) w_j
w_i = w_i / ||w_i||
output W = [w_1 w_2 ... w_C]
output Ŝ = W X̃
```

In the pseudocode above, N is the number of samples of each signal, C is the number of the source signals, and the function $g(\cdot)$ is defined by:

$$g(u) = \tanh(u) = \frac{e^u - e^{-u}}{e^u + e^{-u}} \tag{9}$$

2.2.6 Feature extraction

After performing ICA, 19 independent components (source signals) $s_1[n]$ to $s_{19}[n]$ are obtained. For each of these components, 3 features of skewness, kurtosis and entropy are calculated. For example, the features of $s_1[n]$ are defined as follows:

$$Sk_1 = \frac{\frac{1}{N}\sum_{n=1}^{N}(s_1[n]-\bar{s}_1)^3}{[\frac{1}{N-1}\sum_{n=1}^{N}(s_1[n]-\bar{s}_1)^2]^{\frac{3}{2}}} \tag{10}$$

$$Kt_1 = \frac{\frac{1}{N}\sum_{n=1}^{N}(s_1[n]-\bar{s}_1)^4}{[\frac{1}{N-1}\sum_{n=1}^{N}(s_1[n]-\bar{s}_1)^2]^2} - 3 \tag{11}$$

$$En = -\sum_{I=1}^{I} p_i \log_2 p_i \tag{12}$$

To find p_i in (12), first, a 256-bin histogram is obtained for $s_1[n]$ sample values. Supposing that N_i and N represent the number of the samples falling in the i^{th} bin and the signal length, respectively, the probability mass function of $s_1[n]$ can be estimated as $p_i = \frac{N_i}{N}$. The number of the histogram bins, I, is 256.

By calculating the 3 features for each of the 19 components estimated with ICA, a total of 57 features are obtained. These features are used to train and test the classification models.

2.2.7 Classification

In this study, we considered the ADHD detection as a binary classification problem. The 57 features extracted from the source signals were the classifier inputs, and the classification output was either the healthy or ADHD label. In this study, an artificial neural network classifier [21] and the k-nearest neighbors [22] were used. The artificial neural network classifier consisted of an input layer with 57 neurons (receiving the 57 features), three hidden layers with 8, 20, and 8 neurons, respectively, and an output layer with 2 neurons.

To evaluate the performance of these two classifiers, we used three metrics: accuracy, sensitivity, and specificity. In this binary classification problem, where there are two positive (ADHD) and negative (healthy) classes, the above three metrics are estimated as follows:

$$A = \frac{TP+TN}{TP+FN+FP+TN} \tag{13}$$

$$S_n = \frac{TP}{TP+FN} \tag{14}$$

$$S_p = \frac{TN}{FP+TN} \tag{15}$$

where A, S_n, and S_p are accuracy, sensitivity and specificity, respectively. Other variables are defined as follows:

- True positive (TP): The number of ADHD patients correctly classified as positive.
- False negative (FN): The number of ADHD patients incorrectly classified as negative.
- False positive (FP): The number of healthy individuals incorrectly classified as positive.
- True negative (TN): The number of healthy individuals correctly classified as negative.

The proposed system was implemented using the MATLAB software. Also, a similar system that extracts features directly from the original signals was implemented and compared with the proposed system.

2.3 System evaluation

Ten-fold cross-validation was used to train and test both classifiers. In this type of validation, the data is randomly divided into 10 parts. The training and testing phases are repeated 10 times. Each time, 9 parts of the data are used for training and the remaining part is used for testing the classifiers. In this study, we randomly divided the 121 subjects into 10 groups. Nine (9) of them contain data from 12 people and one group contains data from 13 people. Ten-fold cross-validation was performed as follows:

- In one iteration, 540 signals (108 people $\times$ 5 signals) were used for training and 65 signals (13 people $\times$ 5 signals) were used for testing the classifiers.
- In the other 9 iterations, 545 signals (109 people $\times$ 5 signals) were used for training and 60 signals (12 people $\times$ 5 signals) were used for testing the classifiers.

Finally, after completing these 10 iterations, the average accuracy, sensitivity, and specificity metrics were calculated, which are presented in section 3.2.

3. Results

3.1 Preprocessing results

Fig. 3 shows the EEG signals of one of the ADHD patients. It is clear that these signals are contaminated with blinking artifacts.

Fig. 4 shows the same signals in Fig. 3 after preprocessing. It can be seen that the blinking artifact has been removed. After preprocessing, the independent components of the EEG signal were obtained by ICA, which are seen in Fig. 5. To ensure the performance of ICA, the correlation coefficients between the preprocessed signals in Fig. 4 were calculated and presented in Fig. 6, which shows that these signals are not independent because the correlation between them is non-negligible. It is worth noting that in Fig. 6, blue indicates low correlation values while yellow indicates high correlation values.

The correlation coefficients between the independent components in Fig. 5 were also calculated. These coefficients are shown in Fig. 7. As expected, the correlations between these independent components are very small. Therefore, the ICA has been successful in separating the independent source signals.

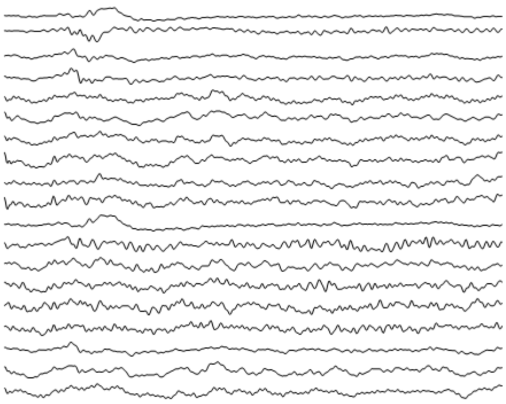

Figure 3. EEG signals from a subject with ADHD.

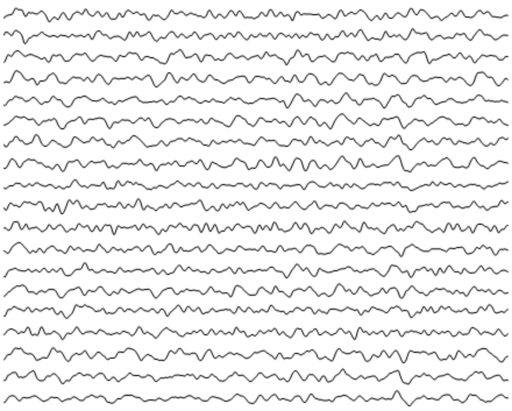

Figure 4. The preprocessed versions of Fig. 3 signals.

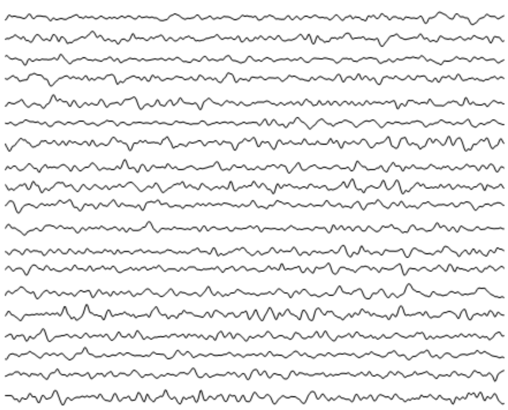

Figure 5. Independent components of the preprocessed signals in Fig. 4.

3.2 System evaluation results

In this section, we present the results of ADHD detection using the KNN and the artificial neural network classifiers. It is worth noting that in this study, we tested values of 1, 3, and 5 for k, and the best results were obtained for k = 3. The results obtained are summarized in Table 1 and Fig. 8. For the KNN classifier with k = 3, the accuracy, sensitivity, and specificity were all equal to 100%. For the artificial

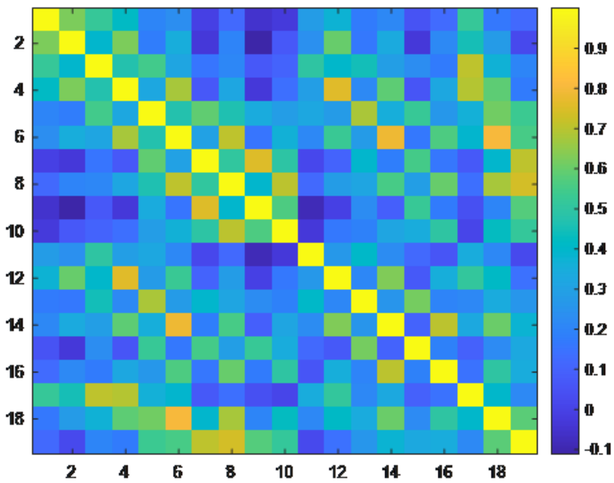

Figure 6. Correlations between the preprocessed EEG signals in Fig. 4.

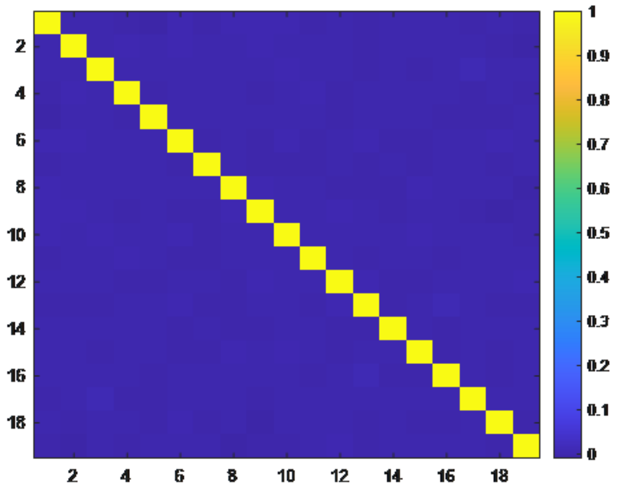

Figure 7. Correlation coefficients between independent components of EEG obtained by ICA.

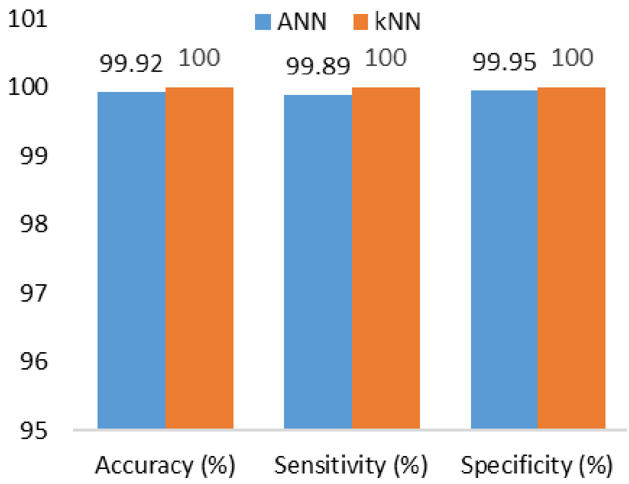

Figure 8. The results of the proposed system using the artificial neural network and the KNN classifier with k = 3.

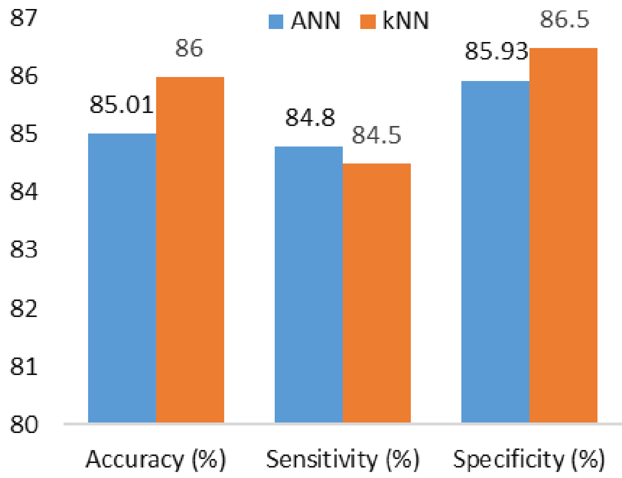

Figure 9. The results of a similar system using the artificial neural network and KNN classifier with k = 3.

neural network, the accuracy, sensitivity, and specificity parameters were equal to 99.92%, 99.89%, and 99.95%, respectively.

3.3 Results of a similar system

To further evaluate the performance of the system, another similar system was developed. This system directly extracts features from the preprocessed signals. The features are exactly the same as the previous ones (entropy, skewness, and kurtosis) for each of the EEG channels. Therefore, we still have 57 input features. To classify the signals into two categories healthy and ADHD, we used the k-nearest neighbor classifiers with k = 3 and the artificial neural network mentioned in section 2.2.7. The results of this system are summarized in Table 2 and Fig. 9.

4. Discussion

We proposed an intelligent system for detecting ADHD disorder from EEG signals in this study. This system is based on a 10 – 20 EEG recording system with 19 channels. First, we preprocess these signals to remove powerline interference and blinking interference. Then, we separate the independent EEG components using independent component analysis. In fact, it is assumed that each EEG signal is a linear combination of independent components generated by 19 sources. These independent components are extracted using ICA. Later, 3 features of Shannon entropy, kurtosis, and skewness are extracted from these signals. Thus, a total of 57 features (3 features in 19 channels) are obtained. Finally, we used the k-nearest neighbors with k = 3, and an artificial neural network to detect ADHD.

To evaluate the system performance, we used a public EEG database containing signals from 61 children with ADHD and 60 healthy children. The accuracy, sensitivity, and specificity metrics for the artificial neural network are 99.92%, 99.89%, and 99.95%, respectively, while all three metrics are 100% for the nearest parameter. The results show that the k-nearest neighbors performed better.

To better evaluate the system, a similar system was also developed that directly estimates features from preprocessed EEG signals. In this case, the accuracy, sensitivity, and specificity metrics were 85.01%, 84.80%, and 85.93%, respectively for the artificial neural network, while the accuracy, sensitivity, and specificity were 86%, 84.5%, and 86.5% for the k-nearest neighbors. By comparing these results with the results of the proposed system, it is clear that the proposed method has improved ADHD detection significantly.

In Table 3, the results of the proposed method are compared with similar studies of ADHD detection from EEGs. Only Alim and Imtiaz (2023) [12] and Maniruzzaman et al. (2022) [11] have used the ADHD database introduced in section 2.1. Other studies have used different datasets. The proposed method has outperformed all other studies except Catherine Joy et al. (2022) [9], which used a small dataset containing the EEG data of only 10 subjects (five healthy, and five ADHD). The results imply that ICA provides more powerful discriminating features for ADHD detection.

It is necessary to explain that the proposed method in this study has limitations, some of which are mentioned below. Though the KNN achieved an accuracy of 100%, it is necessary to evaluate the proposed method on different datasets. In addition, in the proposed method, features extracted from delta, theta, alpha, beta, and gamma waves were not used.

Table 1. The results of the proposed system using the artificial neural network and the KNN classifier with k = 3.

Classifier	Specificity	Sensitivity	Accuracy
ANN	99.95 %	99.89 %	99.92 %
KNN	100 %	100 %	100 %

Table 2. The results of a similar system using artificial neural network and the KNN classifier with k = 3.

Classifier	Specificity	Sensitivity	Accuracy
ANN	85.93 %	84.80 %	85.01 %
KNN	86.50 %	84.50 %	86.00 %

Table 3. Comparison of the results of the proposed method with similar studies of ADHD detection.

Study	Features	Classifier	Accuracy (%)
[12]	δ, θ, α, and β waves	SVM	94.00
[11]	Time domain and morphological features	SVM	94.20
[3]	wavelet entropy, correlation dimension, Lyapunov exponent	KNN	96.00
[4]	Higuchi's, Katz', and Petrosian fractal dimensions, approximate entropy	MLP	93.65
[5]	Higuchi's, Katz', and Sevcik fractal dimensions, Lyapunov exponent	MLP	96.70
[9]	Shannon -, fuzzy-, log energy-, permutation-, and SURE-entropy	MLP	99.82
Proposed method	Shannon entropy, skewness, and kurtosis of ICA components	KNN	100
		ANN	99.92

These waves are of great importance in medicine and have been used in many studies. Therefore, it seems that one of the essential tasks is to use features extracted from EEG waves and compare them with ICA features. This can determine the effect of ADHD on EEG waves.
Also, concatenation of the ICA features with other linear/nonlinear features should be considered in future studies. Classification models based on deep learning, support vector machines, and decision trees can also be used.

5. Conclusions

The method proposed for ADHD detection is based on feature extraction from independent components of EEGs obtained by ICA. This method was very successful in discriminating ADHD patients from healthy subjects. This implies that the features extracted from the independent components of EEG are more powerful than other linear/nonlinear features in ADHD detection applications.

Authors contributions
Authors have contributed equally in preparing and writing the manuscript.

Availability of data and materials
The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflict of interests
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] R. A. Barkley. "Attention-deficit hyperactivity disorder.". *Guilford Press*, 2003.

[2] J. D. Bronzino and D. R. Peterson. "Principles of electroencephalography.". *Biomedical Engineering Fundamentals*, pages 445–456, 2006.

[3] F. Ghassemi, M. Hassan-Moradi, M. Tehrani-Doost, and V. Abootalebi. "Using non-linear features of EEG for ADHD/normal participants' classification.". *Procedia-Social and Behavioral Sciences*, 32:148–152, 2012.

[4] M. R. Mohammadi, A. Khaleghi, A. M. Nasrabadi, S. Rafieivand, M. Begol, and H. Zarafshan. "EEG classification of ADHD and normal children using non-linear features and neural network.". *Biomedical Engineering Letters*, 6:66–73, 2016.

[5] A. Allahverdy, A. K. Moghadam, M. R. Mohammadi, and A. M. Nasrabadi. "Detecting ADHD children using the attention continuity as nonlinear feature of EEG.". *Frontiers in Biomedical Technologies*, 3(1-2):28–33, 2016.

[6] S. Markovska-Simoska and N. Pop-Jordanova. "Quantitative EEG in children and adults with attention deficit hyperactivity disorder: comparison of absolute and relative power spectra and theta/beta ratio.". *Clinical EEG and neuroscience*, 48(1):20–32, 2017.

[7] S. Khoshnoud, M. A. Nazari, and M. Shamsi. "Functional brain dynamic analysis of ADHD and control children using nonlinear dynamical features of EEG signals.". *Journal of integrative neuroscience*, 17(1):17–30, 2018.

[8] A. E. Alchalabi, S. Shirmohammadi, A. N. Eddin, and M. Elsharnouby. "Focus: detecting ADHD patients by an EEG-based serious game.". *IEEE Transactions on Instrumentation and Measurement*, 67(7):1512–1520, 2018.

[9] R. Catherine Joy, S. Thomas George, A. Albert Rajan, and M. S. P. Subathra. "Detection of ADHD from EEG signals using different entropy measures and ANN.". *Clinical EEG and Neuroscience*, 53 (1):12–23, 2022.

[10] B. TaghiBeyglou, A. Shahbazi, F. Bagheri, S. Akbarian, and M. Jahed. "Detection of ADHD cases using CNN and classical classifiers of raw EEG.". *Computer Methods and Programs in Biomedicine Update*, 2:100080, 2022.

[11] M. Maniruzzaman, J. Shin, M. A. M. Hasan, and A. Yasumura. "Efficient feature selection and machine learning based ADHD detection using EEG signal.". *Computers, Materials & Continua*, 72(3): 5179–5195, 2022.

[12] A. Alim and M. H. Imtiaz. "Automatic identification of children with ADHD from EEG brain waves.". *Signals*, 4(1):193–205, 2023.

[13] A. M. Nasrabadi, A. Allahverdy, M. Samavati, and M. R. Mohammadi. "EEG data for ADHD/Control children.". *IEEE Dataport*, 10: 978–980, 2020.

[14] D. J. Stein, K. A. Phillips, D. Bolton, K. W. M. Ful-ford, J. Z. Sadler, and K. S. Kendler. "What is a mental/psychiatric disorder? From DSM-IV to DSM-V.". *Psychological medicine*, 40(11):1759–1765, 2010.

[15] A. Ekhlasi, A. M. Nasrabadi, and M. R. Mohammadi. "Direction of information flow between brain regions in ADHD and healthy children based on EEG by using directed phase transfer entropy.". *Cognitive Neurodynamics*, 15(6):975–986, 2021.

[16] S. Phadikar, N. Sinha, and R. Ghosh. "Automatic eyeblink artifact removal from EEG signal using wavelet transform with heuristically optimized threshold.". *IEEE Journal of Biomedical and Health Informatics*, 25(2):475–484, 2020.

[17] S. G. Mallat. "A Theory for Multiresolution Signal Decomposition: The Wavelet Repre-sentation.". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7):674–693, 1989.

[18] S. Haykin and Z. Chen. "The cocktail party problem.". *Neural computation*, 17(9):1875–1902, 2005.

[19] A. Hyvärinen, J. Hurri, and P. O. Hoyer. "Independent component analysis.". *Springer London*, 2009.

[20] H. Abdi. "The eigen-decomposition: Eigenvalues and eigenvectors.". *Encyclopedia of Measurement and Statistics*, pages 304–308, 2007.

[21] A. D. Dongare, R. R. Kharde, and A. D. Kachare. "Introduction to artificial neural network.". *International Journal of Engineering and Innovative Technology*, 2(1):189–194, 2012.

[22] O. Kramer. "K-nearest neighbors.". *Dimensionality Reduction with Unsupervised Nearest Neighbors*, 2013.