IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 14, No. 3, June 2025, pp. 2302~2314
ISSN: 2252-8938, DOI: 10.11591/ijai.v14.i3.pp2302-2314

Hybrid method for optimizing emotion recognition models on electroencephalogram signals

I Made Agus Wirawan[1], Kadek Yota Ernanda Aryanto[1], I N. Sukajaya[2], Ni Nyoman Mestri Agustini[3], Dewi Arum Widhiyanti Metra Putri[4]
[1]Data Science Lab, Department of Informatic Engineering, Faculty of Engineering and Vocational, Universitas Pendidikan Ganesha, Singaraja, Indonesia
[2]Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Pendidikan Ganesha, Singaraja, Indonesia
[3]Department Neurology and Medical Rehabilitation, Faculty of Medicine, Universitas Pendidikan Ganesha, Singaraja, Indonesia
[4]Education of Children, Adolescents, Adults and Their Characteristics Lab, Faculty of Education, Universitas Pendidikan Ganesha, Singaraja, Indonesia

Article Info

Article history:

Received Aug 30, 2024
Revised Feb 3, 2025
Accepted Mar 15, 2025

Keywords:

3D cube
Continuous capsule network
Differential entropy
Electroencephalogram
Emotion recognition
Modified weighted mean filter
Relative difference

ABSTRACT

Two critical factors that need to be studied in emotion recognition are the differences in electroencephalogram (EEG) signal patterns caused by participant characteristics and EEG signals with spatial information. These factors significantly affect the resulting accuracy. The model proposed in this study can consider these factors. This model consists of the modified weighted mean filter method for the basic EEG signal smoothing process, the differential entropy method for the feature extraction process, the relative difference method for the baseline reduction, the 3D cube method for feature representation, and the continuous capsule network method for the classification process. Based on testing on three public datasets, this hybrid method can overcome factors affecting emotion recognition accuracy. This statement is based on the accuracy produced by this model, which outperformed the accuracy validated in previous studies.

Corresponding Author:

I Made Agus Wirawan
Data Science Lab, Department of Informatic Engineering, Faculty of Engineering and Vocational
Universitas Pendidikan Ganesha
Udayana Road, Singaraja, Bali, Indonesia
Email: imade.aguswirawan@undiksha.ac.id

1. INTRODUCTION

It is important to recognize emotional reactions in humans to understand human mental states and performance [1], [2]. Some emotion recognition approaches can be performed externally. However, internal recognition of emotions, primarily via electroencephalogram (EEG) signals, has several advantages, such as: i) having spatial information that represents human affective experiences [3], and ii) respond to early changes in human emotional reactions [4], [5]. Although various methods have been proposed to improve the performance of artificial intelligence (AI) models in recognizing emotions, interference in the EEG signal and loss of spatial information from the EEG signal can result in low model accuracy [6]. Therefore, examining several appropriate methods to produce optimal AI models to solve these two problems is essential.

Generally, two essential aspects are considered in developing an emotion recognition model on EEG signals: participant and EEG signal characteristics [6]. Participant characteristics include personality traits, gender, culture, and intellectual abilities. This aspect produces different EEG signal pattern variances for each participant. If this EEG signal pattern variance is not addressed, it will impact low emotion recognition accuracy [7]. To overcome this problem, Wirawan *et al.* [8] proposed a baseline reduction process using the relative difference method. The baseline reduction process reduces the experimental signal features with the baseline signal features [7]. The experimental signal is a signal that represents an emotional condition. Meanwhile, the baseline signal represents a neutral condition [9]. Ideally, neutral conditions have lower signal amplitude values than emotional conditions [10], [11]. However, the baseline signal does not represent a neutral condition because the baseline and experimental signals still have the same amplitude values [12]–[14]. The high amplitude of the baseline signal is caused by internal and external artefacts/interference [9], [15]–[17].

Wirawan *et al.* [18] have successfully developed and implemented the modified weighted mean filter method to optimize the baseline reduction process with the relative difference method. However, this model needs to consider the spatial characteristics of EEG signals [18]. Several researchers have studied the aspects of EEG signal characteristics from the feature extraction process, feature representation, and classification. In the feature extraction process, the differential entropy method can represent spatial information and is stable in emotion recognition compared to other features [7], [19]–[21]. In the feature representation process, the 3D cube method can maintain spatial information between channels and between theta, alpha, beta, and gamma frequencies [22]. In terms of classification, several researchers have studied the multi-class support vector machine [23], graph regularized extreme learning machine [24], convolution neural network [25], [26], and the combination of long short-term memory with convolution neural network [27]. However, of the several classification methods studied, the capsule network method has been proven to represent spatial information in the classification process [22]. Wirawan *et al.* [22] proposed a continuous capsule network method to preserve spatial information from EEG signals for classification. However, the noise in EEG signals in this model has yet to be overcome.

The proposed problem-solving approach for the first problem is to apply the modified weighted mean filter method to remove interference in the baseline signal and apply the relative difference method to eliminate interference in the experimental signal. For the second problem, this research proposes the differential entropy method to extract baseline and experimental signal features, the 3D cube method to represent experimental signal features, and the continuous capsule network method for emotion classification. This hybrid method is expected to improve the performance of emotion recognition models on EEG signals. The model was tested with three emotion datasets, such as DEAP, DREAMER, and AMIGOS, to measure the performance of the proposed model. These three datasets have different characteristics regarding the amount of data, number of channels, and recording scenarios. To produce an optimal model for recognizing emotions via EEG signals.

2. METHOD

Based on the studies, the researchers [18], [22] propose a model consisting of several methods, such as the modified weighted mean filter method for removing baseline EEG signal artefacts, the differential entropy method for feature extraction, the relative difference for baseline reduction, the 3D cube method for feature representation, and the continuous capsule network method for the emotion classification process. This experiment will be conducted on three public datasets. The flow diagram related to the research procedures that will be implemented is in Figure 1. Based on Figure 1, the research procedure can be described in the subsection as follows.

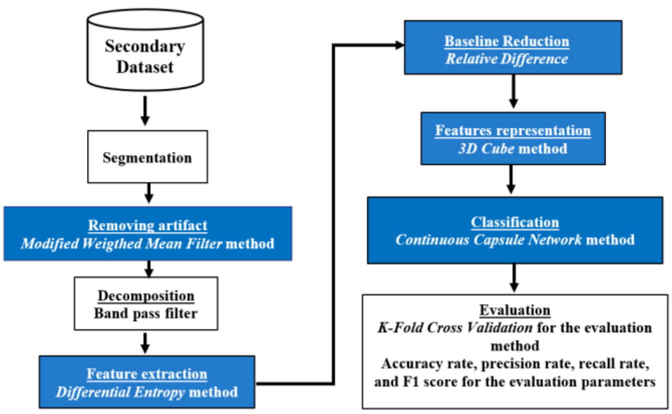

Figure 1. Research flow diagram

2.1. Collection of datasets

The datasets used in this research are DEAP, DREAMER, and AMIGOS [12]–[14]. These three datasets use different EEG recording tools, recording durations, and recording scenarios. These differences in characteristics will be instrumental in testing the reliability of the model being developed.

2.2. Segmentation

The segmentation process separates the baseline signal and the trial signal. The baseline signal is an EEG signal that represents the participant's neutral condition. The trial signal is an EEG signal that represents the participant's emotional reaction. In the DEAP dataset, the first three seconds are the baseline signals, and the fourth to sixty-third seconds are the experimental signals. In DREAMER and AMIGOS, the first five seconds are the baseline signal, and the following seconds are the trial signal [12]–[14]. Every second of EEG signal data on DEAP, DREAMER, and AMIGOS consists of a 128 Hz sampling rate. Figure 2 is an illustration of segmentation in the AMIGOS dataset. Segmentation is performed on all channels. The DEAP dataset contains 32 channels, while the DREAMER and AMIGOS datasets contain 14 [12]–[14].

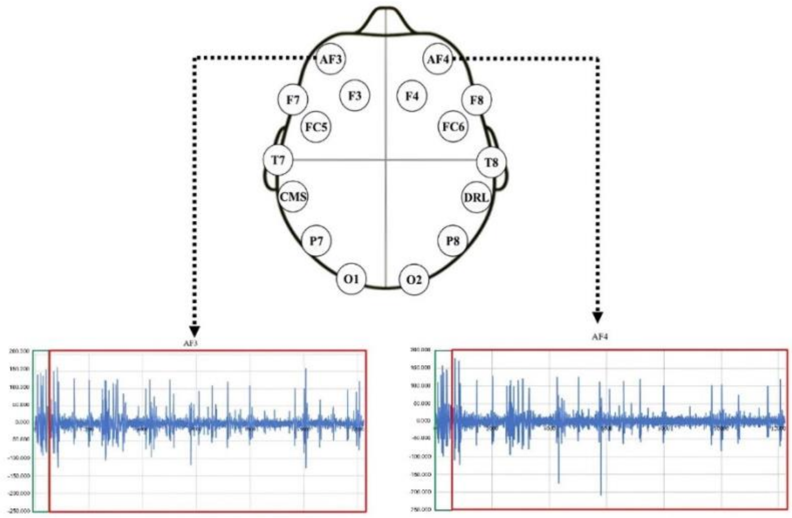

Figure 2. The EEG signal segmentation process for AF3 and AF4 channels on the AMIGOS dataset

2.3. Removing artifacts

The modified weighted mean filter method removes artefacts by smoothing the baseline signal. However, before that, the baseline signal data was normalized using the Z-score normalization method [5], [28]. This process aims to produce weight values from a baseline signal that is typically distributed (0<x<1) [25]. This weight value will be used in the modified weighted mean filter method (1).

$$z_j = \left(\frac{\sum_{i=-n}^{n} w_{j+i} x_{j+i}}{(2n+1)\sum_{i=-n}^{n} W_{j+i}}\right) \quad (1)$$

The value $j=n,\ n+1,\ n+2....,\ m+2n,\ n$ is the window length, while the value m is the amount of data. $\sum_{i=-n}^{n} w_{j+i}$ is a process to normalize the weight values to meet the requirements $\sum w_j = 1$ [29].

2.4. Decomposition

Decomposition aims to separate four frequency bands in the baseline and trial signals. This process uses the Butterworth filter method (3rd order), where the EEG signal frequency filtering process is based on the low-pass and high-pass range (bandpass filter). The decomposition process was carried out for all channels for baseline and trial signals [7].

2.5. Feature extraction

The feature extraction process is carried out to obtain relevant EEG signal features. Each frequency band performs it every second (128 Hz sampling rate) for baseline and trial signals. This research uses the differential entropy method (2).

$$h_i(X) = \frac{1}{2}\log(2\pi e \delta^2{}_i(X)) \tag{2}$$

Where e is Euler's constant, $\delta^2{}_i$ is the ith variance index, X is the theta, alpha, beta, or gamma frequency band, and h_i is the differential entropy value for the EEG signal at the ith index.

2.6. Baseline reduction

This process aims to produce feature values on trial signals that can characterize participants' emotional reactions to classify [7], [30]. This study uses the relative difference method to reduce the trial signal features with the baseline signal features [8], [31]. As an initial stage, the baseline signal's average feature value for each frequency band was calculated (3).

$$BaseMean(X) = \frac{\sum_{k=1}^{N} BaseMean_k(X)}{N} \tag{3}$$

$BaseMean_k(X)$ is the differential entropy feature value of frequency band X at the k^{th} index for the baseline signal. $BaseMean(X)$ is the average differential entropy feature value for frequency band X in the baseline signal. The value k=1, 2,.., N, where N defines the number of differential entropy feature values for each experiment, channel, and participant from the baseline signal [8]. The mean value obtained is then used for the baseline reduction process for the trial signal (4).

$$Final_j(X) = \frac{Exper_j(X)}{BaseMean(X)} \tag{4}$$

Where $Exper_j(X)$ is the feature value of frequency band j (X) represents the baseline reduction value of frequency band X at the jth index for the EEG experimental signal. Value j=1, 2, .., M, where M is the number of feature values for each experiment, channel, and participant in the trial signal.

2.7. Feature representation

After the baseline reduction process, the trial signal feature values are represented using the 3D cube method. The 3D cube is based on the international system 10-20 standard for channel placement on the scalp. The international system standard 10-20 maps spatial information between channels by representing trial signal features into a 9×9 matrix. The size of this representation is based on the maximum number of EEG channels placed on the head. This representation aims to represent spatial information between adjacent channels. Figure 3 represents differential entropy features mapped into a 9×9 matrix [7]. This representation is based on each frequency band, where red, yellow, green, and blue represent gamma, beta, alpha, and theta frequencies [7]. The DEAP and DREAMER datasets produce 2400 and 3728 3D cubes for each participant. Furthermore, each participant generated 6192 3D cubes for long experiments on the AMIGOS dataset and 1394 3D cubes for short experiments.

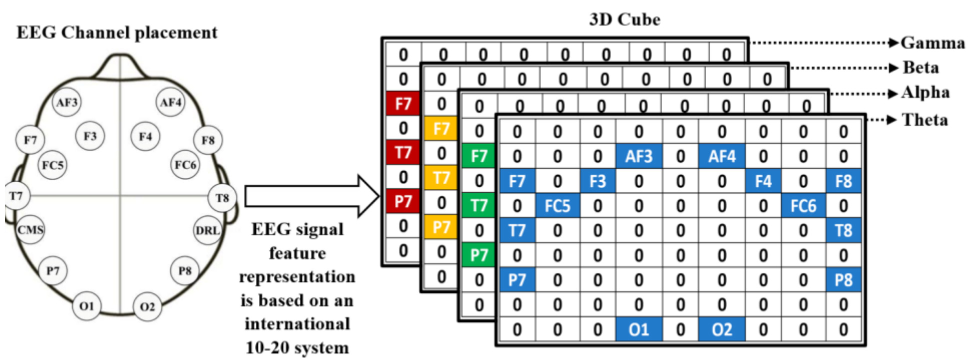

Figure 3. Feature representation for each frequency band using a 3D cube

2.7. Classification

The continuous capsule network method represents spatial information in the classification process. The architecture of the continuous capsule network method is shown in Figure 4 [22]. The continuous capsule network architecture in Figure 4 is divided into three parts: continuous convolution, primary capsule, and emotion capsule.

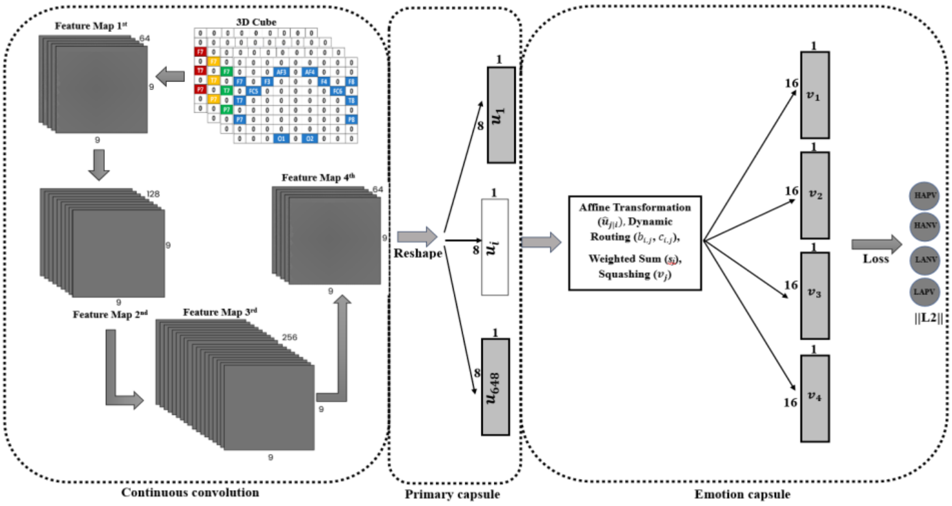

Figure 4. Continuous capsule network architecture [22]

2.7.1. Continuous convolution

At this stage, this process is carried out four times to produce four feature maps. The feature maps will be created in the 1st convolution, 2nd convolution, 3rd convolution, and 4th convolution with sizes of 9×9×64, 9×9×128, 9×9×256, and 9×9×64, respectively. The filters used are 2×2×4 with 64 for 1st convolution, 2×2×4 with 128 for 2nd convolution, 2×2×4 with 256 for 3rd convolution, and 1×1×4 with 64 for 4th convolution. Furthermore, the stride value in each convolution is 1. In addition, the activation for each convolution is rectified linear unit (ReLU). In the continuous convolution approach, the padding used is the SAME and does not use pooling, so the dimensions of the size of the feature maps produced in all convolutions remain the same; only the depth of the feature map changes. This process aims to maintain the spatial information of all EEG channels according to the position in the first data input [7]. The number of parameters generated in the 1st convolution, the 2nd convolution, the 3rd convolution, and the 4th convolution are 1,088 (64×(2×2×4)+1), 32.896 (128×(2×2×64)+1), 131.328 (256×(2×2×128)+1), and 16.448 (64×(1×1×256)+1).

2.7.2. Primary capsule

At this stage, the feature map generated from the 4th convolution is separated into eight blocks, each block with a size 9×9×8. Next, each block is reshaped into a layer of 8×1 capsules (u_i) so that one block will produce 81 capsules [22]. The primary capsule represents the number of capsules at the lower level i (u_i) with a size of 1×8, so there are 648 capsule layers generated from the eight blocks.

2.7.3. Emotion capsule

The affine transformation and dynamic routing processes are carried out at this stage. First, the emotion capsule process begins by initializing j=1 and i=1. The variable j represents the number of emotion classes to be recognized (v_j), while the variable i represents the number of capsule data (u_i). When the variables i and j are 1, the capsule data u_1 will be the first input data in the emotion capsule process and related to the 1st output class (v_1). The following process of the emotion capsule is an affine transformation. However, before this process is carried out, the initial initialization process of the weight matrix $W_{i,j}$ is carried out first. The affine transformation process will produce $\hat{u}_{j,i}$ (vector). After $\hat{u}_{j,i}$ is obtained, the routing initialization process ($r=0$) is carried out and continued with the initialization of the value $b_{i,j}$=0. The r value determines how many dynamic routing processes are carried out. The dynamic routing process is carried out by default thrice [22]. The dynamic routing process consists of calculating the coupling coefficient $c_{i,j}$, weighted sum (s_j), activating squashing (v_j), and the updated value of $b_{i,j}$. The $b_{i,j}$ value determines the coupling coefficient $c_{i,j}$ value. After the value of $c_{i,j}$ is obtained, the weighted sum process is carried out to produce the vector value s_j. The value of s_j obtained is further activated with the squashing function to make the vector value v_j. After the value of the vector v_j is received, the routing value (r) is checked. The dynamic routing process is repeated if the value of r <3. This process begins with updating the value of $b_{i,j}$, then calculating the weighted sum (s_j) and the squashing activation process (v_j). The final value of v_j in the 3rd routing is then processed to determine the value of Norm $\|v_j\|$ and L2 Regularization. Determining the value of $\|v_j\|$ (Norm) is the end of the emotion capsule process for the first capsule data (u_1) in the 1st epoch.

Furthermore, for the emotion capsule process on the following capsule data (u_i), the batch value checking process is carried out first; if the batch value is a multiple of 2, then the weight value update process $W_{i,j}$ is carried out. The weight update process uses the RMSProp method with momentum =0.09. After checking the batch value, the next step is to check the value of variable i. If the value of variable i is less than 648 (i <648), then the increment process (i++) of variable i is carried out. This increment process determines the following capsule data (u_i) to be processed in the emotion capsule. After that, the affine transformation and dynamic routing processes are executed again. This process is repeated until the 648th capsule data. When the 648th capsule data has finished executing, the next step is to check the value of variable j. If variable j is less than 4 (j <4), the value increment process of variable j is carried out (j++). This increment process determines the class v_j that will be processed next in the emotion capsule. This process will stop when 648 data capsules (u_i) have been trained for all emotion classes (j =4). Figure 5 only describes the emotion capsule flow for 648 data capsules (u_i) in 4 classes (v_j) in 1st epoch [30], [32], [33]. The four classes output, namely: high arousal positive valence (HAPV), high arousal negative valence (HANV), low arousal negative valence (LANV), and low arousal positive valence (LAPV).

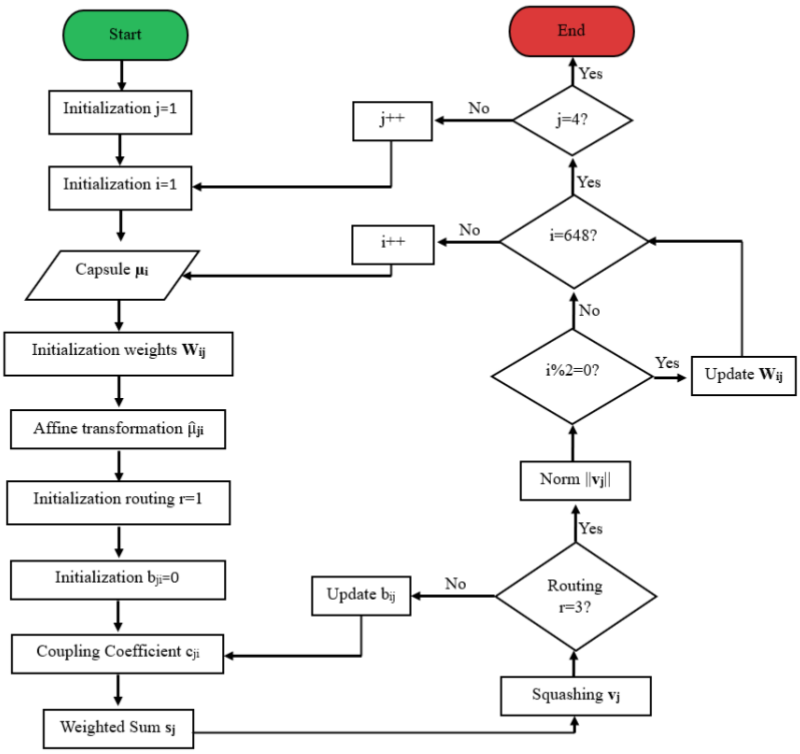

Figure 5. Flowchart emotion capsule

2.9. Evaluation method

This study uses the k-fold cross-validation evaluation method (k=10) to evaluate the proposed model. This model will measure the accuracy, precision, recall, and F1 score values for the three public datasets studied: DEAP, DREAMER, and AMIGOS. This value is used as the basis for measuring the performance of the proposed model.

3. RESULTS AND DISCUSSION

The hybrid model can produce high accuracy on three public datasets based on the evaluation method. This model is also superior to several emotion recognition models in previous studies. Figure 6 presents the proposed model's accuracy, precision, recall, and F1 score values for recognizing four emotion classes from each participant on the DEAP dataset.

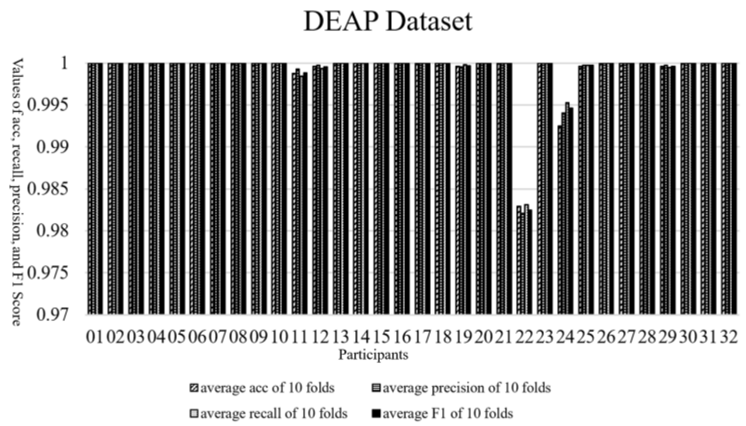

Figure 6. Accuracy, precision, recall, and F1 score graphs for recognizing four emotion classes on the DEAP dataset

Based on Figure 6, several participants with ID 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 13, 14, 15, 16, 17, 18, 20, 21, 23, 26, 27, 28, 30, 31, and 32 managed to achieve accuracy, precision, recall, and F1 score of 100%. In contrast, the lowest accuracy, precision, recall, and F1 were achieved by participant ID 22 at 98.29%, 98.20%, 98.31%, and 98.24%. Overall, the average accuracy, precision, recall, and F1 scores achieved were 99.91%, 99.92%, 99.92%, and 99.91%, respectively. In addition, standard deviation values for accuracy, precision, recall, and F1 are 0.32, 0.33, 0.31, and 0.32, respectively. So, it can be stated that the distribution of accuracy, precision, recall, and F1 score of each participant is closer to the average value compared to the model proposed in [18], [22]. Furthermore, Figure 7 presents the proposed model's accuracy, precision, recall, and F1 score for each participant on the DREAMER dataset.

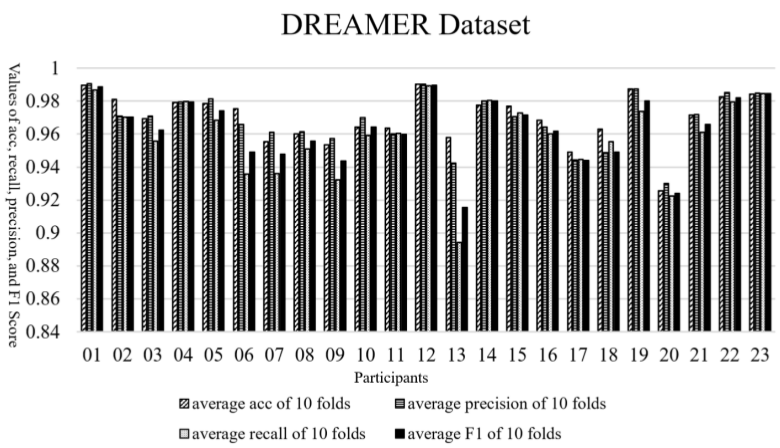

Figure 7. Accuracy, precision, recall, and F1 score graphs for recognizing four emotion classes on the DREAMER dataset

Based on Figure 7, participants with ID 12 achieved accuracy, recall, and F1 scores of 99.03%, 98.93%, and 98.97%, respectively, while participants with ID 01 achieved the highest precision at 99.07%. On the other hand, the lowest accuracy and precision were achieved by participant ID 20 at 92.55% and 92.99%, respectively. In comparison, the lowest recall and F1 scores were achieved by participant ID 13 at 89.43% and 91.55%, respectively. Although none achieved 100% accuracy, overall, the average accuracy, precision, recall, and F1 were achieved at 96.98%, 96.82%, 95.89%, and 96.28%, respectively. Based on the standard deviation values for accuracy, precision, recall, and F1 score were achieved at 1.53, 1.62, 2.32, and 1.97, respectively. So, it can be stated that the distribution of accuracy, precision, recall, and F1 score of each participant is quite close to the average value compared to the model proposed in [18], [22].

Finally, in Figure 8, 13 participants with ID 1, 2, 3, 6, 7, 9, 19, 20, 25, 26, 28, 31, and 34 achieved accuracies, precision, recall, and F1 score of 100% each. Furthermore, 19 other participants produced an average accuracy, precision, recall, and F1 score above 99%. Overall, the average data for accuracy, precision, recall, and F1 score were 99.94%, 99.93%, 99.93%, and 99.93%. In addition, the standard deviation values for accuracy, precision, recall, and F1 score are 0.10, 0.11, 0.13, and 0.12, respectively. So, it can be stated that the distribution of accuracy, precision, recall, and F1 score from each participant is closer to the average value compared to the model proposed in [18], [22]. Furthermore, to measure the success of the implementation of the proposed model, the average accuracy values in this experiment are compared with the average accuracy values from several previous studies, especially for the recognition of four emotion classes, as in Table 1.

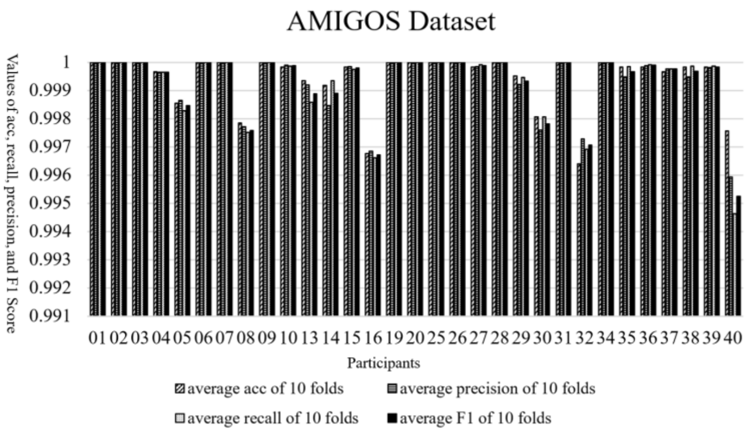

Figure 8. Accuracy, precision, recall, and F1 score graphs for recognizing four emotion classes on the AMIGOS dataset

Table 1. Comparison of the average accuracy of the four emotion classes in this study with several previous studies

Method	DEAP (%)	DREAMER (%)	AMIGOS (%)
Feature extraction: angle plot, classification: multi-class support vector machine [23]	81.67	-	-
Feature extraction: differential entropy, classification: graph regularized extreme learning machine [24]	69.67	-	-
Feature extraction: pearson correlation coefficient, classification: convolution neural network [26]	73.1		
Feature extraction: power spectral density, feature representation: multidimensional feature image, classification: LSTM + CNN [27]	75.21		
Feature extraction and classification: CNN, baseline reduction: difference. Feature representation: 3D cube [25]	93.53	-	95.95
Smoothing: MWMF, extraction feature: differential entropy, representation features: 3D cube, baseline reduction: relative difference, classification: CNN [18]	97.14	89.71	99.59
Extraction feature: differential entropy, baseline reduction: difference, representation features: 3D cube, classification: continuous capsule network [22]	91.35	94.23	96.20
Smoothing: modified weighted mean filter. Extraction feature: differential entropy, Baseline reduction: relative difference, feature representation: 3D cube, classification: continuous capsule network (purposed model)	99.91 ± 0.32	96.98 ± 1.53	99.94 ± 1.10

Although several feature extraction methods and classification methods based on machine learning and deep learning have been applied in previous studies, the accuracy of emotion recognition has not reached optimal (below 85%) [23], [24], [26], [27]. The variation of the trial signal patterns produced causes this problem. So, the classification method cannot recognize the pattern formed. To overcome this problem, Zhao *et al.* [25] used the difference method to study the baseline reduction approach. The reduced trial signal feature values can produce more representative signal patterns so that the EEG signal pattern can be recognized better. However, of several baseline reduction methods, the relative difference method is superior to the difference method [8]. Wirawan *et al.* [18] has successfully applied the relative difference method optimized with the modified weighted mean filter method. However, the CNN method used in this study cannot consider the characteristics of the EEG signal. Wirawan *et al.* [22] use of the continuous capsule network method has been proven to consider the characteristics of the EEG signal. However, Wirawan *et al.* [22] did not apply the relative difference and modified weighted mean filter methods. Based on this study, the proposed model combines the models from [18], [22] to improve emotion recognition accuracy. Based on experiments conducted on three public datasets, the proposed model in this study yields higher accuracy than the models in previous studies in recognizing four classes of emotions. Therefore, further application research can use this model to evaluate students' understanding involving emotional reactions from students during the learning process [34], [35]. In addition to emotion recognition, this model can also be used in different domains, such as detecting epilepsy [36], [37], and motor imagery [38].

However, based on the tests conducted on the DREAMER dataset, the model proposed in this study produces slightly lower accuracy values than the other two secondary datasets, such as DEAP and AMIGOS. This problem is caused by the amplitude value of the EEG signal in the DREAMER dataset being above average. Figures 9 and 10 show visualizations of the baseline signals from several participants on DEAP and AMIGOS datasets.

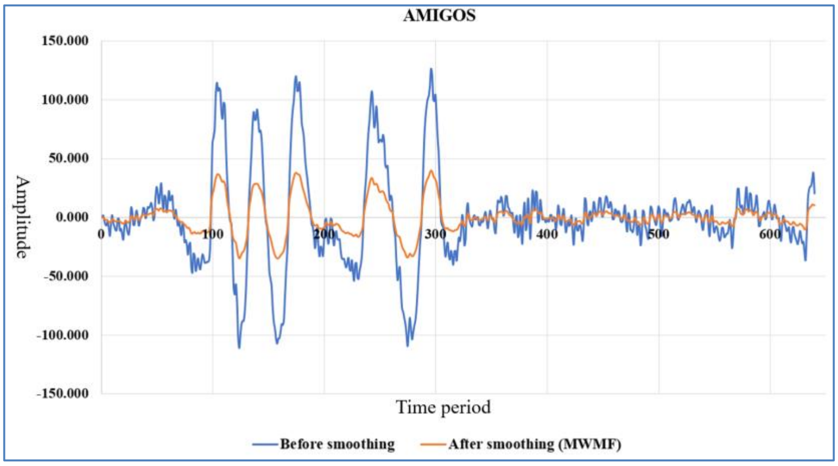

Figure 9. The smoothing process using the MWMF method on the baseline signal from the second participant in the AMIGOS dataset

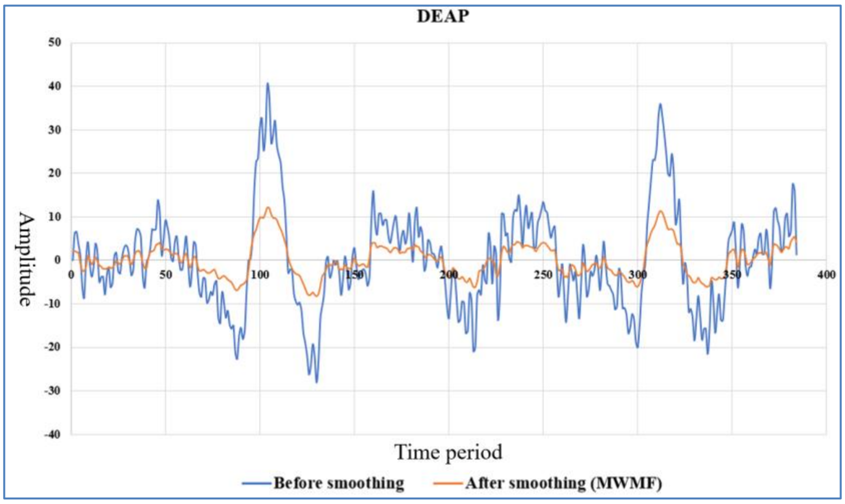

Figure 10. The smoothing process using the MWMF method on the baseline signal from the eighth participant in the DEAP dataset

After removing its artefacts using the modified weighted mean filter method, the amplitude of the baseline signal in the DEAP and AMIGOS datasets ranges ±150 μV. In contrast, the amplitude of the baseline signal in the DREAMER dataset has a range of ±1,600 µV. Figure 11 presents a visualization of the baseline signal in the tenth participant before and after smoothing on the DREAMER dataset.

However, ideally, the EEG wave in participants in normal conditions is 50-100 µV [39], [40]. Based on the experiments, it can be stated that the amplitude of the baseline signal in the DREAMER dataset still contains artefacts, even though the modified weighted mean filter method has been applied. These artefacts cause the accuracy of emotion recognition in the DREAMER dataset to be slightly lower than in the DEAP and AMIGOS datasets.

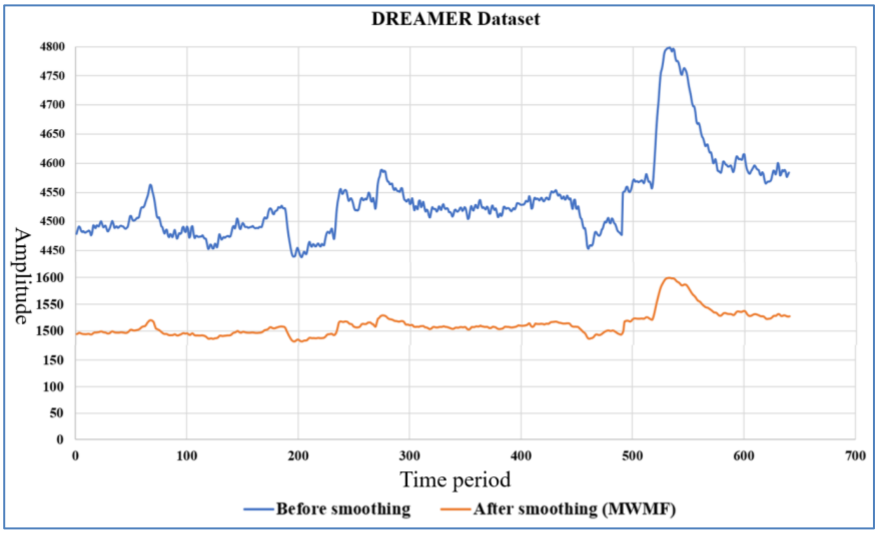

Figure 11. The smoothing process using the MWMF method on the baseline signal from the tenth participant in the DREAMER dataset

4. CONCLUSION

Two critical factors to consider in emotion recognition are differences in EEG signal patterns caused by participant characteristics and EEG signals having spatial information. These two factors significantly affect the resulting accuracy. The proposed model in this study can consider these factors. This model consists of the modified weighted mean filter method for the EEG baseline signal smoothing process, the differential entropy method for the feature extraction process, the relative difference method for baseline reduction, the 3D cube method for feature representation, and the continuous capsule network method for the classification process. Applying the modified weighted mean filter and relative difference methods can overcome differences in trial signal patterns. Meanwhile, the differential entropy, 3D cubes, and continuous capsule network methods can consider spatial information from the trial signal. Based on testing on three public datasets, these methods can improve the emotion recognition model than the model proposed in previous studies. However, based on experiments conducted on three public datasets, it can be stated that the amplitude of the baseline signal in the DREAMER dataset still contains artefacts, even though the modified weighted mean filter method has been applied. This artefact causes the accuracy of emotion recognition in the DREAMER dataset to be slightly lower than in the DEAP and AMIGOS datasets. Therefore, the future challenge is to study the appropriate weight value in the modified weighted mean filter method for the amplitude of the baseline EEG signal above standard and to study the technique of removing artefacts for experimental EEG signals.

FUNDING INFORMATION

This work was supported entirely by Funding for Fiscal Year 2024, No: 374/UN48.16/LT/2024, from the Directorate of Research, Technology, and Community Service, Ministry of Education, Culture, Research, and Technology, Indonesia.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
I Made Agus Wirawan	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Kadek Yota Ernanda Aryanto		✓	✓	✓	✓			✓		✓			✓	
I N. Sukajaya		✓	✓	✓	✓			✓		✓			✓	
Ni Nyoman Mestri Agustini					✓	✓	✓	✓		✓		✓	✓	
Dewi Arum Widhiyanti Metra Putri					✓	✓	✓	✓		✓		✓	✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are openly available in:

- The DEAP dataset openly available in [https://eecs.qmul.ac.uk/mmv/datasets/deap/download.html] at https://doi.org/10.1109/T-AFFC.2011.15, reference number [12].
- The DREAMER dataset that openly available in [https://zenodo.org/records/546113] at https://doi.org/10.1109/JBHI.2017.2688239, reference number [13].
- The AMIGOS dataset openly available in [https://www.eecs.qmul.ac.uk/mmv/datasets/amigos/download.html] at https://doi.org/10.1109/TAFFC.2018.2884461, reference number [14].

REFERENCES

[1] N. M. Krishna *et al.*, "An efficient mixture model approach in brain-machine interface systems for extracting the psychological status of mentally impaired persons using EEG signals," *IEEE Access*, vol. 7, pp. 77905–77914, 2019, doi: 10.1109/ACCESS.2019.2922047.

[2] L. Shu *et al.*, "A review of emotion recognition using physiological signals," *Sensors*, vol. 18, no. 7, 2018, doi: 10.3390/s18072074.

[3] X. Hu, J. Chen, F. Wang, and D. Zhang, "Ten challenges for EEG-based affective computing," *Brain Science Advances*, vol. 5, no. 1, pp. 1–20, 2019, doi: 10.1177/2096595819896200.

[4] T. Song, W. Zheng, P. Song, and Z. Cui, "EEG emotion recognition using dynamical graph convolutional neural networks," *IEEE Transactions on Affective Computing*, vol. 11, no. 3, pp. 532–541, 2020, doi: 10.1109/TAFFC.2018.2817622.

[5] K. Kawintiranon, Y. Buatong, and P. Vateekul, "Online music emotion prediction on multiple sessions of EEG data using SVM," in *2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, Jul. 2016, pp. 1–6, doi: 10.1109/JCSSE.2016.7748921.

[6] I. M. A. Wirawan, R. Wardoyo, and D. Lelono, "The challenges of emotion recognition methods based on electroencephalogram signals: A literature review," *International Journal of Electrical and Computer Engineering*, vol. 12, no. 2, pp. 1508–1519, 2022, doi: 10.11591/ijece.v12i2.pp1508-1519.

[7] Y. Yang, Q. Wu, Y. Fu, and X. Chen, "Continuous convolutional neural network with 3D input for EEG-based emotion recognition," in *Neural Information Processing*, Cham: Springer, 2018, pp. 433–443.

[8] I. M. A. Wirawan, R. Wardoyo, D. Lelono, S. Kusrohmaniah, and S. Asrori, "Comparison of baseline reduction methods for emotion recognition based on electroencephalogram signals," in *2021 Sixth International Conference on Informatics and Computing (ICIC)*, 2021, pp. 1–7, doi: 10.1109/ICIC54025.2021.9632948.

[9] K. Gasper, L. A. Spencer, and D. Hu, "Does neutral affect exist? How challenging three beliefs about neutral affect can advance affective research," *Frontiers in Psychology*, vol. 10, 2019, doi: 10.3389/fpsyg.2019.02476.

[10] N. Zhuang, Y. Zeng, K. Yang, C. Zhang, L. Tong, and B. Yan, "Investigating patterns for self-induced emotion recognition from EEG signals," *Sensors*, vol. 18, no. 3, 2018, doi: 10.3390/s18030841.

[11] S. Narayana, R. R. V. Prasad, and K. Warmerdam, "Mind your thoughts: BCI using single EEG electrode," *IET Cyber-Physical Systems: Theory and Applications*, vol. 4, no. 2, pp. 164–172, 2019, doi: 10.1049/iet-cps.2018.5059.

[12] S. Koelstra *et al.*, "DEAP: A database for emotion analysis; Using physiological signals," *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 18–31, 2012, doi: 10.1109/T-AFFC.2011.15.

[13] S. Katsigiannis and N. Ramzan, "DREAMER: A database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-shelf devices," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 1, pp. 98–107, 2018, doi: 10.1109/JBHI.2017.2688239.

[14] J. A. Miranda-Correa, M. K. Abadi, N. Sebe, and I. Patras, "AMIGOS: A dataset for affect, personality and mood research on individuals and groups," *IEEE Transactions on Affective Computing*, vol. 12, no. 2, pp. 479–493, 2021, doi: 10.1109/TAFFC.2018.2884461.

[15] T. Xu, Y. Zhou, Z. Wang, and Y. Peng, "Learning emotions EEG-based recognition and brain activity: a survey study on BCI for intelligent tutoring system," *Procedia Computer Science*, vol. 130, pp. 376–382, 2018, doi: 10.1016/j.procs.2018.04.056.

[16] J. Cheng *et al.*, "Emotion recognition from multi-channel EEG via deep forest," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 2, pp. 453–464, 2021, doi: 10.1109/JBHI.2020.2995767.

[17] K. Gasper, "Utilizing neutral affective states in research: theory, assessment, and recommendations," *Emotion Review*, vol. 10, no. 3, pp. 255–266, 2018, doi: 10.1177/1754073918765660.

[18] I. M. A. Wirawan, R. Wardoyo, D. Lelono, and S. Kusrohmaniah, "Modified weighted mean filter to improve the baseline reduction approach for emotion recognition," *Emerging Science Journal*, vol. 6, no. 6, pp. 1255–1273, 2022, doi: 10.28991/ESJ-2022-06-06-03.

[19] J. Li, Z. Zhang, and H. He, "Hierarchical convolutional neural networks for EEG-based emotion recognition," *Cognitive Computation*, vol. 10, no. 2, pp. 368–380, 2018, doi: 10.1007/s12559-017-9533-x.

[20] H. Jiang and J. Jia, "Research on EEG emotional recognition based on LSTM," in *Bio-inspired Computing: Theories and Applications*, 2020, pp. 409–417.

[21] D. W. Chen *et al.*, "A feature extraction method based on differential entropy and linear discriminant analysis for emotion recognition," *Sensors*, vol. 19, no. 7, 2019, doi: 10.3390/s19071631.

[22] I. M. A. Wirawan, R. Wardoyo, D. Lelono, and S. Kusrohmaniah, "Continuous capsule network method for improving electroencephalogram-based emotion recognition," *Emerging Science Journal*, vol. 7, no. 1, pp. 116–134, 2023, doi: 10.28991/ESJ-2023-07-01-09.

[23] M. Z. Soroush, K. Maghooli, S. K. Setarehdan, and A. M. Nasrabadi, "A novel EEG-based approach to classify emotions through phase space dynamics," *Signal, Image and Video Processing*, vol. 13, no. 6, pp. 1149–1156, 2019, doi: 10.1007/s11760-019-01455-y.

[24] W. L. Zheng, J. Y. Zhu, and B. L. Lu, "Identifying stable patterns over time for emotion recognition from eeg," *IEEE Transactions on Affective Computing*, vol. 10, no. 3, pp. 417–429, 2019, doi: 10.1109/TAFFC.2017.2712143.

[25] Y. Zhao, J. Yang, J. Lin, D. Yu, and X. Cao, "A 3D convolutional neural network for emotion recognition based on EEG signals," in *2020 International Joint Conference on Neural Networks (IJCNN)*, 2020, pp. 1–6, doi: 10.1109/IJCNN48605.2020.9207420.

[26] H. Mei and X. Xu, "EEG-based emotion classification using convolutional neural network," in *2017 International Conference on Security, Pattern Analysis, and Cybernetics (SPAC)*, 2017, pp. 130–135, doi: 10.1109/SPAC.2017.8304263.

[27] Y. Li, J. Huang, H. Zhou, and N. Zhong, "Human emotion recognition with electroencephalographic multidimensional features by hybrid deep neural networks," *Applied Sciences*, vol. 7, no. 10, 2017, doi: 10.3390/app7101060.

[28] A. Aytekin, "Comparative analysis of normalization techniques in the context of MCDM problems," *Decision Making: Applications in Management and Engineering*, vol. 4, no. 2, pp. 1–25, 2021, doi: 10.31181/dmame210402001a.

[29] B. Sahoo, R. N. Behera, and P. K. Pattnaik, "A comparative analysis of multi-criteria decision making techniques for ranking of attributes for e-governance in India," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 3, pp. 65–70, 2022, doi: 10.14569/IJACSA.2022.0130311.

[30] Y. Liu *et al.*, "Multi-channel EEG-based emotion recognition via a multi-level features guided capsule network," *Computers in Biology and Medicine*, vol. 123, 2020, doi: 10.1016/j.compbiomed.2020.103927.

[31] D. Lelono, H. Nuradi, M. R. Satriyo, T. W. Widodo, A. Dharmawan, and J. E. Istiyanto, "Comparison of difference, relative and fractional methods for classification of the black tea based on electronic nose," in *2019 International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM)*, 2019, pp. 1–7, doi: 10.1109/CENIM48368.2019.8973308.

[32] S. Sabour, N. Frosst, and G. E. Hinton, "Dynamic routing between capsules," *arXiv:1710.09829*, vol. 30, Oct. 2017, doi: 10.48550/arXiv.1710.09829.

[33] H. Chao, L. Dong, Y. Liu, and B. Lu, "Emotion recognition from multiband EEG signals using CapsNet," *Sensors*, vol. 19, no. 9, 2019, doi: 10.3390/s19092212.

[34] E. Henritius, E. Löfström, and M. S. Hannula, "University students' emotions in virtual learning: a review of empirical research in the 21st century," *British Journal of Educational Technology*, vol. 50, no. 1, pp. 80–100, 2019, doi: 10.1111/bjet.12699.

[35] M. Taub, R. Sawyer, J. Lester, and R. Azevedo, "The impact of contextualized emotions on self-regulated learning and scientific reasoning during learning with a game-based learning environment," *International Journal of Artificial Intelligence in Education*, vol. 30, no. 1, pp. 97–120, 2020, doi: 10.1007/s40593-019-00191-1.

[36] W. O. Tatum *et al.*, "Clinical utility of EEG in diagnosing and monitoring epilepsy in adults," *Clinical Neurophysiology*, vol. 129, no. 5, pp. 1056–1082, 2018, doi: 10.1016/j.clinph.2018.01.019.

[37] M. L. Pirgit and S. Beniczky, "EEG and semiology in the elderly: a systematic review," *Seizure: European Journal of Epilepsy*, 2024, doi: 10.1016/j.seizure.2024.09.003.

[38] I. M. A. Wirawan *et al.*, "Acquisition and processing of motor imagery and motor execution dataset (MIMED) for six movement activities," *Data in Brief*, vol. 56, 2024, doi: 10.1016/j.dib.2024.110833.

[39] M. Teplan, "Fundamentals of EEG measurement," *Measurement Science Review*, vol. 2, pp. 1–11, 2002.

[40] V. Roy and S. Shukla, "Designing efficient blind source separation methods for EEG motion artifact removal based on statistical evaluation," *Wireless Personal Communications*, vol. 108, no. 3, pp. 1311–1327, 2019, doi: 10.1007/s11277-019-06470-3.

BIOGRAPHIES OF AUTHORS

I Made Agus Wirawan is an Associate Professor in the Master of Computer Science Study Program at Universitas Pendidikan Ganesha, Bali, Indonesia. He earned bachelor's, master's, and doctorate degrees in the Department of Computer Science and Electronics, Faculty of Mathematics and Natural Sciences, Universitas Gajah Mada, Yogyakarta, Indonesia. His areas of research interest include machine learning, artificial intelligence, and deep learning. He can be contacted at email: imade.aguswirawan@undiksha.ac.id.

Kadek Yota Ernanda Aryanto is an Associate Professor in the Master of Computer Science Study Program at Universitas Pendidikan Ganesha, Bali, Indonesia. He received a bachelor's and master's degree from the Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia. Later, he received a Ph.D. in radiology from the University of Groningen, The Netherlands, in 2016. His primary research interests include medical informatics, data science, intelligent systems, and information security. He can be contacted at email: yota.ernanda@undiksha.ac.id.

I N. Sukajaya is an Associate Professor in the Mathematics Department of Universitas Pendidikan Ganesha. He achieved a Dr. degree at the Department of Electrical Engineering – Institut Teknologi Sepuluh Nopember (ITS) in 2017, and his master's degree was completed in 1999 at the Department of Informatics Engineering – Institut Teknologi Bandung. His research topics are data mining, student profiling, and severe games in education. He is also a reviewer in some qualified international journals. He can be contacted at email: nyoman.sukajaya@undiksha.ac.id.

Ni Nyoman Mestri Agustini a senior lecturer at the Department Neurology and Medical Rehabilitation, Faculty of Medicine, Universitas Pendidikan Ganesha, Bali, Indonesia. She earned a medical bachelor's degree (S.Ked.) and a medical doctor (dr.) from the Faculty of Medicine, Universitas Udayana, Bali, Indonesia. She also earned a specialist of neurology in the Department of Neurology, Faculty of Medicine, Universitas Udayana, Bali, Indonesia. Her areas of research interest include neurology and neurobehavior. She can be contacted at email: nyoman.mestri@undiksha.ac.id.

Dewi Arum Widhiyanti Metra Putri is a senior lecturer at the Department of Education, Psychology, and Guidance, Faculty of Educational Science, Universitas Pendidikan Ganesha, Bali, Indonesia. She earned a bachelor's degree (S.Psi.) from the Faculty of Psychology, Universitas Gadjah Mada, Yogyakarta, Indonesia. She obtained a master's degree (M.A.) from the Faculty of Psychology, Universitas Gadjah Mada, Yogyakarta, Indonesia. She earned a doctorate (Dr) degree from the Faculty of Psychology, Universitas Gadjah Mada, Yogyakarta, Indonesia. The areas of her research interest include developmental psychology and cultural psychology. She can be contacted at email: arum-widhiyanti@undiksha.ac.id.