



Modified Weighted Mean Filter to Improve the Baseline Reduction Approach for Emotion Recognition

I Made Agus Wirawan^{1,2}, Retantyo Wardoyo^{3*}, Danang Lelono³,
Sri Kusrohmaniah⁴

¹ Doctoral Program Department of Computer Science and Electronics, Faculty of Mathematics and Natural Science, Universitas Gadjah Mada, Yogyakarta 55281, Indonesia.

² Education of Informatics Engineering Department, Faculty of Engineering and Vocational, Universitas Pendidikan Ganesha, Singaraja 81116, Indonesia..

³ Department of Computer Science and Electronics, Faculty of Mathematics and Natural Science, Universitas Gadjah Mada, Yogyakarta 55281, Indonesia.

⁴ Department of Psychology, Faculty of Psychology, Universitas Gadjah Mada, Yogyakarta 55281, Indonesia.

Abstract

Participants' emotional reactions are strongly influenced by several factors such as personality traits, intellectual abilities, and gender. Several studies have examined the baseline reduction approach for emotion recognition using electroencephalogram signal patterns containing external and internal interferences, which prevented it from representing participants' neutral state. Therefore, this study proposes two solutions to overcome this problem. Firstly, it offers a modified weighted mean filter method to eliminate the interference of the electroencephalogram baseline signal. Secondly, it determines an appropriate baseline reduction method to characterize emotional reactions after the smoothing process. Data collected from four scenarios conducted on three datasets was used to reduce the interference and amplitude of the electroencephalogram signals. The result showed that the smoothing process can eliminate interference and lower the signal's amplitude. Based on the three baseline reduction methods, the Relative Difference method is appropriate for characterizing emotional reactions in different electroencephalogram signal patterns and has higher accuracy. Based on testing on the DEAP dataset, these proposed methods achieved accuracies of 97.14, 99.70, and 96.70% for the four categories of emotions, the two categories of arousal, and the two categories of valence, respectively. Furthermore, on the DREAMER dataset, these proposed methods achieved accuracies of 89.71, 97.63, and 96.58% for the four categories of emotions, the two categories of arousal, and the two categories of valence, respectively. Finally, on the AMIGOS dataset, these proposed methods achieved accuracies of 99.59, 98.20, and 99.96% for the four categories of emotions, the two categories of arousal, and the two categories of valence, respectively.

Keywords:

Electroencephalogram;
Emotion Recognition;
Modified Weighted Mean Filter;
Differential Entropy; 3D Cube;
Baseline Reduction; Difference;
Relative Difference; Fractional Difference;
Convolution Neural Network.

Article History:

Received: 09 June 2022
Revised: 18 August 2022
Accepted: 26 August 2022
Available online: 13 September 2022

1- Introduction

The participants' emotional reactions are strongly influenced by several factors such as personality traits, intellectual abilities, and gender [1–3]. These tend to generate different EEG signal patterns in each of them, affecting the accuracy of emotion recognition. Yang et al. (2018) studied a baseline reduction process using the Difference method to represent emotional reactions in different EEG signal patterns [4]. This is performed by subtracting the experiment's EEG signal

* **CONTACT:** rw@ugm.ac.id

DOI: <http://dx.doi.org/10.28991/ESJ-2022-06-06-03>

© 2022 by the authors. Licensee ESJ, Italy. This is an open access article under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<https://creativecommons.org/licenses/by/4.0/>).

using the average baseline. The baseline signal represents a neutral condition with a lower amplitude value than meditation, concentration, and emotional states [5–7]. Its reduction process is equivalent to the value of the experiment's EEG signal features that represent different emotional reactions. Based on the application of CNN methods to classify emotions into two categories, the baseline reduction approach proposed by Yang et al. (2018) [4] has been proven to be highly accurate. Several other studies have also analyzed different methods, for instance, Yang et al. (2018) [8] investigated baseline reduction using the difference method. Fortunately, the CNN and LSTM methods were used for feature extraction and to classify emotions into two categories, namely arousal and valence. Its use for the classification process aims to characterize the EEG signal with respect to time series.

Cheng et al. (2020) reviewed a baseline reduction approach using the Difference method. Furthermore, a multi-grained cascade forest (gcForest) method was used to classify emotions into two categories [9]. To consider the spatial information of the EEG signal in the baseline reduction approach, Liu et al. (2020), proposed the Capsule Network method. This technique was also used to classify emotions into two categories [10]. According to Zhao et al. (2020) [11], the baseline reduction approach was used to discern four emotional categories. This study further used the CNN method to classify emotions into high arousal and positive valence, high arousal and negative valence, low arousal and positive valence, and low arousal and negative valence. Wirawan et al. (2021) studied three baseline reduction methods, namely difference, relative, and fractional differences. This research stated that these three basic reduction methods failed to produce significant accuracy. There is a need for further studies to be carried out to determine an appropriate baseline reduction method for characterizing emotional reactions with respect to different EEG signal patterns [12].

Although several baseline reduction methods have been applied to deep learning, recording the EEG signal without external or internal interference is complex even though the participants are calm [5, 9]. Internal disturbances experienced during calm conditions included Electrooculograms (EOG), Electrocardiograms (ECG), Electromyograms (EMG), and emotional reactions [5, 13]. External interference is often caused by electrode power lines. This disturbance constantly increases the EEG baseline signal amplitude fluctuation in the long term [14–17]. It also leads to its inability to represent a neutral condition. On the contrary, Narayana et al. (2019) [6], and Zhuang et al. (2018) [7] stated that neutral states are usually fulfilled if the EEG signal amplitude is lower than the concentration and emotional conditions. It is essential to examine the appropriate method used to reduce the interference encountered when recording the baseline EEG signal in this study. Several approaches have been employed to eliminate interference or artefacts, including regression, Wavelet Transform, Independent Component Analysis (ICA) [18–21], and Principal Component Analysis (PCA). However, these algorithms focus only on detecting and removing some of them, such as EOG, ECG, and EMG.

Another approach this is commonly used is smoothing. This process improves data quality by converting the noise into a smoother signal. Among several others, the Mean Filter method reduces interference in EEG signals by smoothing the short-term amplitude fluctuations. This method has several advantages, such as producing a low MSE value and low computational time compared to the other methods [22]. In contrast, the Mean Filter method cannot reduce the amplitude of the EEG baseline signal in the long term [23, 24], thereby leading to the use of weighted values. This was derived from the baseline EEG signal amplitude. It was used to maintain its pattern irrespective of the fact that the amplitude was lowered. Considering that the baseline EEG signal has non-stationary characteristics, the Z-score normalization can be used to determine its weight, as well as normalize its data [25]. Modification of the Mean Filter approach by adding a weight value based on the normalization of the Z-Score is known as the Modified Weighted Mean Filter method.

The earlier description led to the formulated problem that the baseline EEG signal is disturbed both internally and externally, therefore, it is unable to represent each participant's neutral condition. This led to the proposition of two main contributions, namely (1) Offering a Modified Weighted Mean Filter method to eliminate the interference of the EEG baseline signal. (2) Determining an appropriate baseline reduction method to characterize emotional reactions in different EEG signal patterns using the feature values after smoothing with the MWMF approach. Based on the literature review from 2015 to 2021, the proposed contribution has never been evaluated in previous studies [3].

2- Related Works

The baseline reduction approach has been proven to increase the accuracy of emotional recognition based on EEG signals [4, 10]. This was detected by subtracting the feature value of the experiment EEG signal by the average baseline. This approach aims to produce signals representing different emotional reactions [12]. The emotional and neutral states are represented by experiment and baseline EEG signals, respectively. The baseline EEG signal was recorded before the participants were given the stimulus medium. They were expected to be calm and free from emotional reactions [16, 26, 27]. Neutral conditions have lower amplitude values than meditation, concentration, and emotional states [6, 7].

Yang et al. (2018) first researched a baseline reduction approach [4]. This process employed the feature values of the experiment EEG signal and the average baseline. These were obtained from the extraction procedure using the Differential Entropy method. This is a suitable approach because it tends to characterize the EEG signal with respect to the time series and its frequency [28, 29]. The value of the DE is obtained using Equation 1 [3, 30]:

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

The value of e is represented as Euler's constant, δ^2 depicts variance, and h_i is the DE for each segment of the EEG signal. The DE feature values from the baseline and experiment EEG signals were used for the reduction process. Yang et al. (2018) stated that the baseline reduction process employs the Difference method. The result of the reduction process is equivalent to the feature value of the experiment EEG signal. Furthermore, Yang et al. (2018) represented the value of the DE feature from an experiment EEG signal in a 3D Cube. This is spatially represented by the frequencies and EEG channels. The DE feature value from the experiment signal described in the 3D Cube is then used as input data in the classification method. The technique proposed by Yang et al. (2018) involves using the CNN method. To avoid the loss of feature padding information from the input data (matrices 9×9), the CNN method proposed by Yang et al. (2018) uses the SAME padding and does not employ pooling. As shown in Figure 1, the 1st 3D Cube data (input data) was convoluted four times. The 1st, 2nd, and 3rd use a 4×4 filter, and the stride value are 1. The 4th convolution uses a 1×1 filter, with a stride value of 1. The feature map of the fourth convolution was reshaped and connected to the hidden layer (fully connected). Each node was connected to the output layer (two outputs) from the hidden one. The convolution process was then activated using a Rectified Linear Unit (ReLU). Based on some tests on the DEAP dataset, average accuracy of 89.45% and 90.24% were obtained for recognizing the valence and arousal emotions, respectively. A baseline reduction approach triggered emotional recognition based on EEG signals to produce a higher accuracy [4, 8].

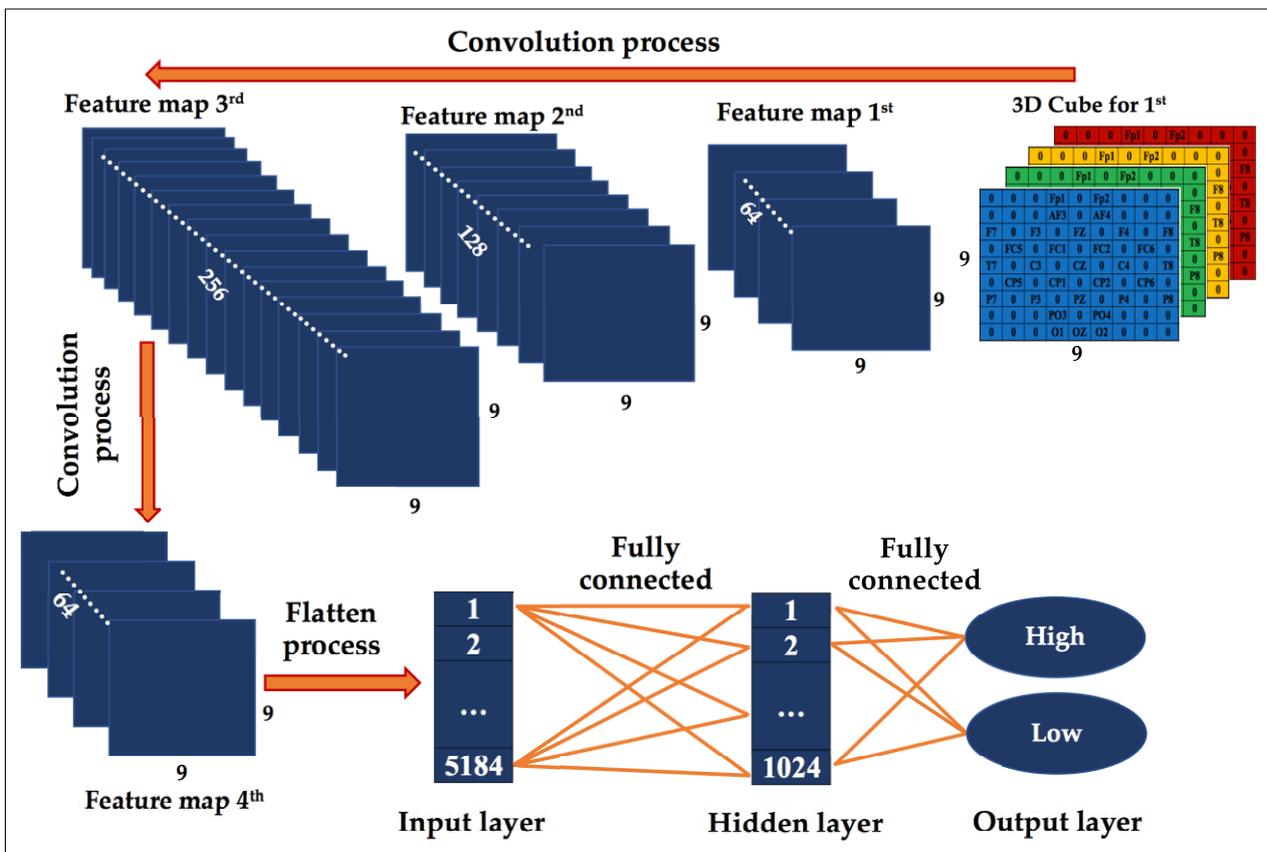


Figure 1. The architecture of the CNN method for four emotion classes [4]

Nevertheless, studies on the baseline reduction approach are still being carried out, such as that by Yang et al. (2018). A combination of CNN and LSTM methods was proposed to optimize the extraction and classification processes. These were used for the extraction process and emotion classification, respectively. Based on tests on the DEAP dataset, the proposed baseline reduction approach produced an accuracy of 90,80% and 91,03% for valence and arousal emotions, respectively [8]. In addition to using a hybrid method for the extraction and classification processes, Cheng et al. (2020) attempted to optimize the baseline reduction approach. This led to the proposed extraction and classification processes using the multi-Grained Cascade Forest (gcForest) technique. Based on tests carried out on the DEAP, the proposed baseline reduction approach accurately identified arousal and valence values of 97.53% and 97.69%, respectively. Relating to the tests on the DREAMER dataset, accuracies of 90.41% and 89.03% for arousal and valence emotions were obtained, respectively [9]. Liu et al. (2020) proposed a Capsule Network method considering the spatial information of EEG signals in the baseline reduction approach. Based on analysis performed on the DEAP dataset, the average accuracies for arousal and valence were 98.31% and 97.97%, respectively. Concerning the DREAMER dataset, accuracies of 94.59% and 95.26% were obtained for arousal and valence, respectively [10]. Its usage to identify two

categories of emotions caused Zhao et al. (2020) [11] to use the baseline reduction approach to identify four of them, namely high arousal and positive valence, high arousal and negative valence, low arousal and positive valence, and low arousal and negative valence. This investigation involved using the CNN method for the extraction and classification processes. The tests carried out on the DEAP, and AMIGOS datasets yielded mean accuracies of 93.53% and 95.86% for the four emotional classes.

In addition to the extraction and classification processes, other efforts to optimize the baseline reduction approaches were carried out by Wirawan et al. (2021). This led to the examination of three baseline reduction methods, namely the Difference, Relative, and Fractional Difference methods [4, 12, 31]. In the Difference method, the baseline reduction process involves subtracting the value of the experiment EEG signal feature by the average feature value of the baseline EEG signal feature. In the Relative Difference method, the baseline reduction is achieved by dividing the feature value of the experiment EEG signal by the average feature value of the baseline EEG signal feature. The Fractional Difference method is a combination of Difference and Relative Difference approaches. The Difference method's baseline reduction process was first applied by Yang et al. (2018) [4]:

- *The Difference Method:* The baseline reduction process is achieved by subtracting the value of the DE feature from the experiment EEG signal by the average on the baseline EEG signal. The Difference method is obtained using Equation 2:

$$Final_{v_j}(x) = Exper_{v_j}(x) - BaseMean_v(x) \quad (2)$$

- *The Relative Difference Method:* The baseline reduction process is achieved by dividing the value of the DE feature from the experiment EEG signal by the average on the baseline EEG signal [12] which is obtained using Equation 3:

$$Final_{v_j}(x) = \frac{Exper_{v_j}(x)}{BaseMean_v(x)} \quad (3)$$

- *The Fractional Difference Method:* The Fractional Difference method is obtained using Equation 4. It combines the Difference and Relative Difference methods. The value of the DE feature in the experiment EEG signal is subtracted from the average on the baseline EEG signal [12]. The result was then divided by the average value of DE features in the baseline EEG signal:

$$Final_{v_j}(x) = \frac{Exper_{v_j}(x) - BaseMean_v(x)}{BaseMean_v(x)} \quad (4)$$

wherein $Exper_{v_j}(x)$ denotes the DE feature value of the x frequency band of the j^{th} index of an experiment EEG signal. $BaseMean_v(x)$ is the baseline EEG signal's average DE feature value of the x frequency band. $Final_{v_j}(x)$ denotes the final DE feature value of the x frequency band at the j^{th} index of an experiment EEG signal. The $BaseMean_v(x)$ calculation process is given in Equation 5:

$$BaseMean_v(x) = \frac{\sum_{j=1}^N Base_{v_j}(x)}{N} \{Base_{v_j}\} \in R^C \quad (5)$$

wherein $Base_{v_j}(x)$ defines the DE feature value of the frequency band x at the j^{th} index of the EEG baseline signals. N denotes the number of DE feature values of EEG baseline signals, while C is the number of channels. The $BaseMean_v(x)$ indicates the average DE features for the frequency band x indicates the average DE features for the frequency band x on the EEG baseline signal. Based on the classification process carried out with the CNN method, tested on the DEAP dataset, the Difference, Relative, and Fractional Difference methods were used to obtain 81.25, 82.16, and 82.10% accuracy of emotional arousal, respectively. Sequentially, these were also used to obtain an accuracy of 80.61, 81.37, and 81.47%, for valence emotions. These three basic reduction methods produced significantly similar accuracies. It is essential to carry out further studies to determine an appropriate baseline approach for characterizing emotional reactions in different EEG signal patterns [12].

Although several reduction methods have been applied to deep learning, recording the baseline EEG signal without external or internal interference is complex even when participants are calm. These internal disturbances include Electrooculograms (EOG), Electrocardiograms (ECG), Electromyograms (EMG), and emotional reactions. External interference is often caused by the electrode power lines, and it constantly increases the fluctuation of the EEG baseline signal amplitude in the long term [14-17]. This causes the signal to be unable to represent a neutral condition. Three public datasets based on EEG signals are often used to investigate emotional recognition, particularly studies that adopted baseline reduction approaches, such as DEAP, DREAMER, and AMIGOS. In accordance with the random selection of these participants, the signal amplitude of the baseline and experiment EEG signals were the same, as shown in Figures 2 to 4.

The raw EEG signals obtained from the first trial with the first participant are shown in Figure 2. In the DEAP dataset, the duration of each experiment was 63 seconds. The first three seconds are the baseline EEG signals, while the remaining 4 to 63 s are that of the experiment. Additionally, the 128 Hz sampling rate represents each second.

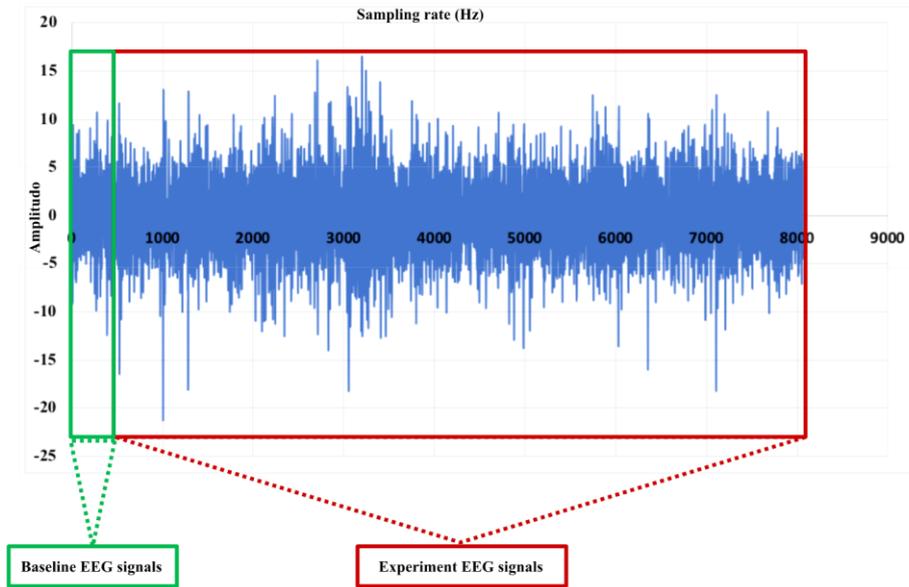


Figure 2. Raw EEG signals for Fp1 from the DEAP dataset [26]

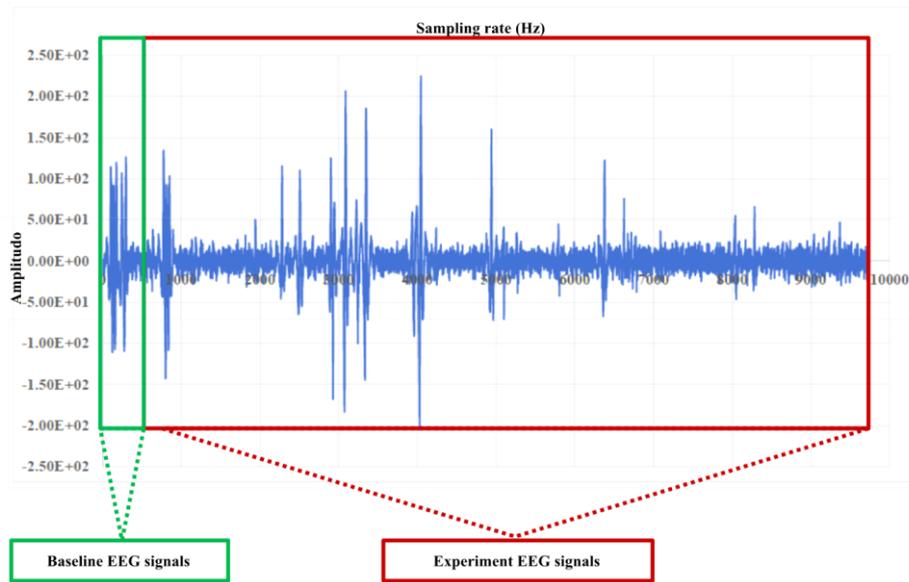


Figure 3. Raw EEG signals for AF3 from the AMIGOS data set [27]

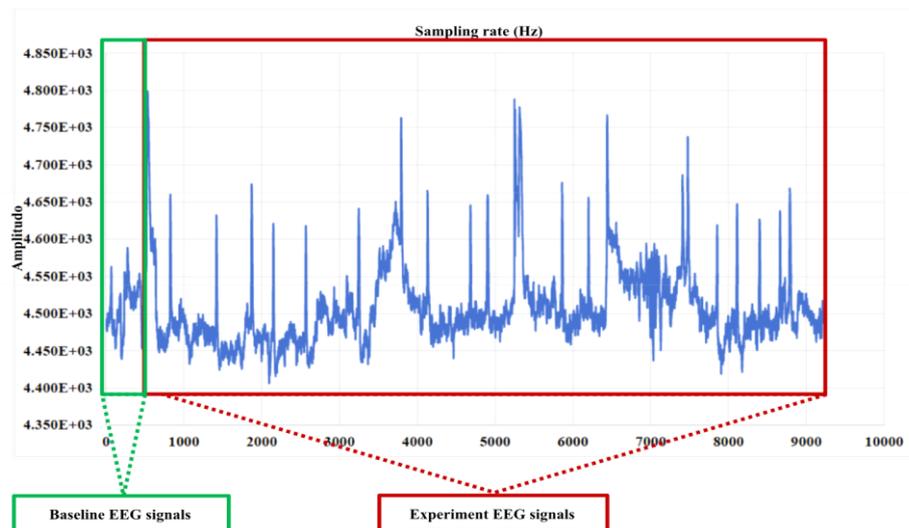


Figure 4. Raw EEG signals for AF3 from the DREAMER data set [16]

Figure 3 shows the raw EEG signals of the eighth participant during the eighth trial. The overall duration of the experiment is 75.9 seconds. In the AMIGOS dataset, the first five seconds are the baseline EEG signals, while the remaining 6 to 75.9 seconds are that of the experiment. Furthermore, the 128 Hz sampling rate represents each second. Figure 4 shows the raw EEG signals obtained from the tenth participant in the tenth trial, and the overall duration was 72 s. In the DREAMER dataset, the first five seconds were the baseline EEG signals, while the remaining 6 to 72 s were that of the experiment. Coincidentally, the 128 Hz sampling rate represents each second. The baseline EEG signal amplitude is same than the experiment EEG signal. Both external and internal interferences cause this problem during the recording process [5, 9, 32]. Internal disturbances include Electrooculograms (EOG), Electrocardiograms (ECG), Electromyograms (EMG), and emotional reactions. The power line current often causes external interference at each electrode [14–16]. This constantly increases the fluctuations of the EEG signal amplitude in the long term, thereby causing the baseline to be unable to represent neutral conditions [16, 17]. Conversely, the studies by Narayana et al. (2019) [6] and Zhuang et al. (2018) [7] and reported that the neutral state is fulfilled if the EEG signal amplitude is less than that of the concentration and emotional states.

This study designed a Modified Weighted Mean Filter (MWMF) method to reduce interference as well as represent the neutral condition in the baseline EEG signal. Interestingly, this technique is a development of the Mean Filter method in which a weight value is added, and its value is discerned using Equation 6:

$$Z_j = \frac{\sum_{i=-n}^n x_{j+i}}{2(n+1)} \quad (6)$$

where x_i denotes the input value of the EEG signal at the i^{th} index, $2(n+1)$ is the window length, and z_j indicates the current output value at the j^{th} index. Since the determination of the window lengths significantly affects the extent of delay and latency, an increase in its value triggers the extent of delay elements and latency, and vice versa [33]. Based on these conditions, the window length used in this study was 1 ($n=1$). The Mean Filter method reduces noise in the baseline EEG signal through a smoothing process [22]. Additionally, it is essential to examine the use of weights in this approach to represent the neutral condition of baseline EEG signals. The purpose was to reduce its amplitude. A neutral state tends to be achieved by reducing the baseline EEG signals. Furthermore, weight values were generated by normalizing its amplitudes. It was also used to maintain the same baseline EEG signal pattern even though the value of the amplitude was lowered. The Z-score method was employed to normalize the baseline EEG signals. Its use for the normalization process tends to overcome the non-stationary values in the EEG signal. Several critical procedures need to be considered when developing EEG signal-based emotional recognition models, such as feature extraction, representation, and classification [3, 34, 35]. The determination of these frameworks refers to the research carried out by Wu et al. (2018), and Wirawan et al. (2021) [4, 12]. Secondary datasets such as DEAP, DREAMER, and AMIGOS, were used to validate the emotional recognition.

3- The Proposed Model

Based on the two contributions proposed in this study, the stages of emotional recognition concerning EEG signals are arranged, as shown in Figure 5.

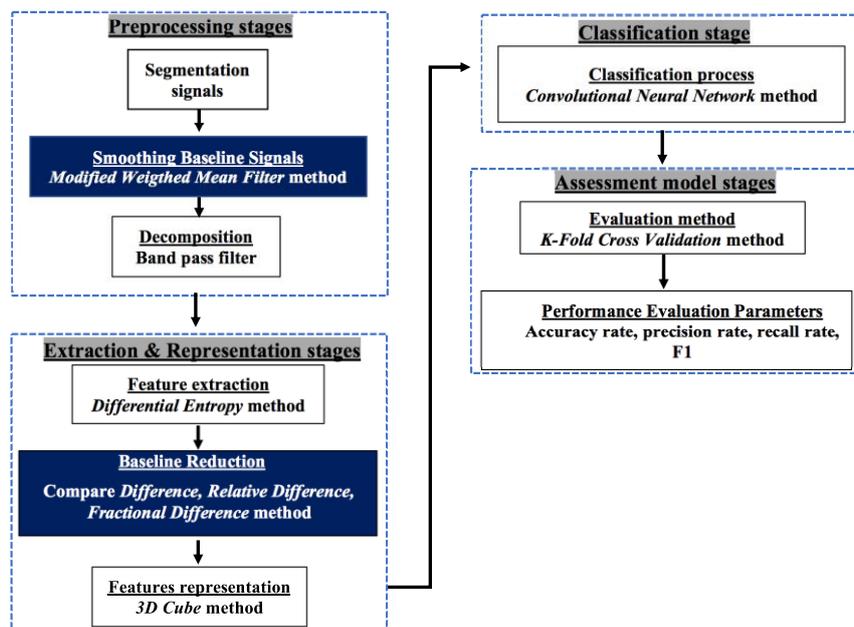


Figure 5. Emotional recognition is based on EEG signal stages

There were four emotional recognition steps, two of which are blue rectangles, and contributed to this study. The first is a smoothing approach, namely, the Modified Weighted Mean Filter technique used to eliminate interference and decrease the amplitude of the baseline EEG signals. The second contribution is to determine which of the three reduction methods is appropriate for characterizing emotional reactions from different EEG signal patterns based on those that have been smoothed. Furthermore, this model uses EEG signal data to recognize the four and two emotional classes. This study applied three public datasets, namely DEAP, DREAMER, and AMIGOS. These have different characteristics regarding the number of channels used, trial mechanisms performed individually or in groups, and the duration. The use of these three datasets was crucial for validating the proposed model.

3-1- Preprocessing Stages

3-1-1- Segmentation

This process divides EEG signals into baseline and experiment segments. Regarding the DEAP dataset, the first three seconds are the baseline signals ($x_0 \dots x_{383}$), while the fourth to the sixty-third s are that of the experiment ($x_{384} \dots x_{8063}$). Meanwhile, in DREAMER and AMIGOS, the first five seconds are the baseline signals, while the next second is that of the experiment. Figure 6 shows the segmentation process for the DEAP dataset on the Fp1 channel for the first participant [4].

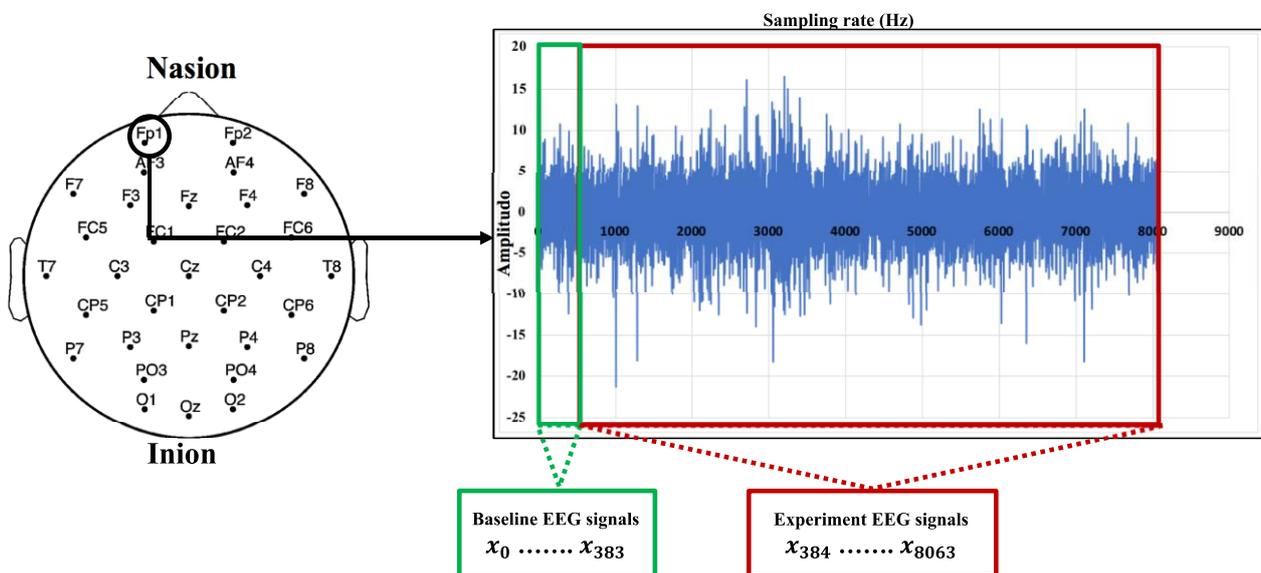


Figure 6. EEG signal segmentation processes on channel Fp1

The segmentation process was performed for each channel during the diverse trials. The DEAP dataset contains 32 channels and 40 trials for each participant. Meanwhile, that of DREAMER includes 14 channels and 18 trials. The AMIGOS dataset consists of 14 channels and 20 trials.

3-1-2- Smoothing Baseline Signal (Contribution)

A neutral state is achieved by eliminate disturbances and decrease the amplitude the baseline EEG signals, and this process is the main contribution of this study. Furthermore, a Modified Weighted Mean Filter (MWMF) method was proposed to eliminate disturbances and decrease the amplitude of the EEG baseline signals. Weight values were generated by normalizing their amplitudes. These were also used to maintain the same baseline EEG signal pattern even though its amplitude was reduced. The Z-score technique was used to normalize the baseline EEG signals and to overcome the non-stationary values. Based on Equation 6, the baseline procedure involving the use of the MWMF method consisted of three processes, namely normalization, padding, and smoothing.

- *Normalization*

The values resulting from the normalization process were converted using an absolute function (ABS). This aims to produce a positive weight value to overcome the problem of outliers when smoothing the baseline EEG signal data. Table 1 shows the values obtained using Z-Score normalization, which lasted for three seconds (384 Hz).

Table 1. Illustration of the normalization process in the baseline EEG signal

Data	Baseline signals (x_i)	Weight (W_i)
1 st Data	x_0	$W_0 = ABS\left(\frac{x_0 - \bar{x}}{\sigma}\right)$
2 nd Data	x_1	$W_1 = ABS\left(\frac{x_1 - \bar{x}}{\sigma}\right)$
3 rd Data	x_2	$W_2 = ABS\left(\frac{x_2 - \bar{x}}{\sigma}\right)$
384 th Data	x_{383}	$W_{383} = ABS\left(\frac{x_{383} - \bar{x}}{\sigma}\right)$

Based on Table 1, Equation 7 is used to calculate the weight value determined using the Z-Score normalization method [25, 36].

$$W_i = ABS\left(\frac{x_i - \bar{x}}{\sigma}\right) \tag{7}$$

\bar{x} is the average value of the EEG baseline signal data, and σ is the standard deviation. x_i is the value of the EEG baseline signal data at the i^{th} index, and W_i is the weighted value at the i^{th} index. The value of $i = 0, 1, 2, 3, \dots, M-1$, where M is the number of data points (sampling rate).

- *Padding*

The padding process is performed by adding a null value (0) both at the beginning and end of the EEG baseline signal data (x_i) and weight value data (W_i). It allows the MWMF method to perform the smoothing operation on all baseline EEG signal data (x_i) and weight value data (W_i). Given that the window length is one (1), the padding process is carried out by adding a value of zero (0) both at the beginning and end of the data. Table 2 shows the padding process for the three s baseline signal on the DEAP dataset. In addition, the 1 (s) EEG signal had 128 sampling rates. At 3 (s), the baseline signal contains 384 sampling rates.

Table 2. Illustration of the padding process in the baseline EEG signal

Data	Baseline Signals (x_i)	Weight (W_i)
1	0 ← Padding	0 ← Padding
2	x_0	W_0
3	x_1	W_1
385	x_{383}	W_{383}
386	0 ← Padding	0 ← Padding

In this process, the three-second baseline EEG signal produces x_i and W_i values of 386 data points for the DEAP dataset. The DREAMER and AMIGOS datasets had x_i and W_i values of 642 data points in five seconds.

- *Smoothing Stage*

After the weight and padding processes were performed, the baseline EEG signal was smoothed using the MWMF method. Table 3 shows the smoothing process for the baseline EEG signal data (DEAP dataset), starting from the 2nd to 385th data points.

Table 3. Illustration of the smoothing process in the baseline EEG signal

Data	Baseline signals (x_i)	Weight (W_i)	Modified Weighted Mean Filter (z_i)
1 st Data	0	0	0
2 nd Data	x_0	W_0	$z_0 = \left[0 + \left(\frac{W_0}{0 + W_0 + W_1}\right)x_0 + \left(\frac{W_1}{0 + W_0 + W_1}\right)x_1\right]/(2 + 1)$
3 rd Data	x_1	W_1	$z_1 = \left[\left(\frac{W_0}{W_0 + W_1 + W_2}\right)x_0 + \left(\frac{W_1}{W_0 + W_1 + W_2}\right)x_1 + \left(\frac{W_2}{W_0 + W_1 + W_2}\right)x_2\right]/(2 + 1)$
385 th Data	x_{383}	W_{383}	$z_{383} = \left[\left(\frac{W_{382}}{W_{382} + W_{383} + 0}\right)x_{382} + \left(\frac{W_{383}}{W_{382} + W_{383} + 0}\right)x_{383} + 0\right]/(2 + 1)$
386 th Data	0	0	0

Based on Table 3, the smoothing process of the EEG baseline signal using the MWMF method is represented by Equation 8:

$$z_j = \left(\frac{\sum_{i=-n}^n W_{j+i} x_{j+i}}{(2n+1) \sum_{i=-n}^n W_{j+i}} \right) \quad (8)$$

In the MWMF method, the values of $j = n, n+1, n+2, \dots, m+2n$, while m is the amount of data. Considering that the total weight needs to meet the requirements of $\sum W_j = 1$, determining the weight value of the Z-Score is carried out as follows $\sum_{i=-n}^n W_{j+i}$ [37].

3-1-3- Decomposition

Decomposition was carried out on the smoothed baseline and experiment EEG signals. A bandpass filter was used to convert both signals into four frequency bands. This method filters the frequency of the EEG signals based on the low and high-pass ranges. Table 4 shows the low and high-pass fields for each frequency band.

Table 4. Range of low - high pass for each frequency band

No	Frequency Band	Low Pass	High Pass
1	Theta	4 Hz	8 Hz
2	Alpha	8 Hz	14 Hz
3	Beta	14 Hz	31 Hz
4	Gamma	31 Hz	45 Hz

Decomposition was performed for all channels on the baseline and experiment EEG signals, as shown in Table 5 for the Fp1 channel.

Table 5. Illustrates the decomposition process of the baseline EEG and experiment signals for the Fp1 channel in one experiment on the DEAP dataset

Frequency band	Baseline EEG signal			Experiment EEG signal		
	z_0	z_{383}	x_{384}	x_{8063}
Theta band	$z_0\text{Theta}$	$z_{383}\text{Theta}$	$x_{384}\text{Theta}$	$x_{8063}\text{Theta}$
Alpha band	$z_0\text{Alpha}$	$z_{383}\text{Alpha}$	$x_{384}\text{Alpha}$	$x_{8063}\text{Alpha}$
Beta band	$z_0\text{Beta}$	$z_{383}\text{Beta}$	$x_{384}\text{Beta}$	$x_{8063}\text{Beta}$
Gamma band	$z_0\text{Gamma}$	$z_{383}\text{Gamma}$	$x_{384}\text{Gamma}$	$x_{8063}\text{Gamma}$

3-2- Extraction and Representation Stages

3-2-1- Extraction Features

The extraction process was carried out to obtain the relevant features of the EEG signal. This research employed the Differential Entropy (DE) method, as shown in Equation 1. The feature extraction was performed every second (128 sampling rate/data) on the smoothed baseline and the experiment EEG signal in each frequency band, as shown in Table 6 for the Fp1 channel.

Table 6. Illustrates the extraction process of the baseline EEG and experiment signals for the Fp1 channel on the DEAP dataset

Frequency band	Baseline EEG signal on 1 st s			1 st DE feature on baseline EEG signal	Experiment EEG signal on 63 th s			63 th DE feature on experiment EEG signal
	z_0	z_{127}		x_{7935}	x_{8063}	
Theta	$z_0\text{Theta}$	$z_{127}\text{Theta}$	$h_1(\text{Theta})$	$x_{7935}\text{Theta}$	$x_{8063}\text{Theta}$	$h_{63}(\text{Theta})$
Alpha	$z_0\text{Alpha}$	$z_{127}\text{Alpha}$	$h_1(\text{Alpha})$	$x_{7935}\text{Alpha}$	$x_{8063}\text{Alpha}$	$h_{63}(\text{Alpha})$
Beta	$z_0\text{Beta}$	$z_{127}\text{Beta}$	$h_1(\text{Beta})$	$x_{7935}\text{Beta}$	$x_{8063}\text{Beta}$	$h_{63}(\text{Beta})$
Gamma	$z_0\text{Gamma}$	$z_{127}\text{Gamma}$	$h_1(\text{Gamma})$	$x_{7935}\text{Gamma}$	$x_{8063}\text{Gamma}$	$h_{63}(\text{Gamma})$

The DEAP dataset produced three DE features for the baseline ($h_1(x) - h_3(x)$) and 60 for the experiment EEG signal ($h_4(x) - h_{63}(x)$) in each channel, frequency band, trial, and participant. The DREAMER and AMIGOS datasets produced five DE features for the baseline EEG signals. On the contrary, the resulting one from the experiment EEG signal corresponds to each participant's trial duration in the respective channels, frequency band, and participant.

3-2-2- Baseline Reduction (Contribution)

After the feature values of the smoothed baseline and the experiment EEG signals were obtained, a reduction process was performed. Three baseline reduction methods were examined to determine the most appropriate one for describing emotional responses from different EEG signal patterns. The DE features from both signals were used for the baseline reduction. This research reviewed three reduction processes, namely the Difference, Relative, and Fractional Difference methods. The average value of the DE features on the smoothed baseline EEG signal for each frequency band in a channel and during an experiment on one participant concerning the DEAP dataset was calculated as the initial step. Based on Equation 5, the calculation process is as follows:

$$BaseMean_v(Theta) = \frac{(h_1(Theta)+h_2(Theta)+h_3(Theta))}{3}$$

$$BaseMean_v(Alpha) = \frac{(h_1(Alpha)+h_2(Alpha)+h_3(Alpha))}{3}$$

$$BaseMean_v(Beta) = \frac{(h_1(Beta)+h_2(Beta)+h_3(Beta))}{3}$$

$$BaseMean_v(Gamma) = \frac{(h_1(Gamma)+h_2(Gamma)+h_3(Gamma))}{3}$$

Next, using Equations 2 to 4 the baseline reduction processes for the Difference, Relative, and Fractional Difference methods for each frequency band are shown in Table 7.

Table 7. Illustration of the baseline reduction using Difference, Relative Difference, and Fractional Difference methods on the DEAP dataset

DE Features	Baseline reduction method		
	Difference	Relative Difference	Fractional Difference
$Final_v_4(Theta)$	$h_4(Theta) - BaseMean_v(Theta)$	$\frac{h_4(Theta)}{BaseMean_v(Theta)}$	$\frac{(h_4(Theta) - BaseMean_v(Theta))}{BaseMean_v(Theta)}$
$Final_v_{63}(Theta)$	$h_{63}(Theta) - BaseMean_v(Theta)$	$\frac{h_{63}(Theta)}{BaseMean_v(Theta)}$	$\frac{(h_{63}(Theta) - BaseMean_v(Theta))}{BaseMean_v(Theta)}$
$Final_v_4(Alpha)$	$h_4(Alpha) - BaseMean_v(Alpha)$	$\frac{h_4(Alpha)}{BaseMean_v(Alpha)}$	$\frac{(h_4(Alpha) - BaseMean_v(Alpha))}{BaseMean_v(Alpha)}$
$Final_v_{63}(Alpha)$	$h_{63}(Alpha) - BaseMean_v(Alpha)$	$\frac{h_{63}(Alpha)}{BaseMean_v(Alpha)}$	$\frac{(h_{63}(Alpha) - BaseMean_v(Alpha))}{BaseMean_v(Alpha)}$
$Final_v_4(Beta)$	$h_4(Beta) - BaseMean_v(Beta)$	$\frac{h_4(Beta)}{BaseMean_v(Beta)}$	$\frac{(h_4(Beta) - BaseMean_v(Beta))}{BaseMean_v(Beta)}$
$Final_v_{63}(Beta)$	$h_{63}(Beta) - BaseMean_v(Beta)$	$\frac{h_{63}(Beta)}{BaseMean_v(Beta)}$	$\frac{(h_{63}(Beta) - BaseMean_v(Beta))}{BaseMean_v(Beta)}$
$Final_v_4(Gamma)$	$h_4(Gamma) - BaseMean_v(Gamma)$	$\frac{h_4(Gamma)}{BaseMean_v(Gamma)}$	$\frac{(h_4(Gamma) - BaseMean_v(Gamma))}{BaseMean_v(Gamma)}$
$Final_v_{63}(Gamma)$	$h_{63}(Gamma) - BaseMean_v(Gamma)$	$\frac{h_{63}(Gamma)}{BaseMean_v(Gamma)}$	$\frac{(h_{63}(Gamma) - BaseMean_v(Gamma))}{BaseMean_v(Gamma)}$

Based on Table. 7, 60 DE features ($Final_v_4 - Final_v_{63}$) were generated for one channel (Fp1), frequency band, trial, participant, and baseline reduction method on the DEAP dataset.

3-2-3- Representation Features

After the baseline reduction process for all frequency bands, EEG channels, and trials for one participant was completed (as shown in Table 7), the feature representation procedure was performed. These were represented based on the International System 10-20 standard. The reduced DE feature of the experiment EEG signal ($Final_v_j^i$) was mapped onto 9×9 matrices for each frequency band. This matrix describes the placement position of all channels on the head. A combination of each frequency band was represented using the 3D Cube method. Figure 7 shows the feature representation process for all channels in the DEAP dataset (32 channels).

The first 3D Cube data represents the value of the first DE feature of the experiment EEG signal ($Final_v_4(Theta)$, $Final_v_4(Alpha)$, $Final_v_4(Beta)$, and $Final_v_4(Gamma)$) for all the channels.

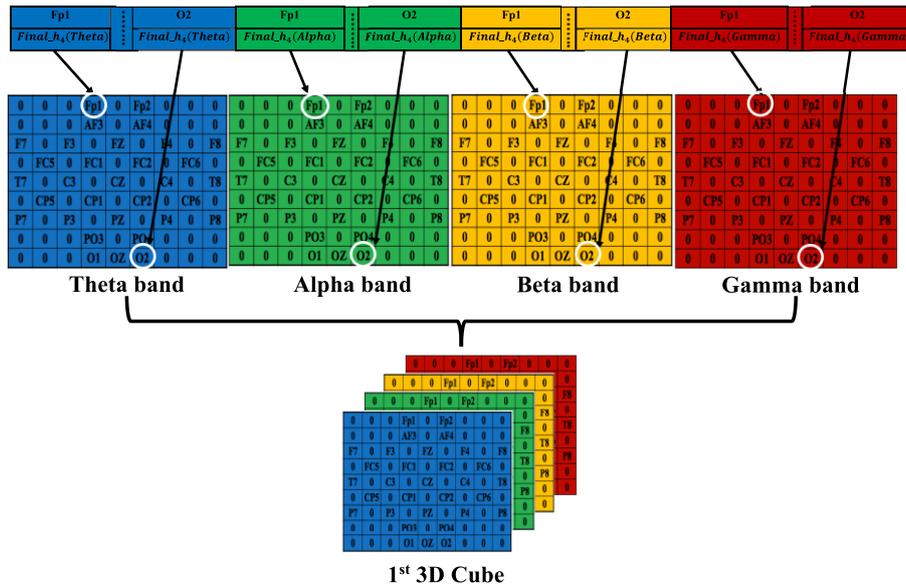


Figure 7. Illustration of the feature representation for each 32 channels in the DEAP dataset using a 3D Cube [3, 4, 38]

3-3- Classification

The DE features of the experiment EEG signal represented in the 3D Cube are used as the input data in the classification process. This study employed the Convolutional Neural Network (CNN) method for this procedure. It involves adopting a CNN architecture based on Yang et al. (2018) that designed two or four emotional categories, as shown in Figure 1. However, the two emotional categories were high and low for arousal and valence emotions. By contrast, the four emotional categories include high arousal and positive valence (HAPV), high arousal and negative valence (HANV), low arousal and negative valence (LANV), and low arousal and positive valence (LAPV). This study uses the L2 regularization and Adam Optimizer methods to calculate and update the loss value. Several parameters were determined, including the learning rate (1e-4), epoch (50), and batch size (128).

3-4- Assessment Model

There were two assessment models, namely evaluation and performance parameters. The evaluation model used in this research was a K-fold cross-validation method, with a value of 10. This was carried out on each participant using three public datasets. The evaluation process results were measured using the following parameters, namely accuracy, precision, recall, and F1 rate.

4- Results and Discussion

Based on the tests carried out on the DEAP, DREAMER, and AMIGOS datasets, the smoothing process performed using the Modified Weighted Mean Filter method tends to reduce the interference and the amplitude of the EEG baseline signal. The aim was to represent the neutral condition of the participants. Figure 8 shows the baseline EEG signal patterns before and after smoothing using the MWMF method.

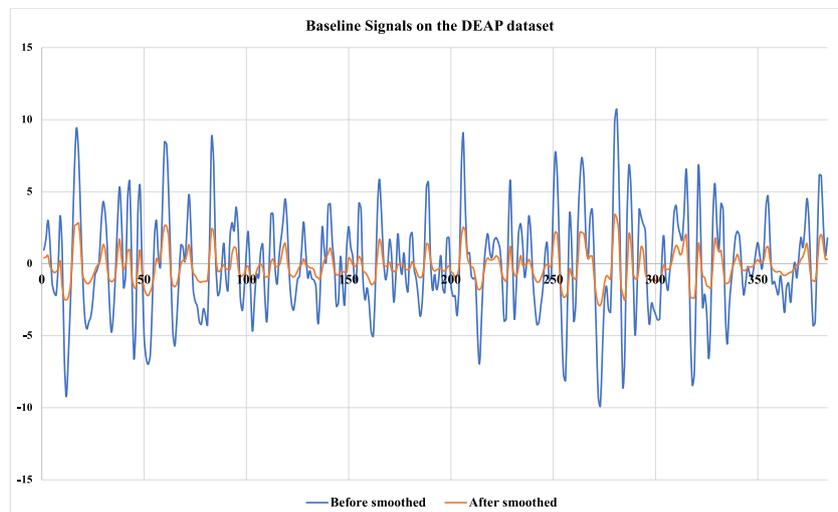


Figure 8. Baseline EEG signal patterns before and after smoothing using the MWMF method

The application of the MWMF method was used to prove that the baseline EEG signal can represent a neutral condition. Furthermore, its smoothed features were used to reduce the experiment EEG signals. This process aims to characterize emotional reactions according to the different signal patterns. The validation process was carried out using four scenarios for each dataset. The essence was to evaluate the ability of the MWMF method to represent neutral conditions on the EEG signal and assess its impact on optimizing the baseline reduction approach.

- The first test scenario involves using a three-second smoothed baseline EEG feature for the DEAP dataset (with MWMF). The DREAMER and AMIGOS datasets use a five-second smoothed baseline EEG feature (with MWMF). Furthermore, this baseline signal feature was used for its reduction. Three basic methods were employed, namely Difference, Relative, and Fraction Differences. In addition, this scenario examined the baseline reduction process utilizing an unsmoothed EEG feature (without MWMF).
- In the second scenario, the smoothing process used the first three seconds of the experiment EEG signal contained in the DEAP dataset (with MWMF). It adopted the first five seconds contained in the DREAMER and AMIGOS datasets (with MWMF). Furthermore, this experiment signal feature was used for baseline reduction. Three basic methods were used: Difference, Relative, and Fraction Differences. In addition, this scenario examined the baseline reduction process utilizing an unsmoothed experiment EEG feature (without MWMF).
- In the third scenario, the smoothing process used the last three seconds of the EEG experiment signal in the DEAP dataset (with MWMF). It also used the last five seconds in the DREAMER and AMIGOS datasets (with MWMF). This experiment signal feature was used for baseline reduction. Three basic methods, namely Difference, Relative, and Fraction Differences, were also adopted. In addition, this scenario examined the baseline reduction process utilizing an unsmoothed experiment EEG feature (without MWMF).
- In the fourth scenario, the smoothing process uses three seconds in the middle of the experiment EEG signal contained in the DEAP dataset (with MWMF). It also employed five seconds in the middle of the signal contained in the DREAMER and AMIGOS datasets (with MWMF). This feature was used for baseline reduction, as well as three basic methods, namely Difference, Relative, and Fraction Differences. In addition, this scenario examined this procedure by utilizing an unsmoothed experiment EEG feature (without MWMF).

The total number of validation processes is twenty-four for each dataset. Figure 9 shows the validation process of the proposed method on the DEAP dataset.

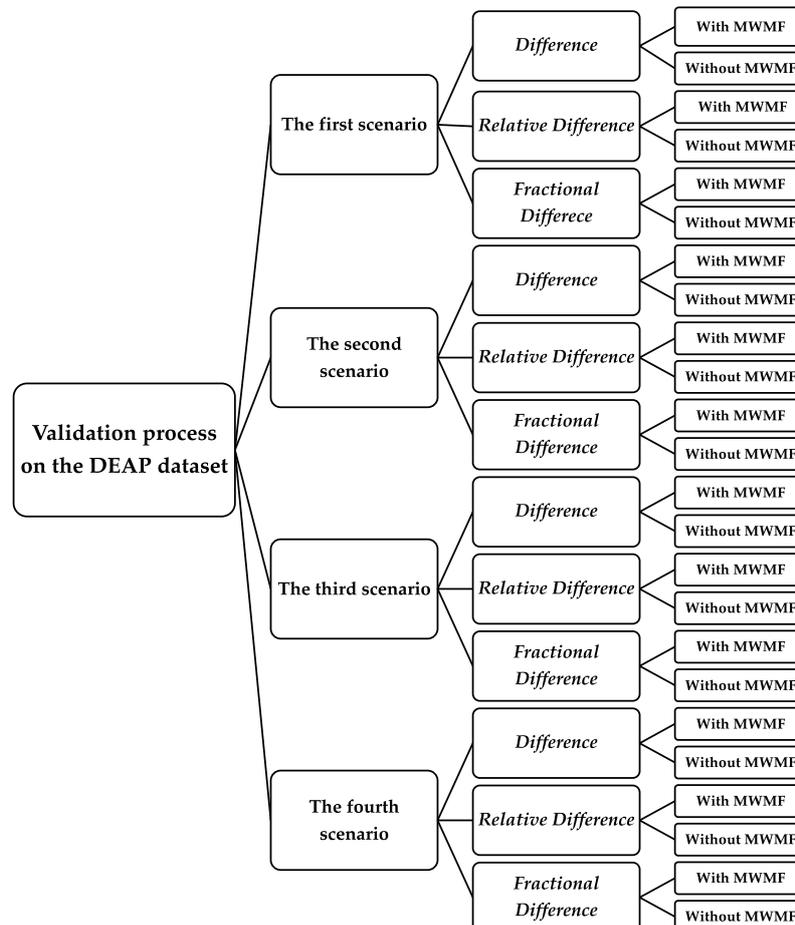


Figure 9. Twenty-four validation procedures were performed for each dataset

Each test results are the average accuracies value for the four emotional classes. Additionally, statistical analysis was carried out using the Wilcoxon test to ascertain the twenty-four validation processes on each dataset. It was also used to determine the significant increase or decrease of the four emotional classes for the three-baseline reduction approach with or without the MWMF method. This led to the development of two hypotheses.

- *Ho*: There is an insignificant difference in the accuracy of the baseline reduction methods with and without the MWMF approach. This condition is true if the 2-tailed value is greater than or equal to the degree of significance ($\alpha = 0.05$), expressed as 2-tailed $\geq \alpha$.
- *Ha*: There is a significant difference in the accuracy of the baseline reduction methods with and without the MWMF approach. This condition is true if the 2-tailed value is smaller than the degree of significance ($\alpha = 0.05$), expressed as 2-tailed $< \alpha$.

Table 8 shows the average accuracy and Wilcoxon test results for the baseline reduction method, with and without the MWMF approach for the three datasets.

Table 8. Average accuracy and statistical test for each validation process on the DEAP, DREAMER, and AMIGOS datasets

No	Scenarios	Baseline Reduction	Average Accuracy		Wilcoxon Test				Hypotheses
			Without MWMF	With MWMF	Positive Ranks	Negative Ranks	Ties	2-tailed	
1	1 st scenario on DEAP	<i>Difference</i>	77,69%	73,37%	4	28	0	0,000	<i>Ha</i> accepted
2	1 st scenario on DEAP	<i>Relative Difference</i>	79,82%	97,14%	32	0	0	0,000	<i>Ha</i> accepted
3	1 st scenario on DEAP	<i>Fractional Difference</i>	79,76%	97,13%	32	0	0	0,000	<i>Ha</i> accepted
4	2 nd scenario on DEAP	<i>Difference</i>	74,27%	70,19%	3	29	0	0,000	<i>Ha</i> accepted
5	2 nd scenario on DEAP	<i>Relative Difference</i>	76,75%	96,54%	32	0	0	0,000	<i>Ha</i> accepted
6	2 nd scenario on DEAP	<i>Fractional Difference</i>	76,70%	96,55%	32	0	0	0,000	<i>Ha</i> accepted
7	3 rd scenario on DEAP	<i>Difference</i>	73,28%	70,30%	5	27	0	0,000	<i>Ha</i> accepted
8	3 rd scenario on DEAP	<i>Relative Difference</i>	76,18%	96,85%	32	0	0	0,000	<i>Ha</i> accepted
9	3 rd scenario on DEAP	<i>Fractional Difference</i>	76,08%	96,85%	32	0	0	0,000	<i>Ha</i> accepted
10	4 th scenario on DEAP	<i>Difference</i>	71,08%	67,84%	7	25	0	0,000	<i>Ha</i> accepted
11	4 th scenario on DEAP	<i>Relative Difference</i>	74,37%	96,44%	32	0	0	0,000	<i>Ha</i> accepted
12	4 th scenario on DEAP	<i>Fractional Difference</i>	74,43%	96,42%	32	0	0	0,000	<i>Ha</i> accepted
13	1 st scenario on DREAMER	<i>Difference</i>	82,89%	80,24%	3	20	0	0,001	<i>Ha</i> accepted
14	1 st scenario on DREAMER	<i>Relative Difference</i>	76,25%	89,71%	22	1	0	0,000	<i>Ha</i> accepted
15	1 st scenario on DREAMER	<i>Fractional Difference</i>	76,24%	89,73%	22	1	0	0,000	<i>Ha</i> accepted
16	2 nd scenario on DREAMER	<i>Difference</i>	81,67%	79,65%	1	22	0	0,000	<i>Ha</i> accepted
17	2 nd scenario on DREAMER	<i>Relative Difference</i>	75,10%	88,78%	22	1	0	0,000	<i>Ha</i> accepted
18	2 nd scenario on DREAMER	<i>Fractional Difference</i>	75,08%	88,77%	22	1	0	0,000	<i>Ha</i> accepted
19	3 rd scenario on DREAMER	<i>Difference</i>	82,22%	79,43%	1	22	0	0,000	<i>Ha</i> accepted
20	3 rd scenario on DREAMER	<i>Relative Difference</i>	75,52%	88,99%	22	1	0	0,000	<i>Ha</i> accepted
21	3 rd scenario on DREAMER	<i>Fractional Difference</i>	75,42%	89,03%	22	1	0	0,000	<i>Ha</i> accepted
22	4 th scenario on DREAMER	<i>Difference</i>	82,82%	80,07%	2	21	0	0,000	<i>Ha</i> accepted
23	4 th scenario on DREAMER	<i>Relative Difference</i>	76,14%	89,26%	23	0	0	0,000	<i>Ha</i> accepted
24	4 th scenario on DREAMER	<i>Fractional Difference</i>	76,27%	89,19%	23	0	0	0,000	<i>Ha</i> accepted
25	1 st scenario on AMIGOS	<i>Difference</i>	78,60%	74,71%	0	31	0	0,000	<i>Ha</i> accepted
26	1 st scenario on AMIGOS	<i>Relative Difference</i>	75,64%	99,59%	31	0	0	0,000	<i>Ha</i> accepted
27	1 st scenario on AMIGOS	<i>Fractional Difference</i>	75,69%	99,59%	31	0	0	0,000	<i>Ha</i> accepted
28	2 nd scenario on AMIGOS	<i>Difference</i>	79,18%	75,95%	0	31	0	0,000	<i>Ha</i> accepted
29	2 nd scenario on AMIGOS	<i>Relative Difference</i>	77,44%	99,88%	31	0	0	0,000	<i>Ha</i> accepted
30	2 nd scenario on AMIGOS	<i>Fractional Difference</i>	77,52%	99,85%	31	0	0	0,000	<i>Ha</i> accepted
31	3 rd scenario on AMIGOS	<i>Difference</i>	85,56%	80,19%	0	31	0	0,000	<i>Ha</i> accepted
32	3 rd scenario on AMIGOS	<i>Relative Difference</i>	84,60%	99,48%	31	0	0	0,000	<i>Ha</i> accepted
33	3 rd scenario on AMIGOS	<i>Fractional Difference</i>	84,39%	99,50%	31	0	0	0,000	<i>Ha</i> accepted
34	4 th scenario on AMIGOS	<i>Difference</i>	83,23%	79,66%	0	31	0	0,000	<i>Ha</i> accepted
35	4 th scenario on AMIGOS	<i>Relative Difference</i>	81,98%	99,74%	31	0	0	0,000	<i>Ha</i> accepted
36	4 th scenario on AMIGOS	<i>Fractional Difference</i>	82,04%	99,75%	31	0	0	0,000	<i>Ha</i> accepted

The validation process was carried out on the DEAP, DREAMER, and AMIGOS datasets to determine the average emotional recognition accuracy of the four classes with and without the MWMF method. These were validated using the

Wilcoxon test, with and without the MWMF method. This analysis has several values supporting the significance test hypothesis, including Positive, Negative Ranks, Ties, and 2-tailed. A Positive Rank indicated the extent of data that experienced absolute accuracy when the baseline reduction method was compared with and without MWMF (with MWMF > without MWMF). A Negative Rank depicts the extent of data that experienced a significant decrease in accuracy when the baseline reduction method was compared with and without MWMF (with MWMF < without MWMF). Ties represent the extent of data that does not accurately change when the baseline reduction method is compared with and without MWMF (with MWMF = without MWMF). The 2-tailed (Asymp. sig.) is the significance value obtained from the baseline reduction method with and without MWMF. Furthermore, it was also used to test the hypotheses. H_a was accepted if the 2-tailed value was < 0.05; otherwise, H_o was accepted.

Based on the validation of the DEAP, DREAMER, and AMIGOS datasets, the Relative and Fractional Difference methods tend to produce higher accuracy when combined with the MWMF method for the four-class emotional recognition. In all scenarios, H_a was accepted because the 2-tailed value is less than 0.05 at 0.000. All participants experienced an increased accuracy in the DEAP and AMIGOS datasets. The Positive Rank values for the DEAP dataset were 32, whereas that of AMIGOS was 31. On the contrary, in the DREAMER dataset, only 1 participant out of 23 others experienced a decrease in accuracy (Negative Rank). However, the majority, namely 22 participants, experienced increased accuracy. Statistical tests concluded that there is a significant increase in accuracy when using a combination of Relative or Fraction Difference and MWMF methods.

The MWMF method tends to reduce the interference and amplitude of the baseline EEG signal. This represents the neutral condition of the participants [6, 7]. Additionally, this approach, as well as a smoothing process, can also eliminate interference and decrease the amplitude of the experiment EEG signal. Both smoothed features resulted in a significant increase in accuracy, as shown in Table 8. The baseline reduction approach can use all EEG signal features to represent emotional reactions in different patterns. The feature values from smoothed signals significantly increased the accuracy of emotional recognition when used in the Relative and Fractional Difference methods. The raw EEG signal does not represent emotional reactions in different patterns when using the Difference method. This is caused by the smoothed features yielding a low value. The adoption of the Difference method did not affect the raw EEG signal pattern, as stated in Equation 2. As shown in Figures 10-A and 10-B, there was no change in the raw EEG signal both before and after the baseline reduction employed this method. The formed pattern was still rough even though this signal has been the baseline reduction and smoothing processes. Compared to the Difference method, applying the MWMF approach to the Relative and Fractional Difference techniques yielded a raw EEG signal with a stable pattern. Figures 10-C and 10-D show the reduced pattern using both methods, producing a more stable signal than before its reduction, as indicated in Figure 10-A.

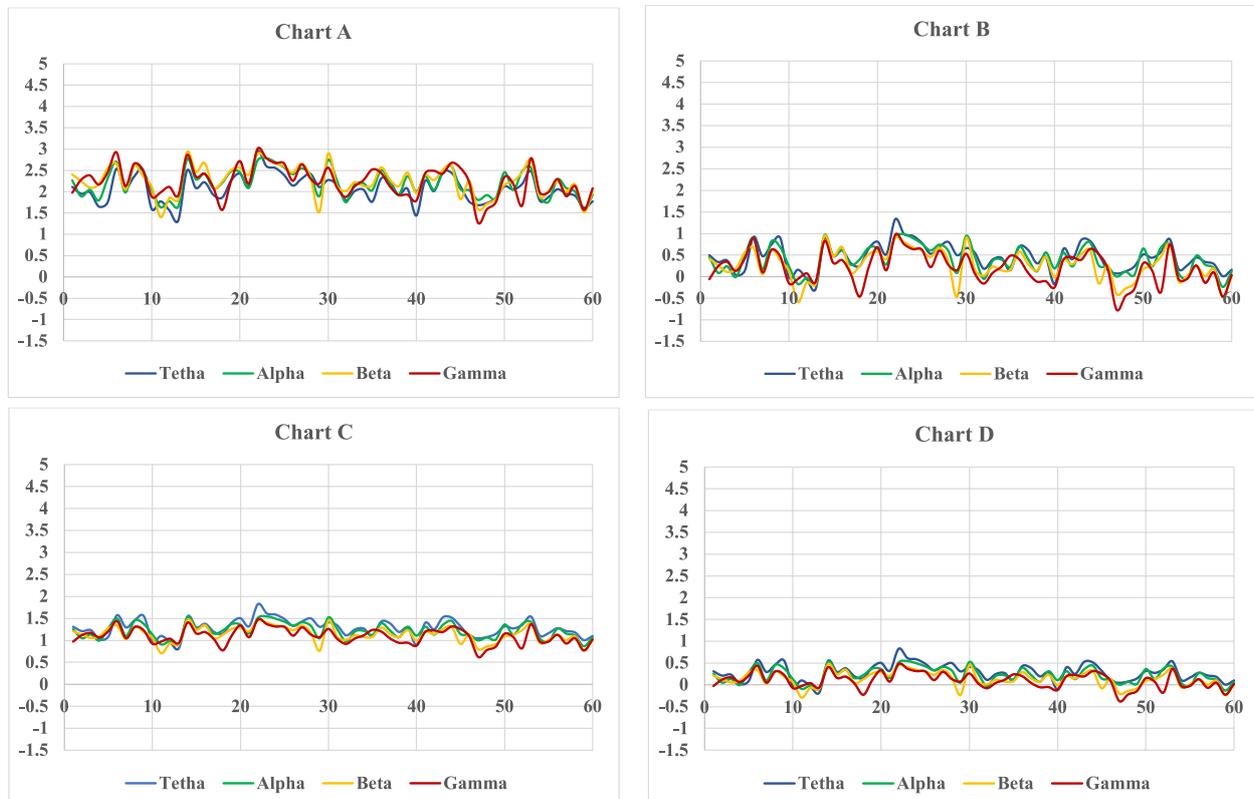


Figure 10. Visualization of experiment EEG signals before and after baseline reduction. (Graph A) Before baseline reduction. (Graph B) After baseline reduction using the Difference method. (Graph C) After baseline reduction using the Relative Difference method. (Chart D) After baseline reduction using the Fraction Difference.

Nevertheless, the Fraction Difference method combines the Difference (subtraction) and Relative Difference (division) approaches. Irrespective of the fact that when combined with the MWMF method, it increases the emotional recognition accuracy, an effective baseline reduction is performed through the division process (as shown in Equation 9) [12]. Therefore, the Relative Difference approach is more appropriate for baseline reduction. The findings of this study are compared with the average accuracy of previous studies, particularly those concerning the recognition of the four emotional categories, as shown in Table 9.

Table 9. Compares the average accuracy of the combined MWMF and Relative Difference methods (in the first scenario) with previous studies to recognize the four classes of emotions

No	Purposed methods	Research	DEAP	DREAMER	AMIGOS
1	Extraction feature using ResNets dan LFCC; Classification using kNN.	Liu et al. (2018) [39]	90.21%	-	-
2	Extraction feature using Angle Space; Classification using MSVM.	Zangeneh et al. (2019) [40]	81.67%	-	-
3	Extraction feature: compare PSD, DE, DASM, RASM, DCAU; Classification using GELM.	Zheng et al. (2016) [41]	69.67%	-	-
4	Baseline reduction using Difference; Extraction feature and Classification using CNN.	Zhao et al. (2020) [11]	93.53%	-	95.95%
5	Smoothing EEG signal using MWMF; Extraction feature using DE; Representation features using 3D Cube; Baseline reduction using Relative Difference; Classification using CNN.	Present Study	97.14%	89.71%	99.59%

Based on the comparison of accuracy in Table 9, the proposed method tends to produce the highest accuracy for the recognition of the four emotional classes than that of Liu et al. (2018), Soroush et al. (2019), and Zheng et al. (2016) [39–41]. Although Zhao et al. (2020) applied a baseline reduction approach and a CNN method, the result obtained was less accurate than the one obtained in the present study [11]. Additionally, this research also examined the recognition of two emotional classes, namely high and low for arousal and valence, respectively. The results obtained from the average accuracy were then compared with previous research, particularly those that applied the baseline reduction approach, as shown in Table 10.

Table 10. Compares the average accuracy of the combined MWMF and Relative Difference methods in the first scenario with previous studies to identify the two classes of emotions (high and low)

No	Purposed methods	Research	DEAP		DREAMER		AMIGOS	
			Arousal	Valence	Arousal	Valence	Arousal	Valence
1	Extraction feature using DE; Representation features using 3D Cube; Baseline reduction using Difference; Classification using CNN.	Yang et al. (2018) [4]	86.05%	89.45%				
2	Baseline reduction using Difference; Extraction features and Classification using CNN + LSTM.	Yang et al. (2018) [8]	91.03%	90.80%				
3	Baseline reduction using Difference; Extraction features and Classification using multi-Grained Cascade Forest (gcForest).	Cheng et al. (2020) [9]	97.53%	97.69%	90.41%	89.03%		
4	Baseline reduction using Difference; Extraction features and Classification using Multi-level Features guided Capsule Network.	Liu et al. (2020) [10]	98.31%	97.97%	95.26%	94.59%		
5	Extraction feature using DE; Representation features using 3D Cube; Baseline reduction compare Difference, Relative Difference, and Fractional Difference; Classification using CNN.	Wirawan et al. (2021) [12]	82.10%	81.47%				
6	Smoothing EEG signal using MWMF; Extraction feature using DE; Representation features using 3D Cube; Baseline reduction using Relative Difference; Classification using CNN.	Present Study	99.70%	96.70%	97.63%	96.58%	98.20%	99.96%

Based on Table 10, the proposed baseline reduction method tends to produce higher accuracy than the ones designed by Yang et al. (2018) and Wirawan et al. (2021) [8, 12]. Although this study referenced the feature extraction, representation, and classification methods from Yang et al. (2018), applying the MWMF method increases the accuracy of emotional recognition [4]. Irrespective of the fact that this method is highly accurate, based on testing the DEAP dataset, it gave a precise, slight result for emotion valence compared to the research carried out by Cheng et al. (2020) and Liu et al. (2020). Both proposed the gcForest and the Capsule Network methods for emotional classification. These techniques can monitor deep learning processes and tend to consider the spatial information in the EEG signal. Future

research needs to design a classification method to address this issue. The CNN method is unable to process spatial information from EEG signals and requires innumerable training sessions [3, 10, 42]. Considering that the three public datasets used had an imbalanced distribution, this study also measures the proposed method's accuracy, precision, recall, and F1 values, as shown in Table 11.

Table 11. Accuracy, precision, recall, and F1 values of the proposed method in Present Study

Datasets	Emotion	Accuracy	Precision	Recall	F1
DEAP	high and low for arousal	99,70%	99,70%	99,67%	99,67%
	high and low for valence	96,70%	96,72%	96,69%	96,40%
	four emotion categories	97,14%	97,40%	96,38%	96,60%
DREAMER	high and low for arousal	97,63%	96,67%	94,53%	95,40%
	high and low for valence	96,58%	96,48%	95,88%	96,13%
	four emotion categories	89,71%	89,16%	87,53%	88,16%
AMIGOS	high and low for arousal	98,20%	97,04%	96,22%	96,49%
	high and low for valence	99,96%	99,95%	99,95%	99,95%
	four emotion categories	99,59%	99,44%	99,53%	99,46%

Despite being applied to unbalanced datasets, the accuracy, precision, recall, and F1 had the same high values. Therefore, the proposed method for emotional recognition based on EEG signals is robust.

5- Conclusion

The baseline EEG signals used in the reduction approach were disturbed both internally and externally. This disturbance led to its inability to represent each participant's neutral conditions. These tend to be achieved if the amplitude of the baseline EEG signal is less than that of the experiment. This research examined a Modified Weighted Mean Filter method for smoothing baseline EEG signals. In addition, this approach can smooth the amplitude of experimental EEG signals. It is aimed at reducing noise and amplitude. Smoothed EEG signals represent neutral states, and their features are used for baseline reduction. This process aims to characterize emotional reactions according to different signal patterns. Among the three methods, Relative Difference is the most appropriate for executing this process. A combination of the Modified Weighted Mean Filter and Relative Difference methods significantly increases the accuracy of the emotional recognition based on the validation and statistical tests carried out on the DEAP, DREAMER, and AMIGOS datasets. Therefore, it was concluded that applying the Modified Weighted Mean Filter method can optimize the baseline reduction process, mainly when the Relative Difference approach is used.

Although the overall emotional recognition accuracy improved significantly, some issues led to its low accuracy in the DEAP and DREAMER datasets. Considering that the use of the CNN method was unable to represent spatial information between parts of an object and its whole in the classification process, it decreased the accuracy of emotional recognition. Therefore, it is crucial to examine the classification method to overcome these problems in future research.

6- Declarations

6-1- Author Contributions

Conceptualization, R.W. and I.M.A.W.; preparation of introduction, I.M.A.W.; preparation of related work, I.M.A.W. and R.W. and D.L. and S.K.; data collection and analysis, I.M.A.W.; preparation of the preprocess, I.M.A.W. and R.W. and D.L. and S.K.; compilation theory of smoothing signals, I.M.A.W. and R.W.; compilation theory of feature extraction, I.M.A.W. and R.W.; compilation of theory Baseline Reduction, I.M.A.W. and D.L.; compilation of the theory feature representation, I.M.A.W. and R.W.; compilation of the theory classification process, I.M.A.W. and R.W. and D.L. and S.K.; formulation of validation theory, I.M.A.W. and R.W.; compilation of theory accuracy measurement, I.M.A.W. and R.W.; development of feature extraction, representation feature, and classification method, I.M.A.W.; preparation of results and discussion, I.M.A.W. and R.W. and D.L. and S.K.; drawing conclusions, I.M.A.W. and R.W. and D.L. and S.K. The published manuscript has been read and approved by all authors.

6-2- Data Availability Statement

This study used three public datasets, including the DEAP, DREAMER, and AMIGOS datasets. Where the DEAP dataset can be freely accessed at: <https://anaxagoras.eecs.qmul.ac.uk/request.php?dataset=DEAP>, the DREAMER dataset is freely accessible at: <https://zenodo.org/record/546113#.YoB2n5MzaRs>, and the dataset AMIGOS is freely accessible at: <https://anaxagoras.eecs.qmul.ac.uk/request.php?dataset=AMIGOS>.

6-3- Funding

The publication of this article was funded by the Universitas Gadjah Mada, through the Rekognisi Tugas Akhir fund no 482/UN1.P.III/KPT/HUKOR/2022.

6-4- Institutional Review Board Statement

Not applicable.

6-5- Informed Consent Statement

Not applicable.

6-6- Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancies have been completely observed by the authors.

7- References

- [1] Zhang, J., Yin, Z., Chen, P., & Nichele, S. (2020). Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review. *Information Fusion*, 59, 103–126. doi:10.1016/j.inffus.2020.01.011.
- [2] Tyng, C. M., Amin, H. U., Saad, M. N. M., & Malik, A. S. (2017). The influences of emotion on learning and memory. *Frontiers in Psychology*, 8. doi:10.3389/fpsyg.2017.01454.
- [3] Made Agus Wirawan, I., Wardoyo, R., & Lelono, D. (2022). The challenges of emotion recognition methods based on electroencephalogram signals: A literature review. *International Journal of Electrical and Computer Engineering*, 12(2), 1508–1519. doi:10.11591/ijece.v12i2.pp1508-1519.
- [4] Yang, Y., Wu, Q., Fu, Y., Chen, X. (2018). Continuous Convolutional Neural Network with 3D Input for EEG-Based Emotion Recognition. *Neural Information Processing. ICONIP 2018, Lecture Notes in Computer Science*, 11307. Springer, Cham, Switzerland. doi:10.1007/978-3-030-04239-4_39.
- [5] Gasper, K., Spencer, L. A., & Hu, D. (2019). Does Neutral Affect Exist? How Challenging Three Beliefs About Neutral Affect Can Advance Affective Research. *Frontiers in Psychology*, 10. doi:10.3389/fpsyg.2019.02476.
- [6] Narayana, S., Prasad, R. R. V., & Warmerdam, K. (2019). Mind your thoughts: BCI using single EEG electrode. *IET Cyber-Physical Systems: Theory and Applications*, 4(2), 164–172. doi:10.1049/iet-cps.2018.5059.
- [7] Zhuang, N., Zeng, Y., Yang, K., Zhang, C., Tong, L., & Yan, B. (2018). Investigating patterns for self-induced emotion recognition from EEG signals. *Sensors (Switzerland)*, 18(3), 1–22. doi:10.3390/s18030841.
- [8] Yang, Y., Wu, Q., Qiu, M., Wang, Y., & Chen, X. (2018). Emotion Recognition from Multi-Channel EEG through Parallel Convolutional Recurrent Neural Network. *2018 International Joint Conference on Neural Networks (IJCNN)*, Rio de Janeiro, Brazil. doi:10.1109/ijcnn.2018.8489331.
- [9] Cheng, J., Chen, M., Li, C., Liu, Y., Song, R., Liu, A., & Chen, X. (2021). Emotion Recognition from Multi-Channel EEG via Deep Forest. *IEEE Journal of Biomedical and Health Informatics*, 25(2), 453–464. doi:10.1109/JBHI.2020.2995767.
- [10] Liu, Y., Ding, Y., Li, C., Cheng, J., Song, R., Wan, F., & Chen, X. (2020). Multi-channel EEG-based emotion recognition via a multi-level features guided capsule network. *Computers in Biology and Medicine*, 123, 103927. doi:10.1016/j.combiomed.2020.103927.
- [11] Zhao, Y., Yang, J., Lin, J., Yu, D., & Cao, X. (2020). A 3D Convolutional Neural Network for Emotion Recognition based on EEG Signals. *2020 International Joint Conference on Neural Networks (IJCNN)*. doi:10.1109/ijcnn48605.2020.9207420.
- [12] Agus Wirawan, I. M., Wardoyo, R., Lelono, D., Kusrohmaniah, S., & Asrori, S. (2021). Comparison of Baseline Reduction Methods for Emotion Recognition Based on Electroencephalogram Signals. *2021 6th International Conference on Informatics and Computing (ICIC 2021)*. doi:10.1109/ICIC54025.2021.9632948.
- [13] Jiang, X., Bian, G. Bin, & Tian, Z. (2019). Removal of artifacts from EEG signals: A review. *Sensors (Switzerland)*, 19(5), 1–18. doi:10.3390/s19050987.
- [14] Usakli, A. B. (2010). Improvement of EEG signal acquisition: An electrical aspect for state of the Art of front end. *Computational Intelligence and Neuroscience*, 2010. doi:10.1155/2010/630649.
- [15] Kawala-Sterniuk, A., Podpora, M., Pelc, M., Blaszczyzyn, M., Gorzelanczyk, E. J., Martinek, R., & Ozana, S. (2020). Comparison of Smoothing Filters in Analysis of EEG Data for the Medical Diagnostics Purposes. *Sensors*, 20(3), 807. doi:10.3390/s20030807.
- [16] Katsigiannis, S., & Ramzan, N. (2018). 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*, 22(1), 98–107. doi:10.1109/JBHI.2017.2688239.

- [17] Athavipach, C., Pan-Ngum, S., & Israsena, P. (2019). A wearable in-ear EEG device for emotion monitoring. *Sensors (Switzerland)*, 19(18), 1–16. doi:10.3390/s19184014.
- [18] Jin, Y.-M., Luo, Y.-D., Zheng, W.-L., & Lu, B.-L. (2017). EEG-based emotion recognition using domain adaptation network. 2017 International Conference on Orange Technologies (ICOT). doi:10.1109/icot.2017.8336126.
- [19] Al-Shargie, F., Tariq, U., Alex, M., Mir, H., & Al-Nashash, H. (2019). Emotion Recognition Based on Fusion of Local Cortical Activations and Dynamic Functional Networks Connectivity: An EEG Study. *IEEE Access*, 7, 143550–143562. doi:10.1109/access.2019.2944008.
- [20] Song, T., Zheng, W., Lu, C., Zong, Y., Zhang, X., & Cui, Z. (2019). MPED: A Multi-Modal Physiological Emotion Database for Discrete Emotion Recognition. *IEEE Access*, 7, 12177–12191. doi:10.1109/access.2019.2891579.
- [21] Thammasan, N., Moriyama, K., Fukui, K., & Numao, M. (2016). Continuous Music-Emotion Recognition Based on Electroencephalogram. *IEICE Transactions on Information and Systems*, E99. D(4), 1234–1241. doi:10.1587/transinf.2015edp7251.
- [22] Kowalski, P. & Smyk, R., 2018. Review and comparison of smoothing algorithms for one-dimensional data noise reduction. 2018 International Interdisciplinary PhD Workshop (IIPHDW). doi:10.1109/iiphdw.2018.8388373.
- [23] Zhong, X., Yin, Z., & Zhang, J. (2020). Cross-Subject emotion recognition from EEG using Convolutional Neural Networks. 2020 39th Chinese Control Conference (CCC). doi:10.23919/ccc50068.2020.9189559.
- [24] Veeramallu, G. K. P., Anupalli, Y., Jilumudi, S. kumar, & Bhattacharyya, A. (2019). EEG based automatic emotion recognition using EMD and Random forest classifier. 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT). doi:10.1109/iccct45670.2019.8944903.
- [25] Kawintiranon, K., Buatong, Y., & Vateekul, P. (2016). Online music emotion prediction on multiple sessions of EEG data using SVM. 2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE). doi:10.1109/JCSSE.2016.7748921.
- [26] Koelstra, S., Mühl, C., Soleymani, M., Lee, J. S., Yazdani, A., Ebrahimi, T., Pun, T., Nijholt, A., & Patras, I. (2012). DEAP: A database for emotion analysis; using physiological signals. *IEEE Transactions on Affective Computing*, 3(1), 18–31. doi:10.1109/T-AFFC.2011.15.
- [27] Miranda-Correa, J. A., Abadi, M. K., Sebe, N., & Patras, I. (2021). AMIGOS: A Dataset for Affect, Personality and Mood Research on Individuals and Groups. *IEEE Transactions on Affective Computing*, 12(2), 479–493. doi:10.1109/TAFFC.2018.2884461.
- [28] Pane, E. S., Wibawa, A. D., & Pumomo, M. H. (2018). Channel Selection of EEG Emotion Recognition using Stepwise Discriminant Analysis. 2018 International Conference on Computer Engineering, Network and Intelligent Multimedia (CENIM). doi:10.1109/CENIM.2018.8711196.
- [29] Chen, D. W., Miao, R., Yang, W. Q., Liang, Y., Chen, H. H., Huang, L., ... & Han, N. (2019). A feature extraction method based on differential entropy and linear discriminant analysis for emotion recognition. *Sensors*, 19(7). doi:10.3390/s19071631.
- [30] Jiang, H., Jia, J. (2020). Research on EEG Emotional Recognition Based on LSTM. *Bio-inspired Computing: Theories and Applications*. BIC-TA 2019, Communications in Computer and Information Science, 1160. Springer, Singapore. doi:10.1007/978-981-15-3415-7_34.
- [31] Lelono, D., Nuradi, H., Satriyo, M. R., Widodo, T. W., Dharmawan, A., & Istiyanto, J. E. (2019). Comparison of Difference, Relative and Fractional Methods for Classification of the Black Tea Based on Electronic Nose. 2019 International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM). doi:10.1109/cenim48368.2019.8973308.
- [32] Xu, T., Zhou, Y., Wang, Z., & Peng, Y. (2018). Learning Emotions EEG-based Recognition and Brain Activity: A Survey Study on BCI for Intelligent Tutoring System. *Procedia Computer Science*, 130, 376–382. doi:10.1016/j.procs.2018.04.056.
- [33] Al-Odienat, A. I., & Al-Mbaideen, A. A. (2015). Optimal length determination of the moving average filter for power system applications. *International Journal of Innovative Computing, Information and Control*, 11(2), 691–705.
- [34] Alarcão, S. M., & Fonseca, M. J. (2019). Emotions recognition using EEG signals: A survey. *IEEE Transactions on Affective Computing*, 10(3), 374–393. doi:10.1109/TAFFC.2017.2714671.
- [35] Bhandari, N. K., & Jain, M. (2020). Emotion recognition and classification using EEG: A review. *International Journal of Scientific and Technology Research*, 9(2), 1827–1836.
- [36] Aytekin, A. (2021). Comparative analysis of normalization techniques in the context of MCDM problems. *Decision Making: Applications in Management and Engineering*, 4(2), 1–25. doi:10.31181/dmame210402001a.
- [37] Divayana, D. G. H., Ariawan, I. P. W., Ardana, I. M., & Wayan Arta Suyasa, P. (2021). Utilization of alkin-wp-based digital library evaluation software as evaluation tool of digital library effectiveness. *Emerging Science Journal*, 5(5), 731–746. doi:10.28991/esj-2021-01308.

- [38] Wardoyo, R., Wirawan, I. M. A., & Pradipta, I. G. A. (2022). Oversampling Approach Using Radius-SMOTE for Imbalance Electroencephalography Datasets. *Emerging Science Journal*, 6(2), 382–398. doi:10.28991/ESJ-2022-06-02-013.
- [39] Liu, N., Fang, Y., Li, L., Hou, L., Yang, F., & Guo, Y. (2018). Multiple feature fusion for automatic emotion recognition using EEG signals. 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). doi:10.1109/ICASSP.2018.8462518.
- [40] Zangeneh Soroush, M., Maghooli, K., Setarehdan, S. K., & Nasrabadi, A. M. (2019). A novel EEG-based approach to classify emotions through phase space dynamics. *Signal, Image and Video Processing*, 13(6), 1149-1156. doi:10.1007/s11760-019-01455-y.
- [41] Zheng, W. L., Zhu, J. Y., & Lu, B. L. (2019). Identifying stable patterns over time for emotion recognition from EEG. *IEEE Transactions on Affective Computing*, 10(3), 417–429. doi:10.1109/TAFFC.2017.2712143.
- [42] Sabour, S., Frosst, N., & Hinton, G. E. (2017). Dynamic routing between capsules. *Advances in neural information processing systems 30 (NIPS 2017)*, 4-9 December, 2017, Long Beach, United States.