Indonesian Journal of Electrical Engineering and Computer Science
Vol. 37, No. 3, March 2025, pp. 1580~1587
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v37.i3.pp1580-1587

Attention deficit and hyperactivity disorder classification in quantitative EEG signals using machine learning algorithms

Syifani Ihfadza Aliyah[1], Sastra Kusuma Wijaya[1], Yetty Ramli[2]
[1]Department of Physics, Faculty of Mathematics and Natural Sciences, Universitas Indonesia, Depok, Indonesia
[2]Department of Neurology, Faculty of Medicine, Universitas Indonesia, Depok, Indonesia

Article Info

Article history:

Received Nov 21, 2023
Revised Oct 2, 2024
Accepted Oct 7, 2024

Keywords:

ADHD
Machine learning classification
Principal component analysis
Quantitative EEG
Wavelet transform

ABSTRACT

Attention deficit and hyperactivity disorder (ADHD) classification method as a quantitative observation has been continually improved to assist medical practitioners. Currently, machine learning algorithms such as k-nearest neighbors (KNN), multilayer perceptron (MLP), and support vector machine (SVM) are widely used. This study proposed a feature extraction method for quantitative electroencephalography (qEEG) data derived from the continuous wavelet transform (CWT) to classify children with ADHD versus healthy subjects. Subsequently, this study compared the performance of the classification pipeline before and after the implementation of principal component analysis (PCA) on the features prior to processing with machine learning algorithms. The results revealed that the overall performance of the classifiers consistently improved after the implementation of PCA. The results highlight the varying impact of PCA on classifier performance, with KNN showing an improvement in testing accuracy from 61.84% to 69.21% following PCA implementation, while the other classifiers showed deterioration in performance. These findings suggest that while PCA may be beneficial for some classifiers, its impact on performance varies depending on the specific characteristics of the dataset and the classifier utilized. Moreover, this study provides insight for future implementation of the classification method for ADHD patients across a more specific clinical range of the spectrum.

Corresponding Author:

Sastra Kusuma Wijaya
Department of Physics, Faculty of Mathematics and Natural Sciences, Universitas Indonesia
Depok, West Java, Indonesia
Email: skwijaya@sci.ui.ac.id

1. INTRODUCTION

Attention deficit and hyperactivity disorder (ADHD) is a neurodevelopmental disorder that affects millions of children and adults worldwide, and is characterized by disruptive inattention, excessive activity, and impulsive actions [1]. ADHD impacts around 5–7% of children and 2–5% of adults globally [2]. The most common methods for psychiatrists, pediatricians, neurologists, and psychologists on ADHD are clinical observations [3]. However, in recent years, methods based on brain electrical signals through spectral analysis have aided healthcare professionals in the diagnosis of ADHD [4]. Electroencephalography (EEG) is a non-invasive method for acquiring electrical activity originating from neurons in the brain that can be measured through the scalp [5]. This procedure involves placing electrodes on the scalp to record EEG signals. EEG has proven to be valuable tool in assisting the quantitative diagnosis of ADHD as it provides information about brain electrical activity [6].

EEG has evolved into quantitative EEG (qEEG), where EEG signals are mapped for their brain activity patterns using digital signals and mathematical algorithms [7]. In 2005, Niedermeyer [8] highlighted the role of qEEG in understanding brain activity patterns and their associations with cognitive disorders. The qEEG signal is a useful tool for measuring and analyzing brain activity that plays a crucial role in identifying specific patterns that indicate certain symptoms and aiding in the treatment of various mental health disorders, such as ADHD [9]. The frequency range of EEG signals vary, but the most common are delta (0.5-4 Hz), theta (4-7 Hz), alpha (8-12 Hz), sigma (12-16 Hz), and beta (13-30 Hz) [10]. The study by Barry *et al.* [11] in 2007 found that children with ADHD exhibited lower brain activity in the beta frequency range (13-21 Hz) in the frontal area. This suggests that children with ADHD experience disruptions in their ability to focus attention and regulate behavior. In 2012, Loo and Makeig [12] found that brain activity patterns of children with ADHD differed from those categorized as neurotypical. They discovered that brain activity in children with ADHD showed lower beta frequency and higher theta frequency activity in the frontal and temporal areas, whereas neurotypical children exhibited more stable brain activity across the entire brain.

The continuous wavelet transform (CWT) has been used to obtain detailed information about time series signals such as qEEG [13]. This methods generates time-frequency and nonlinear features that can serves as a quantitative tool to detect the activity of human brain [14]. QEEG studies in individuals with ADHD have indicated specific differences on brain activity such as lower brain activity levels and distinct frequencies [15]. Various machine learning algorithms, such as k-nearest neighbor (KNN), convolutional neural network (CNN), random forest, and support vector machine (SVM), have been applied to classify EEG signals as normal or indicative of ADHD or mental disorders such as depression [16]-[18]. In applying these algorithms, EEG signals data are collected and preprocessed before being input into the model. The model is then trained using qEEG signal data from individuals with and without ADHD, enabling it to distinguish between the two groups [19]. The use of machine learning algorithms to process qEEG signals offers numerous advantages, such as the ability to handle large datasets with high accuracy and the ability to address individual variability in the data [20]. Principal component analysis (PCA) is one of the most common techniques paired with machine learning algorithms to improve data variability, thereby enhance the classifier performance [21]. These studies demonstrate that machine learning can assist medical practitioner in the early clinical diagnosis of ADHD [22].

2. METHOD

This study involved six stages, as illustrated in Figure 1. The stages included qEEG data acquisition, preprocessing, processing, PCA, classification model selection, and model evaluation. During EEG data acquisition, the focus was on gathering data from various electrode placements as sources of EEG signals. During the preprocessing stage, the EEG signals were filtered to remove any unwanted noise. In the next step, the processing stage involved of the CWT, followed by feature extraction from the data where the qEEG data was transformed to extract features needed for the classifier models. The processed qEEG data now consisting of these features, then went through the PCA stage, where PCA was performed to reduce the low variance data and emphasize higher-variance features. Subsequently, the machine learning classification employed three types of classifier models: KNN, multilayer perceptron (MLP), and SVM. After selecting the classification model, the study proceeded to the model evaluation stage, where the performance metrics were examined to assess the effectiveness of the models.

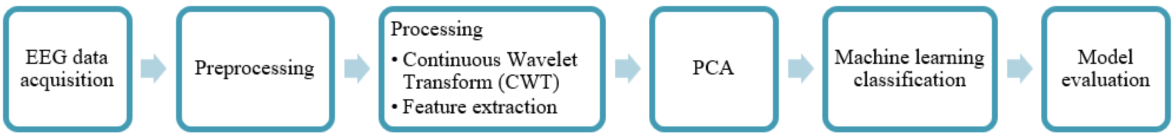


Figure 1. Flow chart of the research stages

2.1. EEG data acquisition

This study used secondary datasets from two sources; Mohammadi *et al.* [23] dataset and the Pereda *et al.* [24] dataset. Both datasets were collected through a similar experiment, which gathered eyes closed and eyes opened qEEG signals from the participants while performing visual and cognitive tasks. Detailed information regarding both datasets is provided in Table 1. The Mohammadi *et al.* [23] dataset was obtained from the IEEE Dataport platform uploaded by Mohammadi *et al.* [23] from the Tehran University of Medical

Sciences, Tehran. The qEEG signal data of the dataset of Mohammadi *et al.* [23] were recorded using an SD-C24 device with a 24-bit ADC at a sampling rate of 128 Hz from 19 channels located at Fz, Cz, Pz, C3, T3, C4, T4, Fp1, Fp2, F3, F4, F7, F8, P3, P4, T5, T6, O1, and O2. Mohammadi *et al.* [23] conducted a study with 60 participants, including 30 subjects with ADHD and 30 healthy subjects. The Pereda *et al.* [24] dataset was sourced from the Figshare platform and uploaded by Pereda *et al.* [24] to the University of La Laguna, Spain. EEG eyes closed and eyes opened data from the 33 participants were recorded using a Nihon Kohden Neurofax EEG-9200 device with a 16-bit ADC at a sampling rate of 256 Hz from eight channels located at Fp1, Fp2, C3, C4, T3, T4, O1, and O2.

Table 1. Quantitative EEG datasets information

Source	Participants		Sampling rate	ADC Bit
	ADHD	Control		
Mohammadi *et al.* [23]	30 subjects (22L; 8P; 9.62 ± 1.75 years old)	30 subjects (25L; 5P; 9.85 ± 1.77 years old)	256 Hz	24 bit
Pereda *et al.* [24]	19 subjects (19L; 8.50 ± 1.74 years old)	14 subjects (14L; 8.21 ± 1.74 years old)	128 Hz	16 bit

2.2. Preprocessing

Raw qEEG data must be preprocessed to ensure the data are free from noise and signal artifacts. Both datasets were converted into data frames and normalized from digital signals to analog voltage data with microvolt magnitudes as expressed in (1). Subsequently, the sensor position montage was arranged according to the international 10-20 standard montage system for 19 and 8 channels. The output of the conversion from a digital signal to analog requires knowledge of the device reference voltage and analog-to-digital converter bit depth, as provided in the dataset information.

$$Voltage\ Reading\ = \frac{Digital\ Reading}{2^{N}-1} \cdot\ Reference\ Voltage \tag{1}$$

2.3. Processing

2.3.1. Continuous wavelet transforms

The wavelet transform is an orthogonal function that serves as a tool to decompose data, functions, or operators into different frequency components and then analyze each component with a resolution adapted to its scale. A signal in time-domain is processed using the wavelet transform in frequency-domain signals within a specified frequency range, generate time-frequency coefficients. One type of wavelet transformation, the CWT, involves processing a signal with a specified continuous frequency rather than discrete frequency intervals. The resulting wavelet coefficient was used as the basis for feature extraction for each channel per subject [25].

CWT enables the time-frequency analysis of a signal, allowing the identification and characterization of features at different scales within the signal. This transformation is based on the use of a scaled continuous and shifted wavelet function [26]. The CWT of a signal f(t) with respect to a wavelet function $\psi(t)$ is defined as the multiplication of the original signal and shifted and scaled wavelet functions in the time domain. The mathematical formula for the CWT is expressed in (2) as follows:

$$CWT\ (a,b) =\ |\alpha|^{-\frac{1}{2}} \cdot \int [f\ (t)] \cdot \psi \left[\frac{t-b}{\alpha}\right] dt \tag{2}$$

where a represents the scale factor, b represents the translation factor, f(t) represents the original signal, and $\psi \left[\frac{t-b}{\alpha}\right]$ represents the scaled and shifted wavelet functions. The type of wavelet used in this study was the Morlet wavelet with the mathematical formula in (3) with:

$$\psi(t) = \exp^{-\frac{t^2}{2}} \cos(5t) \tag{3}$$

2.3.2. Feature extraction

Feature extraction involved the extraction of the features from the EEG signals of each channel per participant in both datasets. The participant qEEG data were divided into frequency bands using the wavelet method with 1-4 Hz for delta, 4-8 Hz for theta, 8-16 Hz for alpha, 16-32 Hz for beta, and 32-64 Hz for gamma bands. The CWT method generates two coefficients:those with low-frequency information are called approximation coefficients, and those with high-frequency information are called detail coefficients. The coefficients from the CWT were computed mathematically to obtain four features; the average of signal power or energy, signal entropy, mean, and standard deviation. Therefore, the total features generated from each participant would be the sum of all channels multiplied by the two coefficients and four features, resulting in a total of 152 features per participant from the Mohammadi *et al.* [23] dataset with 19 channels and 64 features per participant from the Pereda *et al.* dataset. The total number of features for each participant was multiplied by five frequency bands. Therefore, the number of inputs for the machine learning models was 300 for the Mohammadi *et al.* [23] dataset and 165 for the Pereda *et al.* [24] dataset.

2.4. Principal component analysis

PCA is capable of reducing clusters of features with low variance, thereby generating higher variance that could significantly increase the performance of classification models [27]. The eigenvalue represents the amount of variance captured by each principal component, indicating the significance of the corresponding eigenvector in describing the variability of the data. The Kaiser-Guttman rule suggests retaining principal components with eigenvalues greater than 1, indicating that they explain more variance than principal components with eigenvalues below the line; they are considered significant for analysis [28]. The resulting components of the PCA were visualized using a scree plot to show how the data variance changed when PCA was applied. Kaiser's line helps ensure the retention of representative principal component of the broader population, which can be relied upon for further analysis and interpretation due to their high variance data. This line is crucial for reliable data analysis because it ensures both the variance and reliability of the feature reduction process.

2.5. Machine learning classification

The classification of subjects falls under the category of supervised learning in the machine learning area [29]. There are various supervised learning algorithms, such as KNN, MLP, and SVM. KNN is the fundamental concept of finding k-nearest data points from the data to be classified and determining the class label that most frequently appears among the k-neighbors [30]. KNN is particularly advantageous when dealing with complex and nonlinear patterns in data, making it suitable for tasks where the underlying distribution is not well understood or is highly irregular. The MLP is a type of neural network with an input layer, hidden layers, and output layer. It utilizes weights and biases to transform input data through multiple layers and learns complex patterns during training. MLPs are commonly employed for tasks such as classification and regression in machine-learning applications, capable of learning complex feature representations and achieve a high predictive accuracy, particularly in large-scale datasets with diverse features [31]. SVM are among the most favored machine learning algorithms for classification and regression. SVM aims to find the best hyperplane that can separate the two classes in the given data. This hyperplane is chosen by maximizing the margin, which is the distance between the hyperplane and the nearest points from each class [32]. SVM is especially effective when managing data with numerous dimensions and can handle non-linear relationships between features using kernel functions. Its ability to find an optimal hyperplane and maximize the margin makes it robust to overfitting and ensures a good generalization performance.

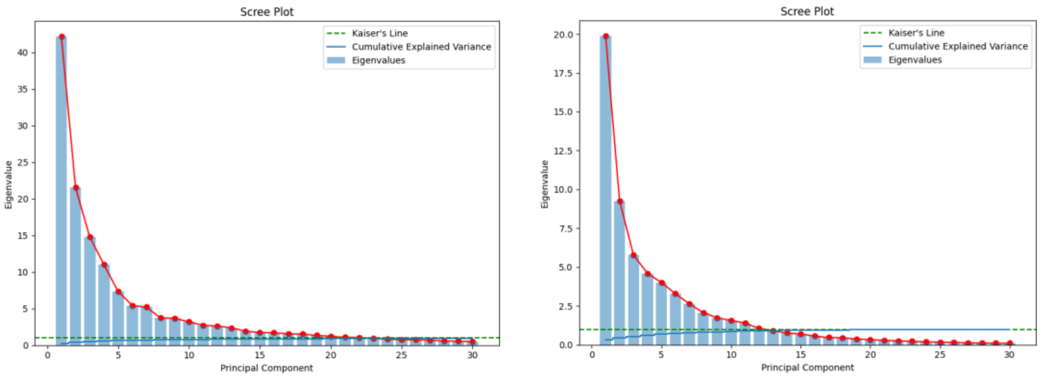
2.6. Model evaluation

Model evaluation is the last step of this research method, and the performance of the machine learning models was evaluated based on the accuracy of the testing and training of feature datasets. The performances of the three classifiers were evaluated using confusion matrices. Additionally, the accuracies of the classifiers before and after the implementation of PCA were compared [33].

3. RESULTS AND DISCUSSION

3.1. Principal component analysis

The features obtained from the Mohammadi *et al.* [23] and Pereda *et al.* [24] datasets were reduced from 152 and 64 features, respectively, and were reduced using PCA to improve model performance. The reduced features were visualized using a scree plot, as illustrated in Figure 2, to show how the data variance changed when PCA was applied. The relationship between the principal components and eigenvalues showed that the number of eigenvalues tended to decrease when the principal component increased, representing a decrease in the amount of variance. Therefore, Kaiser's line at an eigenvalue of one was utilized to determine the principal components with high-variance. Figure 2(a) illustrates the PCA scree plot of the Mohammadi *et al.* [23] dataset, which shows that there are 22 high-variance components above the Kaiser's line. Figure 2(b) shows that there are 12 high variance components derived from the Pereda *et al.* [24] dataset. These components were used as features for classifiers to improve the performance modeling. Although processing additional components with lower variance may lengthen the training process, it does not substantially alter the model results and can therefore be disregarded.

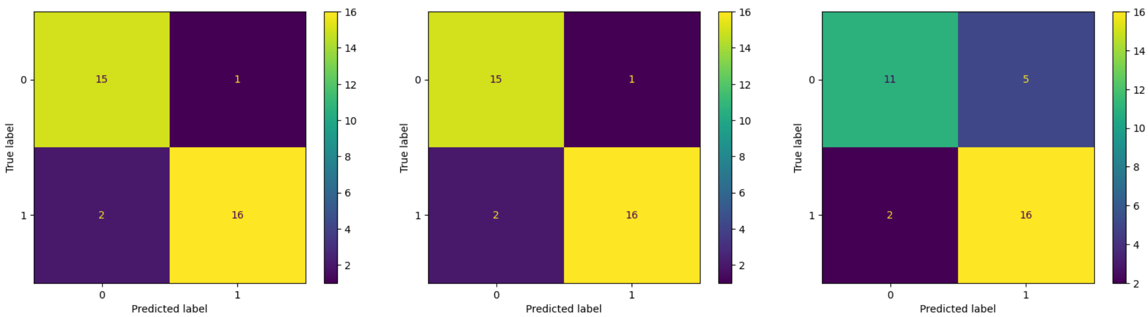


(a) (b)

Figure 2. Scree plot of (a) 22 components and (b) 12 components

3.2. Machine learning classification

The reduced data were then randomly partitioned into testing and training sets at ratios of 20% and 80%, respectively. Subsequently, the SVM, KNN, and MLP classifiers were utilized to classify ADHD. Figure 3 shows the confusion matrices containing the four values of the model prediction outcome from the three classifiers. True-positive refers to the number of patients correctly diagnosed with ADHD and true-negative refers to the number of patients correctly diagnosed as neurotypical. False-positive describes the number of patients incorrectly diagnosed with ADHD and false-negative describes the number of patients incorrectly diagnosed as neurotypical. A positive value is represented by "1" and a negative value is represented by "0". Figure 3(a) illustrates the confusion matrix of the SVM, Figure 3(b) the KNN, and Figure 3(c) the MLP. The performance of a classifier is considered good when the confusion matrix exhibits a larger output in the true-positive and true-negative cells.



(a) (b) (c)

Figure 3. Confusion matrix of classifiers: (a) SVM, (b) KNN, and (c) MLP

4. MODEL EVALUATION

The classifiers used in the model were the MLP, SVM, and KNN. Three classifiers were employed to compare the models for classifying qEEG data and acquire the most accurate and robust model for classification. The Mohammadi *et al.* [23] and Pereda *et al.* [24] datasets were combined to improve model generalization and increase the amount of training data. The combined dataset was then processed through classifiers both with PCA and without PCA to evaluate the effect of PCA. The impact of PCA feature reduction on the performance of classifiers was analyzed from the provided data in Tables 2 and 3. Across the three classifiers, applying PCA resulted in a decrease in accuracy for both training and testing datasets compared to scenarios without PCA as shown in Table 2. This suggests that the PCA feature reduction might not have effectively captured the underlying patterns in the data for this study. When PCA was not applied, the highest training accuracy of 95.68% was achieved by the MLP classifier, but its testing accuracy dropped significantly to 59.21%, indicating potential overfitting. Similarly, a high training accuracy of 87.31% was achieved by the SVM, but it exhibited a lower testing accuracy of 53.95%, suggesting some degree of overfitting. In contrast, KNN displayed moderate training accuracy of 74.07% and the highest testing accuracy of 61.84% among the three classifiers without PCA. However, after PCA was implemented, the accuracy of all classifiers decreased except the KNN. The MLP classifier still achieved the highest training accuracy of 86.27% but experienced a significant drop in testing accuracy with 53.95%. The accuracy of SVM also declined, reaching 80.17% in training and 51.31% in testing with PCA applied. Interestingly, the KNN testing accuracy improved to 69.21% with PCA, indicating a potential improvement in generalization. While MLP consistently outperformed SVM and KNN in term of training accuracy but suffered from overfitting issues. SVM, on the other hand, demonstrated greater robustness to overfitting but showed lower overall accuracy overall, particularly when PCA was used. Ultimately, the choice between applying PCA or not depends on the specific requirements of the classification task and the trade-offs between feature reduction and accuracy.

Table 2. Performance comparison of classifiers without PCA

Classifier	Parameters	Accuracy	
		Training (%)	Testing (%)
MLP	hidden layer (34, 34, 34)	95.68	59.21
SVM	34 35 support vectors	87.31	53.95
KNN	4 neighbors	74.07	61.84

Table 3. Performance comparison of classifiers with PCA

Classifier	Parameters	Accuracy	
		Training (%)	Testing (%)
MLP	hidden layer (34, 34, 34)	86.27	53.95
SVM	34 35 support vectors	80.17	51.31
KNN	4 neighbors	77.34	69.21

The comparison of classifiers for ADHD diagnosis from other research, as shown in Table 4, revealed varying performance levels across different stimulation tasks. For instance, in a study by Alchalabi *et al.* [34], the SVM classifier achieved an exceptional accuracy of 98.6% during a focused gaming task. Similarly, Mohammadi *et al.* [23] reported an accuracy of 93.7% using an MLP classifier during a visual cognitive task, whereas Yang *et al.* [35] obtained an accuracy of 89.3% with a KNN classifier during a motor task with interference. Consistent with these findings, this research focused on utilizing EEG signals for ADHD diagnosis through quantitative analysis and machine learning algorithms, while also exploring the efficacy of signal attributes such as power, entropy, average, and standard deviation using signal processing techniques like CWT. These results provide valuable insights into the potential of EEG-based classification methods for ADHD diagnosis and highlight the importance of further research in this area.

Table 4. Classifier comparisons on classifying ADHD

Stimulation	Classifier	Accuracy	Reference
Focused gaming	SVM classifier	98.6%	Alchalabi *et al.* [34]
Visual cognitive task	MLP classifier	93.7%	Mohammadi *et al.* [23]
Motoric task with interference	KNN classifier	89.3%	Yang *et al.* [35]

5. CONCLUSION

This study demonstrated the feasibility of using qEEG signals for ADHD classification through quantitative analysis and machine learning algorithms, extracting features such as power, entropy, average, and standard deviation via the CWT. PCA aided in the extraction of high variance features, reducing overfitting and enhancing classification accuracy. However, the impact of PCA varied depending on the dataset and classifier utilized. Notably, the SVM classifier outperformed the others, achieving a 53.95% testing accuracy despite its lower training accuracy of 87.31%, showcasing robust generalization. Conversely, the MLP classifier's high training accuracy of 95.68% dropped significantly to 59.21% in testing, indicating potential overfitting issues. The KNN classifier performed competitively, with a 61.84% testing accuracy,which notably improved to 69.21% with PCA, suggesting enhanced generalization.

This study offers valuable insights for optimizing ADHD diagnosis using qEEG signals, emphasizing classifier robustness and generalization. The findings could assist healthcare professionals in improving diagnosis accuracy and quantifying ADHD within a clinical spectrum. Future research in refining machine learning hyperparameters could further enhance classifier performance, contributing to more effective ADHD classification methods tailored to the specific clinical range of the disorder.

ACKNOWLEDGEMENTS

This study was supported by the Universitas Indonesia Research Fund (Grant PUTI UI Q3 No. NKB-235/UN2.RST/HKP.05.00/2023).

REFERENCES

[1] American Psychiatric Association, "Diagnostic and statistical manual of mental disorders (DSM-5)," *American Psychiatric Pub*, 2013.
[2] G. Polanczyk, "The worldwide prevalence of ADHD: a systematic review and metaregression analysis," *American Journal of Psychiatry*, vol. 164, no. 6, p. 942, Jun. 2007, doi: 10.1176/appi.ajp.164.6.942.
[3] T. H. Eom and Y. H. Kim, "Clinical practice guidelines for attention-deficit/hyperactivity disorder: recent updates," *Clinical and Experimental Pediatrics*, vol. 67, no. 1, pp. 26–34, Jan. 2024, doi: 10.3345/cep.2021.01466.
[4] A. Alim and M. H. Imtiaz, "Automatic identification of children with ADHD from EEG brain waves," *Signals*, vol. 4, no. 1, pp. 193–205, Feb. 2023, doi: 10.3390/signals4010010.
[5] S. Beniczky and D. L. Schomer, "Electroencephalography: basic biophysical and technological aspects important for clinical applications," *Epileptic Disorders*, vol. 22, no. 6, pp. 697–715, Dec. 2020, doi: 10.1684/epd.2020.1217.
[6] R. Garg and R. Verma, "EEG based classification techniques for ADHD diagnosis: a review," *Journal of medical systems*, vol. 40(4), pp. 305–314, 2016.
[7] M. Nuwer, "Assessment of digital EEG, quantitative EEG, and EEG brain mapping: Report of the American Academy of Neurology and the American Clinical Neurophysiology Society," *Neurology*, vol. 49, no. 1, pp. 277–292, Jul. 1997, doi: 10.1212/WNL.49.1.277.
[8] E. Niedermeyer and F. L. da Silva, "Electroencephalography: basic principles, clinical applications, and related fields (5th ed.)," *Lippincott Williams & Wilkins*, 2005.
[9] R. A. Barkley, "Attention-deficit hyperactivity disorder: a handbook for diagnosis and treatment (3rd ed.)," *Guilford Press*, 2006.
[10] C. S. Nayak and A. C. Anilkumar, "EEG Normal Waveforms," 2024.
[11] R. J. Barry, A. R. Clarke, S. J. Johnstone, and C. R. Brown, "EEG differences in children between eyes-closed and eyes-open resting conditions," *Clinical Neurophysiology*, vol. 120, no. 10, pp. 1806–1811, 2009, doi: 10.1016/j.clinph.2009.08.006.
[12] S. K. Loo and S. Makeig, "Clinical Utility of EEG in Attention-Deficit/Hyperactivity Disorder: A Research Update," *Neurotherapeutics*, vol. 9, no. 3, pp. 569–587, Jul. 2012, doi: 10.1007/s13311-012-0131-z.
[13] S. Altun, A. Alkan, and H. Altun, "Automatic diagnosis of attention deficit hyperactivity disorder with continuous wavelet transform and convolutional neural network," *Clinical Psychopharmacology and Neuroscience*, vol. 20, no. 4, pp. 715–724, Nov. 2022, doi: 10.9758/CPN.2022.20.4.715.
[14] A. Khaleghi, P. M. Birgani, M. F. Fooladi, and M. R. Mohammadi, "Applicable features of electroencephalogram for ADHD diagnosis," *Research on Biomedical Engineering*, vol. 36, no. 1, pp. 1–11, Mar. 2020, doi: 10.1007/s42600-019-00036-9.
[15] R. J. Chabot and G. Serfontein, "Quantitative electroencephalographic analysis of boys with attention-deficit hyperactivity disorder: selected findings," *Journal of Child Neurology*, vol. 11, no. 2, pp. 81–89, 1996.
[16] Y. C. Chang, T. Y. Chen, C. C. Lu, and H. C. Chu, "A review of EEG-based classification techniques for attention deficit hyperactivity disorder," *Frontiers in neuroengineering, 11, 22.*, 2018.
[17] R. A. Apsari and S. K. Wijaya, "Electroencephalographic spectral analysis to help detect depressive disorder," in *IBIOMED 2020 - Proceedings of the 37th International Conference on Biomedical Engineering*, Oct. 2020, pp. 13–18, doi: 10.1109/IBIOMED50285.2020.9487614.
[18] 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*, vol. 6, no. 2, pp. 66–73, 2016, doi: 10.1007/s13534-016-0218-2.
[19] M. Jalili, M. A. R. Kordestani, and H. Adeli, "EEG-based classification of attention deficit hyperactivity disorder using deep learning: A review," *Journal of medical systems*, vol. 44(11), p. 547, 2020.
[20] Pawan and R. Dhiman, "Machine learning techniques for electroencephalogram based brain-computer interface: a systematic literature review," *Measurement: Sensors*, vol. 28, p. 100823, Aug. 2023, doi: 10.1016/j.measen.2023.100823.
[21] R. Kottaimalai, M. P. Rajasekaran, V. Selvam, and B. Kannapiran, "EEG signal classification using principal component analysis with neural network in brain computer interface applications," in *2013 IEEE International Conference on Emerging Trends in Computing, Communication and Nanotechnology, ICE-CCN 2013*, Mar. 2013, pp. 227–231, doi: 10.1109/ICE-CCN.2013.6528498.
[22] T. Chen, I. Tachmazidis, S. Batsakis, M. Adamou, E. Papadakis, and G. Antoniou, "Diagnosing attention-deficit hyperactivity disorder (ADHD) using artificial intelligence: a clinical study in the UK," *Frontiers in Psychiatry*, vol. 14, Jun. 2023, doi: 10.3389/fpsyt.2023.1164433.
[23] 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*, vol. 6, no. 2, pp. 66–73, May 2016, doi: 10.1007/s13534-016-0218-2.
[24] E. Pereda, M. García-Torres, B. Melián-Batista, S. Mañas, L. Méndez, and J. J. González, "The blessing of dimensionality: feature selection outperforms functional connectivity-based feature transformation to classify ADHD subjects from EEG patterns of phase synchronisation," *PLoS ONE*, vol. 13, no. 8, p. e0201660, Aug. 2018, doi: 10.1371/journal.pone.0201660.
[25] I. Daubechies, *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics, 1992.

[26] S. Mallat, *A Wavelet Tour of Signal Processing*, 2 nd. 1999.
[27] F. Pedregosa *et al.*, "Scikit-learn: machine learning in python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011, [Online]. Available: http://jmlr.csail.mit.edu/papers/v12/pedregosa11a.html%5Cnhttp://arxiv.org/abs/1201.0490.
[28] H. F. Kaiser, "On cliff's formula, the kaiser-guttman rule, and the number of factors," *Perceptual and Motor Skills*, vol. 74, no. 2, pp. 595–598, Apr. 1992, doi: 10.2466/pms.1992.74.2.595.
[29] M. Mohri, A. Rostamizadeh, and A. Talwalkar, *Foundations of Machine Learning*, 2 nd. 2018.
[30] H. Abbas, M. Hussain, N. Shahid, and S. Raza, "ADHD detection and classification using KNN and SVM based on EEG signals," *International Journal of Advanced Computer Science and Applications, 9(10)*, vol. 9, no. 10, pp. 491–497, 2018.
[31] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2009.
[32] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, Sep. 1995, doi: 10.1007/bf00994018.
[33] N. Arora, S. Srivastava, R. Agarwal, V. Mehndiratta, and A. Tripathi, "Diabetes mellitus prediction using machine learning within the scope of a generic framework," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 32, no. 3, pp. 1724–1735, Dec. 2023, doi: 10.11591/IJEECS.V32.I3.PP1724-1735.
[34] 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*, vol. 67, no. 7, pp. 1512–1520, Jul. 2018, doi: 10.1109/TIM.2018.2838158.
[35] J. Yang, W. Li, S. Wang, J. Lu, and L. Zou, "Classification of children with attention deficit hyperactivity disorder using PCA and k-nearest neighbors during interference control task," 2016, pp. 447–453.

BIOGRAPHIES OF AUTHORS

Syifani Ihfadza Aliyah is a researcher at Biomedical Instrumentation Research Group of Universitas Indonesia. She received her bachelor's degree in physics with the specialization in systems and instrumentation physics. Her research areas are signal processing, electroencephalography, and classification program with supervised learning. She is a recipient of several academics' awards such as IISMA Scholarship 2022 sponsored by the Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia to Western University, Canada. Also, she received a PUTI Q3 2023 research grant from Universitas Indonesia to research on ADHD classification from qEEG signal using machine learning. Her research interest includes signal processing, analysis, and classification of qEEG data. She can be contacted through her email: syifani.ihfadza@ui.ac.id.

Dr. Sastra Kusuma Wijaya earned his Doctorate from Okayama University, Japan, specializing in Natural Sciences and Engineering. His academic dedication spans diverse domains, with current interests in Physics Instrumentation, data acquisition, and biosignal processing. Dr. Wijaya is a pioneer in developing measurement tools, utilizing cutting-edge technologies for precision, and broadening experimental methodologies. His pivotal contributions to data acquisition involve innovative approaches in capturing, processing, and interpreting data across scientific domains. Particularly focused on biosignal processing, especially in biomedical signals, he employs advanced techniques to unravel complexities in biological systems. He leads research integrating machine learning into biosignal processing, enhancing accuracy, and efficiency in biomedical data analysis. He can be contacted at email: skwijaya@sci.ui.ac.id.

Dr. Yetty Ramli received her Doctorate in Biomedicine from Institut Pertanian Bogor, Indonesia. She has been serving at the University of Indonesia's Neurology Department since 2003. Dr. Ramli holds the esteemed position of Consultant in Pediatric Neurology, making significant contributions to patient care at Cipto Mangunkusumo Hospital, a prominent government institution. Her active participation in various conferences and events, including presentations at prestigious platforms like Neuroradiology Australia and the Brain Injury Third Three Continents Conference in Singapore, underscores her commitment to addressing challenges in pediatric neurology. Her research interest includes neurology, neurological theraphy, brain injury, and stem cell. She can be contacted at email: yettyramli@yahoo.com.