IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 29 March 2024, accepted 27 May 2024, date of publication 4 June 2024, date of current version 26 June 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3409384

RESEARCH ARTICLE

Enhancing EEG-Based Emotion Recognition Using Asymmetric Windowing Recurrence Plots

DWI WAHYU PRABOWO[1,2], (Graduate Student Member, IEEE), NOOR AKHMAD SETIAWAN[1], (Member, IEEE), JOHAN DEBAYLE[1,3], (Senior Member, IEEE), AND HANUNG ADI NUGROHO[1], (Senior Member, IEEE)

[1]Department of Electrical and Information Engineering, Faculty of Engineering, Universitas Gadjah Mada, Yogyakarta 55281, Indonesia
[2]Department of Information System, Faculty of Computer Science, Universitas Darwan Ali, Sampit 74323, Indonesia
[3]MINES Saint-Étienne, CNRS, UMR 5307 LGF, Centre SPIN, Saint-Étienne, France

Corresponding author: Hanung Adi Nugroho (adinugroho@ugm.ac.id)

This work was supported by the High-Impact Publication Assistance Program at Universitas Gadjah Mada and the Article Processing Charges (APC) IEEE-UGM Token.

ABSTRACT Time-series classification (TSC) has been widely utilized across various domains, including brain-computer interfaces (BCI) for emotion recognition through electroencephalogram (EEG) signals. However, traditional methods often struggle to capture the complex emotional patterns present in EEG data. Recent advancements in encoding techniques have provided promising avenues for improving emotion recognition. This study introduces asymmetric windowing recurrence plots (AWRP) as a novel encoding technique to efficiently encapsulate the dynamic characteristics of EEG signals into texture-rich image representations. This study systematically compares the impact of conventional thresholded and unthresholded recurrence plots (RP) versus the proposed AWRP in emotion recognition tasks. Empirical validations conducted across benchmark datasets, such as DEAP and SEED, demonstrate that the AWRP method achieves classification accuracies of 99.84% and 99.69%, respectively, outperforming existing state-of-the-art methodologies. This study emphasizes the significance of input formulation, highlighting that richer input textures, as provided by AWRP, significantly enhance emotion recognition performance while ensuring computational memory usage efficiency. These findings have significant implications in the domain of EEG-based emotion recognition and offer a novel perspective that can guide future research.

INDEX TERMS Time-series classification (TSC), electroencephalogram (EEG), asymmetric windowing recurrence plots (AWRP), emotion recognition, brain–computer interfaces (BCI).

I. INTRODUCTION

Time-series classification (TSC) has become increasingly critical for analyzing sequential data across various domains, including financial markets, healthcare, manufacturing, and environmental studies. Its ability to uncover complex patterns in temporal datasets has contributed significantly to advancements in these fields. In the domain of brain-computer interfaces (BCI) and emotion recognition, the application of TSC has been particularly impactful. By analyzing electroencephalogram (EEG) signals, a non-invasive method for recording brain activity is essential for accurately interpreting human emotions. Integrating TSC with EEG technology facilitates a more objective understanding of emotional states, which is crucial for enhancing BCI applications and developing practical solutions in this area [1], [2], [3], [4].

Research on emotion recognition has explored various modalities, including facial expressions [5], [6], gestures [7], [8], voice styles [9], [10], [11], and physiological signals, such as electrocardiography (ECG) [12], [13] and galvanic skin response (GSR) [14], [15]. Among these, EEG signals have unique advantages. EEG signals, which measure electrical activity in the brain, offer direct insights into the neural underpinnings of emotional states. This specificity and directness make EEG signals a particularly promising modality for emotion recognition, offering the potential for more accurate detection of emotional states than other methods. Furthermore, EEG signals are particularly

The associate editor coordinating the review of this manuscript and approving it for publication was Laura Celentano.

noteworthy because of their intrinsic linkage to the activity of the central nervous system, which makes them resilient to falsification and manipulation [16]. This underscores the potential of EEG signals as a reliable modality for emotion recognition.

Despite recent advancements in deep learning for EEG-based emotion recognition, these methods often overlook the complex nonlinear dynamics inherent in EEG data. Most existing approaches rely on linear input representation techniques, which may not fully capture the intricacies of the nonlinear characteristics of EEG signals [17], [18], [19], [20], [21]. Thus, effective input formulation strategies for EEG-based emotion recognition are urgently required. As Prabowo et al. [4] highlighted, proper input formulation is essential for enhancing the performance of classification models in recognizing human emotions.

To address the complexity of EEG signals, recent studies have transformed them into image-like formats before utilizing convolutional neural networks (CNN) for classification. This method leverages the proficiency of CNNs in image classification and noise reduction, thereby simplifying the model training process for EEG data analysis. Although this approach has facilitated the handling of EEG data, it may not fully capture the nonlinear and nonstationary nature of EEG signals [22], [23], [24], [25], [26]. Recurrence plots (RPs) are a powerful method for encoding time-series data, and are particularly useful for revealing the hidden complex nature of nonlinear and nonstationary signal patterns [27], [28], [29], [30], [31], [32].

This study investigated the effects of thresholded versus unthresholded RP on emotion recognition by examining their impact on the texture and informational content of the resulting image representations. A critical gap is that both thresholded and unthresholded RPs produce symmetric images, leading to redundancy and increased computational memory requirements [33].

To overcome these challenges, this study introduces a novel encoding technique called the asymmetric windowing recurrence plots (AWRP). Unlike traditional RPs, AWRP generates asymmetric images by applying windowing and asymmetric RP generation. This approach is innovative in its ability to efficiently capture the dynamic properties of EEG signals in textured image form, thereby minimizing computational memory demands while enhancing emotion recognition accuracy. This method addresses the redundancy issue and offers a more detailed representation of EEG data.

The primary contributions of this study are twofold.

1) First, it offers a detailed examination of thresholded and unthresholded RP in emotion recognition, focusing on how texture variations affect classification accuracy. This analysis aims to optimize the performance of emotion recognition systems.
2) Second, this study introduces AWRP, a novel technique for generating EEG signal image representations. This technique improves the classifier performance by enhancing the input texture quality without imposing significant computational memory demands.

The effectiveness of AWRP was validated through a CNN-based classifier and evaluated on well-known datasets (DEAP [34] and SEED [35]) using subject-specific 10-fold cross-validation. This evaluation process demonstrates the advantages of AWRP over existing methodologies.

The remainder of this paper is organized as follows. Related work is reviewed in Section II, followed by details of the methodology in Sections III. The results are presented in Section IV, followed by the conclusions in Section V. Each section was carefully articulated to convey insights, comparisons, and directions for future research.

II. RELATED WORKS

Various methodologies have characterized the advancement of EEG-based emotion recognition systems to enhance their accuracy and robustness. Initial efforts focused on the inherent time-series nature of EEG data. Iyer et al. [17] explored this approach by segmenting EEG signals into frequency bands, computing differential entropy, and employing a combination of CNN and long short-term memory (LSTM) networks for emotion detection. This method highlights the significance of frequency-specific brain activity in emotion detection. Although LSTM architectures are adept at capturing temporal dependencies, the overall method does not explicitly address the nonlinear dynamics within EEG signals.

Wirawan et al. [36] proposed a continuous capsule network to preserve spatial information in EEG signals, thus minimizing data loss and emphasizing the importance of spatial integrity. They segmented EEG signals per second and utilized bandpass filters for frequency-band decomposition. Differential entropy (DE) was subsequently applied to feature extraction. However, this approach, being primarily linear, may not adequately encapsulate the nonlinear dynamic characteristics of EEG signals.

Further advancements were made by Liu et al. [37], who developed a global-to-local feature aggregation network (GLFANet) to enhance the extraction of global and local features from EEG signals. This might capture aspects of the EEG's nonlinear dynamics. Nonetheless, the reliance on Butterworth bandpass filters and the focus on DE features indicate a fundamentally linear approach to feature extraction, which may not entirely embrace the nonlinearity of the signal.

Jiang et al. [19] introduced an attention mechanism-based multiscale feature fusion network (AM-MSFFN) to tackle the challenges posed by low signal-to-noise ratios and the variability of EEG signals among individuals. Although this approach can potentially capture some dynamics, explicitly handling nonlinear dynamics is not evident unless spatial-temporal convolution and multiscale feature fusion blocks are specifically designed to address such complexities.

These studies underscore the need for adaptive and personalized emotion recognition systems.

Continuing this trajectory, Gong et al. [21] extracted DE features and implemented an attention mechanism for transformer architecture that may capture dynamic behavior. However, the linear nature of DE features may not be sufficient to encapsulate the full extent of EEG signal nonlinearity. Similarly, Song et al. [20] explored hybrid models that merge convolutional modules and transformer architecture. However, convolution modules are not designed explicitly to handle the nonlinear dynamics of EEG signals. Li et al. [18] focused on improving feature extraction and model generalization for real-world applications through the spatial-temporal feature fused convolutional graph attention network (STFCGAT). However, they also rely on DE features, which are insufficient for handling the nonlinear dynamic properties of EEG signals.

In the context of depression recognition using EEG signals, Shen et al. [38] and Shen et al. [39] presented novel methodologies aimed at optimizing EEG signal spatial information through adaptive channel fusion and targeting the nonlinear and nonstationary characteristics of EEG data using empirical mode decomposition (EMD). Despite EMD's ability to capture these complex characteristics, Chen et al. [40] noted challenges such as mode mixing and difficulty in distinguishing between noise and intrinsic features, which can lead to misleading interpretations.

The transition from time-series analysis to the transformation of EEG data into image-like representations has emerged as a response to the limitations of traditional time-series methods in capturing the complex nonlinear dynamics of emotional states. By encoding EEG data into images, the field aims to harness the robust feature extraction capabilities of CNNs. Zhang et al. [22] and Asghar et al. [41] studied this area, demonstrating the adaptability of EEG data for image-based analysis utilizing spectrograms. Subsequent studies by Farokhah et al. [24] and Huang et al. [30] further explored the benefits of combining frequency-based transformations (scalograms) and recurrence plots with deep learning models, emphasizing their utility in stable classification performance and unveiling complex EEG signal dynamics.

The aforementioned methodologies primarily rely on linear strategies and do not explicitly address the nonlinear dynamics of EEG signals during the input formulation stage, thereby potentially overlooking these complex characteristics prior to classification. This critical insight reveals a significant gap in the current understanding of and approaches to EEG data analysis. In particular, the use of recurrence plots, which are promising because of their ability to handle nonlinear dynamics, has been limited by the production of symmetric images, leading to redundancy and increased computational demands, as noted by Zhang et al. [33]. In addition, the literature lacks a comprehensive exploration of thresholded versus unthresholded recurrence plots, and the computational implications of encoding lengthy EEG time-series into recurrence plots present a substantial gap. This study addresses these issues by introducing AWRP encoding and evaluating its effectiveness in enhancing image representations for improved classifier performance and efficient computational memory resource management. This innovative approach positions AWRP as a significant advancement in EEG-based emotion recognition that effectively bridges the existing research gaps.

TABLE 1. EEG Signals datasets for emotion recognition.

	DEAP	SEED
Subject	32 subjects	15 subjects
Experiment	1 session	3 sessions
Trial	40 trials	15 trials
Data	Preprocessed	Preprocessed
Channel	32 channels	62 channels
Sampling rate	128 Hz	200 Hz
Class	HAHV, HALV, LAHV, LALV	Positive, Neutral, Negative

III. MATERIALS AND METHODS

This section outlines the datasets and methodology employed in this study.

A. DATASETS

The effectiveness of the proposed AWRP encoding method in EEG-based emotion recognition tasks was evaluated using the preprocessed DEAP [34] and SEED [35] datasets. The DEAP dataset, designed for emotion analysis, comprises EEG and peripheral physiological signals from 32 participants who watched 40 one-minute-long music videos, resulting in 1280 trials. Each trial was labeled with valence, arousal, dominance, and liking on a scale from 1 to 9. The EEG signals, recorded with 32 electrodes and sampled at 512 Hz, underwent preprocessing for artifact removal and were downsampled to 128 Hz. For this study, the emotion recognition task focused on valence and arousal, categorizing them into high and low, resulting in four multiclass classification tasks: low arousal-low valence (LALV), low arousal-high valence (LAHV), high arousal-low valence (HALV), and high arousal-high valence (HAHV). The SEED dataset, exclusively focusing on EEG signals, captured data from 15 participants exposed to 15 film clips to elicit positive, neutral, and negative emotions across the three sessions. Recorded using a 62-channel EEG acquisition system at a 1000 Hz sampling rate, the EEG signals underwent preprocessing steps, such as eye blink artifact removal and bandpass filtering, and were downsampled to 200 Hz. With approximately four minutes of trial duration and film clip length, the dataset offers diverse emotional states for validating the proposed encoding methods in emotion recognition tasks. A summary of the DEAP and SEED dataset details is presented in Table 1

B. RECURRENCE PLOTS

Recurrence plots serve as a potent graphical tool for analyzing dynamic systems and elucidating the recurrence of similar states over time within a system. These plots yield visual insights into the inherent patterns and correlations in time-series data, thereby aiding in identifying deterministic or chaotic dynamics within a system. Each time-series can be represented by an m-dimensional trajectory, where m denotes the number of variables describing the system's state at each time point. The time-delay embedding technique is pivotal for RP analysis because it facilitates the reconstruction of a high-dimensional representation of the original system from time-series data. With a time delay of τ, this technique enables the construction of an m-dimensional trajectory representing the system states s_i, thereby providing a detailed view of the system's behavior for enhanced analytical accuracy [27], [42]. For a time-series signal comprising n data points, denoted by $x_1, \ldots, x_n$, the m-dimensional trajectories of the system states, s_i, can be generated through time-delay embedding with τ, as expressed in

$$\vec{s}_i = (x_i, x_{i+\tau}, \ldots, x_{i+(m-1)\tau}), \quad \forall_i \in \{1, \ldots, n-(m-1)\tau\}. \tag{1}$$

A thresholded RP, denoted by RP_T, is obtained by applying a threshold θ to the distance matrix $D_{ij} = \|\vec{s}_i - \vec{s}_j\|$, where $\forall_{i,j} \in \{1, \ldots, n-(m-1)\tau\}$. This process yields a binary matrix, as expressed in

$$RP_T(i,j) = \begin{cases} 1, & \text{if } D_{ij} \leq \theta, \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

C. INFORMATION RETENTION IN THRESHOLDED-UNTHRESHOLDED RECURRENCE PLOTS

The concept of information retention is pivotal in the analysis of time-series data, particularly when the RP is used in a classification framework. Two types of RPs are commonly employed: thresholded and unthresholded RPs. This section aims to mathematically demonstrate that the unthresholded RP, denoted by RP_U, inherently retains more information than RP_T. In RP_U, the distance matrix D_{ij} is fully represented, where each element D_{ij} corresponds to the distance between points i and j in the phase space. The entropy $H(RP_U)$ of RP_U can be calculated with

$$H(RP_U) = -\sum_k p(k) \log_2 p(k). \tag{3}$$

Here, $p(k)$ is the probability of distance k occurring in D_{ij}, and the sum is over all unique distances.

In contrast, the entropy $H(RP_T)$ of an RP_T is inherently limited by the binary nature of its elements and can be calculated as in

$$H(RP_T) = -[p(1) \log_2 p(1) + p(0) \log_2 p(0)]. \tag{4}$$

Given that $H(RP_U)$ is based on a richer set of distance values than the binary set in $H(RP_T)$, it is evident that $H(RP_T)$ has the potential to capture more information. Therefore, it can be mathematically asserted that RP_U retains more information than RP_T.

Algorithm 1 AWRP Encoding Procedure

1: **Input:** Time-series data $X = \{x_1, x_2, \ldots, x_T\}$, positive integer m
2: **Output:** AWRP images
3: **Process Starts:**
4: $n \leftarrow 2m^2$ {Determine the number of windows}
5: Partition X into n non-overlapping windows $\{W_1, W_2, \ldots, W_n\}$, each of equal length
6: **for** $k = 1$ to n **do**
7: Generate unthresholded RP (RP_{U_k}) for window W_k
8: Decompose (RP_{U_k}) into $U(RP_{U_k})$ and $L(RP_{U_k})$
9: **end for**
10: **for** $l = 1$ to $\frac{n}{2}$ **do**
11: Construct ARP_l using the upper and lower triangular components of adjacent RPs:
12: $ARP_l \leftarrow \left[L(RP_{U_{2l-1}}) \mid U(RP_{U_{2l}})\right]$
13: **end for**
14: Arrange $\frac{n}{2}$ ARPs in an $m \times m$ grid to form the final AWRP:

$$AWRP = \begin{bmatrix} ARP_{(1,1)} & ARP_{(1,2)} & \cdots & ARP_{(1,m)} \\ ARP_{(2,1)} & ARP_{(2,2)} & \cdots & ARP_{(2,m)} \\ \vdots & \vdots & \ddots & \vdots \\ ARP_{(m,1)} & ARP_{(m,2)} & \cdots & ARP_{(m,m)} \end{bmatrix}$$

15: Where each $ARP_{(i,j)}$ is given by:

$$ARP_{(i,j)} = ARP_{(i-1)\cdot m+j}$$

16: return AWRP image

D. PROPOSED ASYMMETRIC WINDOWING RECURRENCE PLOTS

As proven in the previous section, RP_U has richer information than RP_T, RP_U is selected to generate the proposed AWRP. The following is a detailed explanation of Algorithm 1 regarding the AWRP encoding procedure:

1) **Windowing of Time-Series Data**
Given a time-series $X = \{x_1, x_2, \ldots, x_T\}$, the series is partitioned into n non-overlapping windows. Each window W_k is a continuous, non-overlapping subset of X and is of equal length. Index k ranges from 1 to n, ensuring that every data point in the time-series is utilized. The number n is determined using the formula $n = 2m^2$ for a positive integer m. This choice ensures that the final AWRP has a square shape. By setting the number of windows to $n = 2m^2$, it is guaranteed that there will be m^2 asymmetric recurrence plots (ARPs) after combining the adjacent windows. Arranging these m^2 ARPs in an $m \times m$ grid results in a square-shaped AWRP. This structure is particularly advantageous

when employing CNN-based classifications for analysis because square-shaped inputs are often preferable for standard CNN architectures.

2) **Generation of Unthresholded RP per Window**
For each window W_k, where $k = 1, 2, \ldots, n$, an unthresholded RP, denoted by RP_{U_k}, is generated. This plot captures the recurrence structure of the states within the respective window. Each RP_{U_k} is a square matrix.

3) **Extraction of Triangular Components**
Each RP_{U_k} is decomposed into its upper and lower triangular components, denoted respectively as $U(RP_{U_k})$ and $L(RP_{U_k})$.

4) **Construction of Asymmetric Recurrence Plots**
A series of $\frac{n}{2}$ ARPs were constructed using the upper and lower triangular components of the adjacent RPs. For $l = 1$ to $\frac{n}{2}$, each ARP_l is formed by concatenating the lower and upper triangular RP_U components, as shown in

$$ARP_l = \left[L(RP_{U_{2l-1}}) \mid U(RP_{U_{2l}})\right]. \tag{5}$$

5) **Expressing the Final AWRP**
Let $ARP_{(i,j)}$ represent the ARP in the i^{th} row and j^{th} column of the grid. The final AWRP then arranges the $\frac{n}{2}$ ARPs in an $m \times m$ grid to form a square matrix. The general form of AWRP can be written as

$$AWRP = \begin{bmatrix} ARP_{(1,1)} & ARP_{(1,2)} & \cdots & ARP_{(1,m)} \\ ARP_{(2,1)} & ARP_{(2,2)} & \cdots & ARP_{(2,m)} \\ \vdots & \vdots & \ddots & \vdots \\ ARP_{(m,1)} & ARP_{(m,2)} & \cdots & ARP_{(m,m)} \end{bmatrix}. \tag{6}$$

For a specific $ARP_{(i,j)}$, it is expressed in

$$ARP_{(i,j)} = ARP_{(i-1)\cdot m+j}. \tag{7}$$

This representation ensures that the AWRP is square-shaped, which is suitable for CNN architectures.

Fig. 1 illustrates how the mathematical model generates an AWRP image with eight windows ($n = 8$). The 1D-EEG signals from the DEAP dataset were encoded without thresholding using dimensions of $m = 5$ and a time delay of $\tau = 150$.

E. INFORMATION RETENTION IN AWRP OVER UNTHRESHOLDED RP

AWRP images were generated, as depicted in Fig. 1 shows rich textures. This distinctive characteristic is beneficial when utilizing a CNN-based classifier, which is known for its robust performance in image classification tasks. This section provides a mathematical justification to show that AWRP has richer textures than RP_U. The AWRP consists of multiple ARPs, each derived from a different window in the time-series. The entropy of the individual ARPs is formulated in

$$H(ARP_l) = -\sum_k p(k|l) \log_2 p(k|l). \tag{8}$$

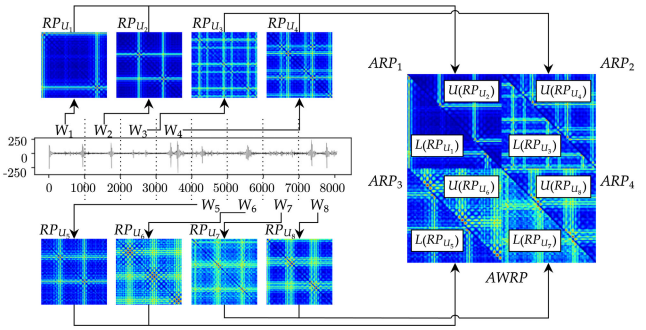

FIGURE 1. The proposed AWRP encoding on DEAP dataset with 8 non-overlap windowing (m=5,τ=150).

Here, $p(k|l)$ is the conditional probability that distance k occurs in the l^{th} ARP.

Because ARPs are initially generated from non-overlapping windows of the time-series, the aggregate entropy of AWRP can be considered as the sum of the entropies of individual ARPs, as formulated in

$$\begin{aligned} H(AWRP) &= \sum_{l=1}^{\frac{n}{2}} H(ARP_l) \\ &= -\sum_{l=1}^{\frac{n}{2}} \sum_k p(k|l) \log_2 p(k|l). \end{aligned} \tag{9}$$

This aggregate entropy is likely to be higher than the entropy of a single RP_U because it captures variations across different time-series segments. To validate this statement, this study conducted empirical validation through an ablation study using a windowing number. The empirical validation of the proposed methodology is described in the experimental setup section.

F. EXPERIMENTAL SETUP

This section explains the emotion recognition pipeline used in this study, as illustrated in Fig. 2. The EEG signal encoding and classification setting was implemented on a high-performance computing device with a 3.60 GHz CPU and DGX-1 V100 32 GB VRAM.

1) EEG SIGNALS ENCODING

This study empirically analyzed thresholded and unthresholded RPs using the proposed AWRP method. For the AWRP, five non-overlapping windows are implemented with window sizes of $n = 8, 18, 32, 50$, and 72. During the encoding process of EEG signals, various values of dimension n and time delay τ were employed. Table 2 provides a detailed overview of the values for dimension n and time delay τ used in each encoding scheme.

2) CLASSIFICATION SETTINGS

This study employs a variant of the CNN proposed by LeCun et al. [43] known as LeNet-5. The LeNet-5

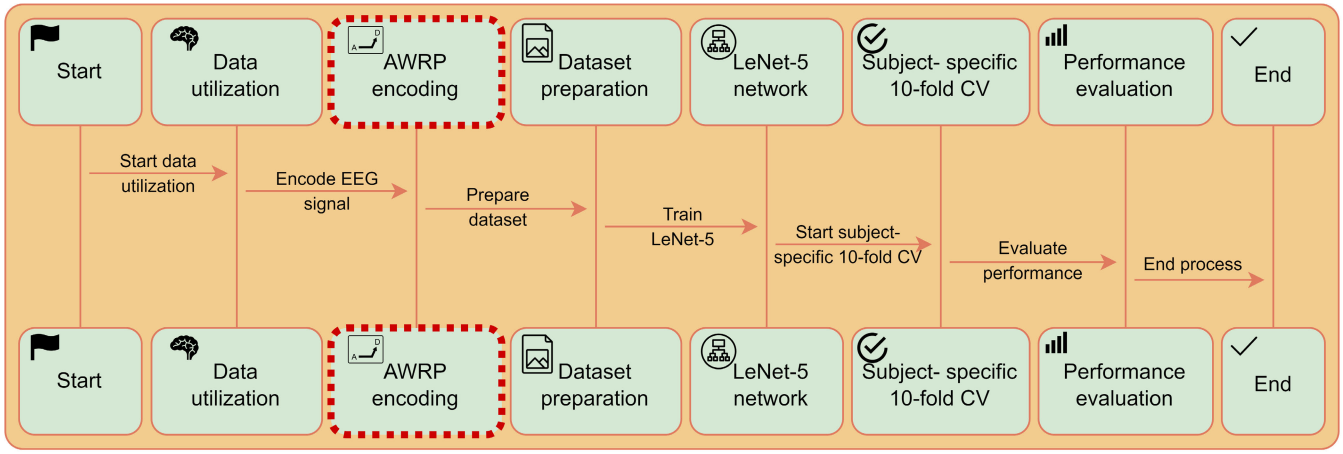

FIGURE 2. Emotion recognition pipeline using the proposed AWRP encoding.

TABLE 2. Dimension and time delay parameter values used in EEG signals encoding.

	DEAP		SEED	
	Dimension	Time delay	Dimension	Time delay
RP_T	9	950	9	950
RP_U	9	950	9	950
$AWRP_8$	5	150	9	550
$AWRP_{18}$	7	50	6	400
$AWRP_{32}$	5	35	5	250
$AWRP_{50}$	4	30	4	200
$AWRP_{72}$	3	21	4	150

†The time delay unit is the discrete-time point within the specified window size.

TABLE 3. Summary of modified LeNet-5 architecture.

Layer Type	Units/Nodes	Activation	Additional Layers
Input	-	-	-
C1: Conv	6 filters of 5×5	Leaky ReLU	Dropout, BatchNorm
S2: Avg Pool	2×2 pool size	-	Dropout, BatchNorm
C3: Conv	16 filters of 5×5	Leaky ReLU	Dropout, BatchNorm
S4: Avg Pool	2×2 pool size	-	BatchNorm
C5: FC	120	Leaky ReLU	Dropout
F6: FC	84	Leaky ReLU	Dropout
Output	Varied (4 or 3)	Softmax	-

architecture comprises two convolutional layers, two average pooling layers, and three fully connected layers. The final layer of the network contains 10 neurons, producing a 10-dimensional vector as the output. This vector represents the predicted probability distribution across ten possible classes corresponding to the digits in handwritten digit recognition. The input layer size was adjusted to 150×150. Given that the DEAP and SEED datasets contain four and three emotion labels, respectively, the number of neurons in the last layer of LeNet-5 was modified to match the number of labels in each dataset. The activation function used for each convolutional layer and the first and second fully connected layers is a leaky rectified linear unit (Leaky ReLU). The softmax function was applied to the final fully connected layer. Dropout layers were inserted following layers C1, S2, and C3, as well as the first and second fully connected layers, to prevent overfitting. Batch normalization layers were added after layers C1, S2, C3, and S4 to accelerate the training process. The feature extraction process is seamlessly integrated into the end-to-end learning framework of LeNet-5. The architecture of LeNet-5 is summarized in Table 3.

3) EVALUATION METHOD

In this study, EEG signal datasets were encoded into image datasets using AWRP. These images were subsequently classified using the LeNet-5 model to evaluate the efficacy of AWRP for emotion recognition. The methodological framework adopted for this evaluation was a subject-specific 10-fold cross-validation approach, as shown in Fig. 3. This approach entailed dividing each subject's AWRP image samples into ten subsets. For each validation fold, one subset was allocated as the validation set to assess the model's performance, whereas the ensemble of the remaining nine subsets constituted the training set. This process was executed ten times, ensuring that each subset was used once as the validation set, thereby facilitating ten distinct evaluations of the model's performance per subject. The results of these iterations were averaged to derive the overall performance metrics. To quantify the model's performance, metrics such as accuracy, F1-score, and Kappa score were computed for each subject across all the folds. To derive the robust performance metrics, the results of these iterations were averaged to produce an overall assessment. The computed metrics included accuracy, F1-score, and Kappa score, which were aggregated for each participant. Consolidated metrics representing the average performance across all subjects were reported in this study. Specifically, these metrics were averaged over 32 subjects for the DEAP dataset and 15 subjects for the SEED dataset, providing a comprehensive view of the effectiveness of the model for emotion recognition from EEG signal images. The Kappa score was computed using (10), where p_0 signifies the observed agreement and p_e denotes the expected agreement.

$$\kappa = \frac{p_0 - p_e}{1 - p_e} \tag{10}$$

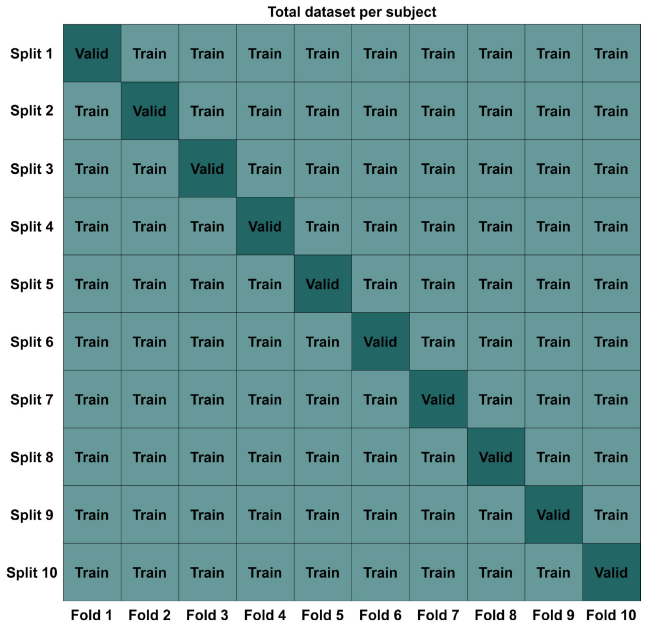

FIGURE 3. The subject-specific 10-fold cross-validation approach.

These metrics served as a critical comparative measure to evaluate the performance of the LeNet-5 model with AWRP inputs compared to those obtained using traditional thresholded and unthresholded recurrence plot inputs. Statistical evaluations, including the paired t-test, Wilcoxon, ANOVA, and Friedman tests, were performed to determine the superiority of the input representation method. The study also included an entropy analysis to assess information retention and complexity as well as to evaluate the computational efficiency of AWRP in terms of memory usage and runtime.

4) HYPERPARAMETER DETAILS

Hyperparameter optimization for the model was performed using the Bayesian optimization method [44], with a focus on tuning two primary hyperparameters: dropout ratio and learning rate. The dropout ratio was tested at two levels, specifically 0.2 and 0.3, while the learning rate was selected from the set $[1 \times 10^{-2}, 1 \times 10^{-3}, 1 \times 10^{-4}]$. A λ^2 regularization term with a coefficient set to 0.01 was incorporated to prevent overfitting during the training process. The Adam optimizer was employed for model compilation using sparse categorical cross-entropy as the loss function to guide the optimization process. Table 4 presents the optimal hyperparameters identified for the AWRP variants (8, 18, 32, 50, and 72) applied to the DEAP and SEED datasets using LeNet-5.

TABLE 4. LeNet-5 Optimal Hyperparameters used to classify $\text{AWRP}_{n=8,18,32,50,72}$ inputs on DEAP and SEED Datasets.

Dataset	Encoding	$d1$	$d2$	$d3$	$d4$	$d5$	lr	Epoch
DEAP	RP_T	0.3	0.2	0.3	0.2	0.2	1×10^{-4}	44
	RP_U	0.3	0.2	0.3	0.3	0.2	1×10^{-4}	45
	AWRP_8	0.3	0.2	0.3	0.3	0.2	1×10^{-4}	35
	AWRP_{18}	0.2	0.2	0.2	0.2	0.2	1×10^{-4}	45
	AWRP_{32}	0.2	0.2	0.3	0.3	0.3	1×10^{-4}	46
	AWRP_{50}	0.2	0.2	0.3	0.3	0.2	1×10^{-4}	48
	AWRP_{72}	0.3	0.2	0.3	0.2	0.3	1×10^{-4}	48
SEED	RP_T	0.3	0.2	0.2	0.2	0.3	1×10^{-4}	46
	RP_U	0.2	0.2	0.3	0.3	0.2	1×10^{-4}	45
	AWRP_8	0.3	0.3	0.3	0.2	0.2	1×10^{-3}	45
	AWRP_{18}	0.3	0.3	0.2	0.3	0.2	1×10^{-4}	43
	AWRP_{32}	0.3	0.2	0.3	0.2	0.2	1×10^{-4}	43
	AWRP_{50}	0.3	0.2	0.2	0.3	0.2	1×10^{-4}	43
	AWRP_{72}	0.3	0.2	0.2	0.3	0.2	1×10^{-4}	50

IV. RESULT AND DISCUSSION

A. CLASSIFICATION PERFORMANCE ON DIFFERENT WINDOW SIZES

1) PERFORMANCE VISUALIZATION ON THE DEAP DATASET

The box plots shown in Fig. 4a, 4b, and 4c illustrate the aggregated performance metrics computed for each participant across all folds in the 10-fold cross-validation of the DEAP dataset. These metrics provide a visual quantification of the efficacy of each input formulation. For accuracy, most methods showed a tight interquartile range with median values close to one, indicating a generally high and consistent performance across subjects. Notably, the RP_T method had a slightly lower median accuracy and wider interquartile range than the other methods, indicating more variability in its performance. The outliers were minimal, with a few observed in the AWRP_{18} and AWRP_{50} methods. F1 Score distributions mirror the patterns observed in accuracy, with RP_T showing more variability. The Kappa statistic, which measures the agreement between predicted and actual classifications corrected for chance, presents a similar trend. Although most methods achieve high median Kappa values, the RP_T method demonstrates increased variability, with its box plot extending lower than those of the others. Outliers are scarcely observed, underscoring the robustness of the findings.

2) PERFORMANCE VISUALIZATION ON THE SEED DATASET

Fig. 4d, 4e, and 4f in the study illustrate the compiled results of the accuracy, F1 Score, and Kappa coefficient metrics for a variety of input formulation methods, applied to the SEED dataset. These box plots summarize the central tendency and variability of the performance of the methods, and were derived by averaging the individual scores of each subject across all folds of the 10-fold cross-validation process. Each method achieved impressive median values above 0.95 for the accuracy and F1 Score, signifying proficient performance in classifying emotional states based on the SEED dataset. A close examination revealed that AWRP_{32} and AWRP_{72} showed tight interquartile ranges across all three performance metrics, underscoring their consistent and reliable outcomes across different subjects. Conversely, the RP_T method manifested a broader distribution of scores, indicating a degree of variability in its performance. No significant outliers were discerned within the box plots, implying that all the methods yielded consistent results

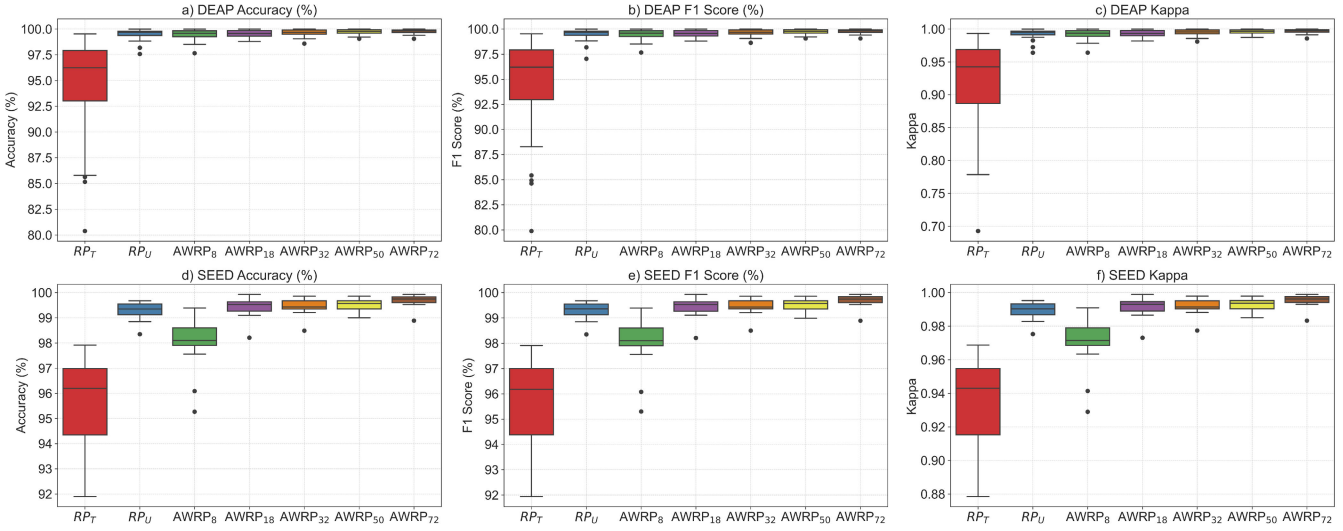

FIGURE 4. The figure presents boxplot distributions for classification performance metrics, accuracy, F1 Score, and Kappa across subject-specific 10-fold cross-validation for methods RP_T, RP_U, and AWRP variants (8, 18, 32, 50, and 72) within the DEAP and SEED datasets segmented into six panels. Panels a) and d) focus on the accuracy distributions for DEAP and SEED, respectively; b) and e) present the F1 Score distributions; and c) and f) illustrate the Kappa distributions. Each boxplot reflects the distribution of metrics aggregated from the individual subject performance, highlighting the subject-dependent validation approach.

without extreme deviations. Therefore, the plots demonstrate that the $AWRP_{32}$, $AWRP_{50}$, and $AWRP_{72}$ methods not only provide high median values but also ensure stability and reliability in their performance across various subjects. This consistency is crucial for developing a robust BCI that can effectively recognize and categorize human emotions and emotional states across individuals.

B. HYPOTHESIS TEST ANALYSIS

1) PAIRED HYPOTHESIS TEST ON CLASSIFICATION PERFORMANCE ON THE DEAP DATASET

In the statistical analysis of the DEAP dataset, a paired t-test was applied to evaluate the performance disparities among various input formulation methods, with each method's performance metrics calculated as the average of the outcomes for each subject across all folds of the cross-validation process. The results in Table 5 confirm the significant differences in performance. The RP_T method consistently underperformed in terms of accuracy, F1 Score, and Kappa compared with visual observations. The AWRP variants displayed significant differences, emphasizing the impact of the window size on the performance. $AWRP_{72}$ emerged as the top performer with the highest average accuracy (99.77%), F1 Score (99.77%), and kappa (0.9966), indicating its superior performance. Further statistical hypothesis testing is required to validate these results with RP_U as the baseline. The p-values from the paired t-tests suggest insufficient evidence to reject the null hypotheses for $AWRP_8$ and $AWRP_{18}$. However, $AWRP_{32}$ shows lower p-values, which provides evidence against the null hypothesis. A significant shift was observed for $AWRP_{50}$ and $AWRP_{72}$, where the p-values were below 0.05, indicating statistically significant performance improvements over RP_U. These results, particularly those for $AWRP_{72}$, highlight the superior classification performance of the DEAP dataset. In summary, $AWRP_{72}$'s high-average performance metrics and significant p-values demonstrate its effectiveness for classification tasks on the DEAP dataset.

TABLE 5. Summary of average performance metrics and paired t-test p-values using the DEAP dataset, aggregated from subject-specific 10-fold cross-validation. The significance was tested at $\alpha = 0.05$ to evaluate the efficacy of the method in emotion classification.

Average Performance Metrics: Method	Accuracy	F1 Score	Kappa
RP_T	94.55%	94.49%	0.9195
RP_U	99.51%	99.49%	0.9928
$AWRP_8$	99.46%	99.46%	0.9922
$AWRP_{18}$	99.54%	99.54%	0.9933
$AWRP_{32}$	99.64%	99.64%	0.9948
$AWRP_{50}$	99.71%	99.71%	0.9958
$AWRP_{72}$	**99.77%**	**99.77%**	**0.9966**
Paired t-test Results (p-values): Comparison	Accuracy	F1 Score	Kappa
RP_U vs $AWRP_8$	0.6047	0.7847	0.6374
RP_U vs $AWRP_{18}$	0.7272	0.6292	0.6880
RP_U vs $AWRP_{32}$	0.1703	0.1705	0.1500
RP_U vs $AWRP_{50}$	0.0347	0.0437	0.0345
RP_U vs $AWRP_{72}$	0.0045	0.0077	0.0040

2) PAIRED HYPOTHESIS TEST ON CLASSIFICATION PERFORMANCE ON THE SEED DATASET

Table 6 displays $AWRP_{72}$ with the highest average values of 99.69% accuracy and an F1 Score of 0.9953, outperforming other AWRP variants. $AWRP_{50}$ and $AWRP_{32}$ also showed noteworthy performance, although it was slightly lower than that of $AWRP_{72}$. $AWRP_8$, but exceeded by its peers,

TABLE 6. Summary of average performance metrics and Wilcoxon test *p*-values using the SEED dataset aggregated from subject-specific 10-fold cross-validation. Significance tested at $\alpha = 0.05$ to evaluate method efficacy in emotion classification.

Average Performance Metrics: Method	Accuracy	F1 Score	Kappa
RP_T	95.73%	95.73%	0.9359
RP_U	99.29%	99.29%	0.9893
AWRP-8	97.96%	97.96%	0.9694
AWRP-18	99.42%	99.42%	0.9914
AWRP-32	99.47%	99.47%	0.9920
AWRP-50	99.52%	99.52%	0.9927
AWRP-72	**99.69%**	**99.69%**	**0.9953**
Paired Wilcoxon Test Results (p-values): Comparison	Accuracy	F1 Score	Kappa
RP_U vs $AWRP_8$	6.10×10^{-5}	6.10×10^{-5}	6.10×10^{-5}
RP_U vs $AWRP_{18}$	3.52×10^{-2}	3.02×10^{-2}	3.02×10^{-2}
RP_U vs $AWRP_{32}$	3.05×10^{-4}	3.05×10^{-4}	3.05×10^{-4}
RP_U vs $AWRP_{50}$	4.28×10^{-3}	4.25×10^{-3}	4.26×10^{-3}
RP_U vs $AWRP_{72}$	6.10×10^{-5}	6.10×10^{-5}	6.10×10^{-5}

still surpassed the baseline RP_U method, which recorded 99.29% accuracy and F1 Score and 0.9893 for kappa. The paired Wilcoxon test results in Table 6 highlight performance discrepancies between the methods. This statistical test was chosen because of its small sample size and non-normal data distribution, making it more suitable than parametric tests. This decision aligns with the characteristics of the dataset, in which the assumption of a normal distribution is inappropriate. The *p*-values obtained by comparing RP_U with various AWRP methods were notably low, often approximately 6.10×10^{-5}. These values, which are significantly below the $\alpha = 0.05$ threshold, clearly indicate the statistical significance of the improvements achieved by AWRP methods over RP_U. Comparisons involving $AWRP_{18}$ also yielded *p*-values well below the α level, confirming its statistically significant superior performance compared to RP_U. In conclusion, the table provides strong evidence of the statistical and practical significance of the AWRP method performance. $AWRP_{72}$ was the most effective, substantiating the efficacy of the AWRP methods, particularly $AWRP_{72}$, in emotion recognition tasks using the SEED dataset.

C. COMPARATIVE ENTROPY ANALYSIS OF ENCODING METHODS

The line graph in Fig. 5 shows the entropy values associated with the various encoding methods applied to the DEAP and SEED datasets. This approach involves computing the average entropy of 100 randomly selected samples from each encoding method. The results indicate that RP_T has notably low entropy values for both datasets, with 0.03 and 0.04 for DEAP and SEED, respectively. This minimal entropy suggests that RP_T encapsulates limited complexity and information content within the generated images. Because of the binary nature of RP_T, significant information may be lost during the encoding process. Consequently, the images generated using this method may lack details and variability,

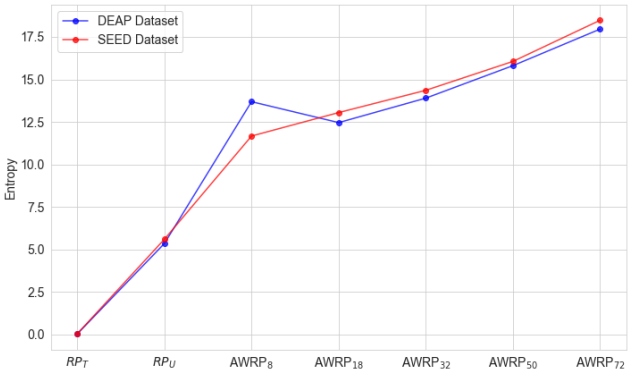

FIGURE 5. The average entropy of encoding methods on DEAP and SEED datasets.

making them less effective for emotion recognition tasks. DEAP and SEED datasets' entropy values were higher for RP_U than for the RP_T, registering 5.35 and 5.60 for RP_U and 4.58 and 4.76 for RP_T. This indicates that RP_U is better at retaining the information during encoding. The graph also shows that all variants of AWRP consistently have higher entropy values than both RP_T and RP_U. The entropy values for AWRP variants range from 13.69 to 17.96 for DEAP and from 11.67 to 18.48 for SEED, indicating that AWRP is better at capturing complex dynamics within EEG signals, making it a promising candidate for EEG-based emotion recognition tasks.

D. MEMORY USAGE AND RUNNING TIME ANALYSIS

To determine the computational efficiency of different signal-to-image encoding methods, the average memory usage and running time for generating 100 image samples were recorded and presented. For the DEAP dataset (Fig. 6), AWRP with eight windows required notably more memory (approximately 103.20 MB) per image generation than the other methods did. Conversely, AWRP with 50 windows utilized the most efficient memory usage, approximately 30.92 MB per average. Although the memory usage among the methods had a relatively minimal standard deviation, indicating a stable performance, the running time gradually increased as the number of windows in AWRP increased, with $AWRP_{72}$ taking the longest time (approximately 1016.65 s). In contrast, for the SEED dataset (Fig. 7) showed higher memory usage for both RP_T and RP_U, averaging 4776.18 MB and 4767.90 MB respectively. $AWRP_{72}$ was the most memory-efficient, using around 408.03 MB an average. The running time trend followed a similar pattern, with RP_U consuming the most time (1737.42 s) and $AWRP_8$ being the fastest (393.31 seconds per image generation). The trade-offs between the traditional and proposed AWRP variants are evident in the memory and running time trends. Although the AWRP methods generally performed better on the SEED dataset, the results varied for the DEAP dataset. Therefore, selecting an appropriate encoding method should consider both the dataset characteristics and the available computational resources.

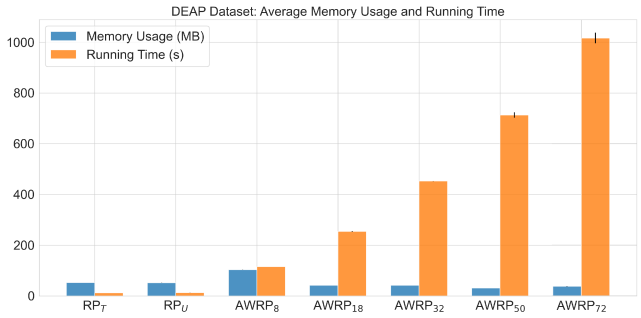

FIGURE 6. The average memory usage and running time of encoding methods in the DEAP dataset.

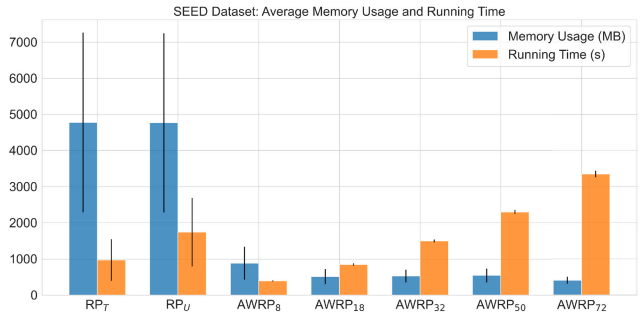

FIGURE 7. The average memory usage and running time of encoding methods in the SEED dataset.

E. AWRP WITH HIGHER WINDOW SIZE

The exceptional performance of $AWRP_{72}$ in analyzing the DEAP and SEED datasets underscores the necessity for a comprehensive examination of window sizes exceeding 72. This exploration is critical for determining whether the remarkable outcomes observed at a window size of 72 represent the apex of potential performance or if further optimization is achievable. Understanding the relationship between window size, denoted by n, and performance is crucial, especially in the context of time-series data analysis. The dimensionality of the analysis window profoundly affects both the interpretability and precision of the results, necessitating a systematic exploration of the performance across varying window sizes. To address this question, this study systematically investigated a spectrum of window sizes: 98, 128, 162, 200, 242, 288, and 338. These specific sizes were selected based on formula $n = 2m^2$ (for $m = 7, 8, \ldots, 13$), motivated by the aim of generating square-shaped AWRP samples. Such geometric considerations are important for the subsequent application of CNNs in the analysis phase, because CNNs prefer square-shaped inputs. The dimensions and time delay parameters corresponding to each window size are listed in Table 7.

1) VISUALIZATION OF PERFORMANCE TRENDS ACROSS AWRP HIGHER WINDOW SIZES

The performance trends shown in Fig. 8 show how the accuracy, F1 score, and Kappa statistics change with varying window sizes, based on the DEAP and SEED datasets. For the DEAP dataset, the accuracy increased slightly with larger window sizes and remained consistent even with larger dimensions, indicating reliable model performance. The reduced variance with larger window sizes also suggests an increased model reliability. The F1 score followed a similar pattern, demonstrating balanced precision and recall, and the Kappa statistic showed a slight upward trend, indicating strong agreement not attributed to chance, which highlights the robustness of the model. The best performance on the DEAP dataset was observed at $AWRP_{338}$ with accuracy, F1 score, and Kappa values of 99.84%, 99.84%, and 0.9977, respectively. In contrast, the performance metrics of the SEED dataset showed diverse characteristics. Initially, the accuracy increased with larger window sizes, plateaus, and then dipped at window sizes of 242, which could be attributed to a decrease in the capacity of the model to generalize. However, it increased again for the window sizes of 288 and 338. The F1 score follows a similar pattern, and the Kappa statistic experiences a significant increase with smaller window sizes, followed by minor fluctuations that maintain a high level of stability at larger sizes, suggesting good predictive performance even with variability. The highest performance on the SEED dataset was observed at $AWRP_{72}$ with an accuracy of 99.69%, and at $AWRP_{128}$ with an F1 score and Kappa values of 99.69% and 0.9953, respectively.

This study reveals the optimal window size range for both datasets, where the performance metrics remain stable, indicating the most suitable balance between performance and computational efficiency. For DEAP, the optimal window size range begins at 72, whereas for SEED, it starts from 98 onwards, with a slight drop in performance at 242. The consistent performance metrics at larger window sizes demonstrate the robustness of the method to variations within the identified range. However, the differences in the trends between the two datasets suggest dataset-specific characteristics, possibly owing to the unique emotional responses or EEG signal characteristics of each dataset.

TABLE 7. Dimension and time delay parameters for AWRP with higher window size.

	DEAP		SEED	
	Dimension	Time delay	Dimension	Time delay
$AWRP_{98}$	3	18	4	100
$AWRP_{128}$	3	18	4	80
$AWRP_{162}$	2	10	3	70
$AWRP_{200}$	2	10	3	70
$AWRP_{242}$	2	8	3	30
$AWRP_{288}$	2	5	3	20
$AWRP_{338}$	2	3	3	30

†The time delay unit is the discrete-time point within the specified window size.

2) STATISTICAL ANALYSIS OF AWRP PERFORMANCE AT HIGHER WINDOW SIZES

To evaluate the scalability and performance of the AWRP method, a range of window sizes was employed, ranging from 72 to 338, and an extreme value of 2592. For the

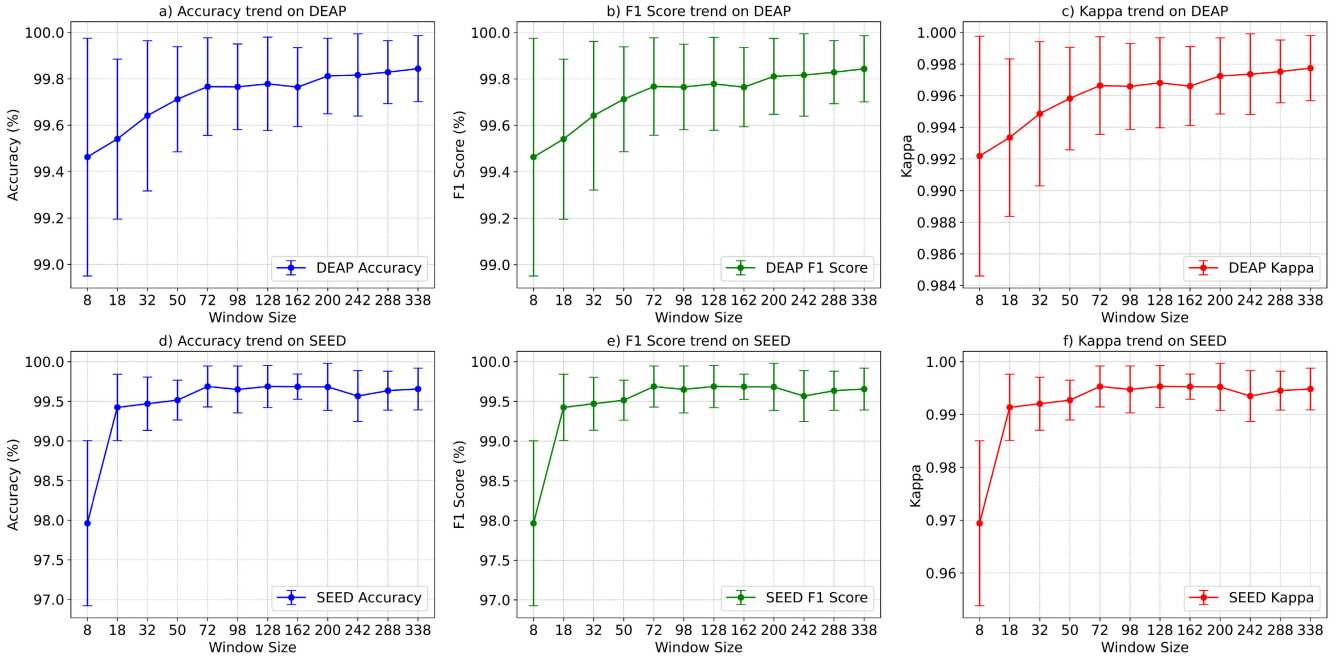

FIGURE 8. Trends in the performance metrics across various window sizes for the DEAP and SEED datasets were derived from subject-specific 10-fold cross-validation. a) Displays the Accuracy trend on DEAP, b) shows the F1 Score trend on DEAP, and c) presents the Kappa trend for DEAP. Similarly, d), e), and f) depict the trends for the Accuracy, F1 Score, and Kappa, respectively, on the SEED dataset.

2592 window size, the dimension and time delay parameters were two and one for the DEAP dataset and three and one for the SEED dataset, respectively. $AWRP_{2592}$ achieved impressive results on both the DEAP and SEED datasets, with accuracy scores of 99.77% on both datasets, F1 Scores of 99.77% for the DEAP dataset and 99.61% for the SEED dataset, and Kappa scores of 0.9966 for the DEAP dataset and 0.9942 for the SEED dataset. This range includes window sizes that have shown promise in the past and extends to a much larger size to test the AWRP method's performance limits. By examining both optimal and extreme window sizes, this study aimed to comprehensively assess the impact of window size on the ability of the AWRP method to maintain its effectiveness across various scales. To assess the impact of window size variations on model performance, statistical analyses were conducted, including ANOVA for the DEAP dataset and Friedman test for the SEED dataset, both with a significance level set at 0.05. The results of these statistical tests are presented in Table 8.

ANOVA on the DEAP dataset revealed F-statistics of 1.0387, 1.0321, and 1.0285 for accuracy, F1 score, and Kappa, respectively, with p-values of 0.4070, 0.4119, and 0.4146. These results suggest that window size does not significantly affect the performance metrics of the model. Similarly, the Friedman test on the SEED dataset produced chi-square values of 5.2349, 5.9533, and 5.2349 for accuracy, F1 score, and Kappa, respectively, with corresponding p-values of 0.7322, 0.6525, and 0.7322, further supporting the conclusion that the differences in performance across the various window sizes were not statistically significant.

TABLE 8. ANOVA and Friedman test results for AWRP performance by window size (98, 128, 162, 200, 242, 288, 338, and 2592) on the DEAP and SEED Datasets, with $\alpha = 0.05$. Performance metrics were based on averages from subject-specific 10-fold cross-validation, highlighting the subject-dependent validation approach.

ANOVA test results on the DEAP dataset:		
Metric	F-Statistic	p-values
Accuracy	1.0387	0.4070
F1 Score	1.0321	0.4119
Kappa	1.0285	0.4146
Friedman test results on the SEED dataset:		
Metric	Chi-Square	p-values
Accuracy	5.2349	0.7322
F1 Score	5.9533	0.6525
Kappa	5.2349	0.7322

The performance of the AWRP method in the tested range of window sizes appeared to be invariant, which is beneficial for practical applications. The absence of significant differences in performance across window sizes allows for flexibility in selecting a window size based on factors other than statistical performance, such as computational constraints or application-specific requirements. The AWRP method also demonstrated robustness against changes in window size, as evidenced by the inclusion of a large window size of 2592 without any significant performance deviation. This suggests that the method can be reliably implemented without the need for extensive window-size optimization.

TABLE 9. Comparison with related works.

Framework	Dataset	Input formulation	Feature Extraction	Classifier	Accuracy
Wirawan et al. [36]	DEAP	Time-series signals	Differential entropy	Continuous capsnet	91.35%
Liu et al. [37]	DEAP	Time-series signals	GLFANet	FCNN	92.92%
Jiang et al. [19]	DEAP	Time-series signals	AM-MSFFN	FCNN	99.32%
This study	DEAP	Image: AWRP	Automated feature extraction	LeNet-5	**99.84%**
Song et al. [20]	SEED	Time-series signals	Conformer	FCNN	95.30%
Jiang et al. [19]	SEED	Time-series signals	AM-MSFFN	FCNN	97.53%
Gong et al. [21]	SEED	Time-series signals	ACTNN	FCNN	98.47%
Li et al. [18]	SEED	Time-series signals	STFCGAT	FCNN	99.11%
This study	SEED	Image: AWRP	Automated feature extraction	LeNet-5	**99.69%**

F. DISCUSSION

1) RESULT FINDINGS

The results from the performance metrics show that AWRP and RP_U are more resilient to noise than RP_U owing to their improved encoding process that retains complex details. This results in more complex and informative encoded images, with higher entropy values and better classification outcomes. These trends suggest that increased entropy enhances the classification performance, making AWRP methods effective in generating distinctive and informative encoded images for accurate emotional state classification. An analysis of the window sizes in AWRP methods and performance metrics reveals an initial improvement in the classification performance with an increase in the window size, but plateaus beyond a certain threshold. This plateau indicates that a window size of 72 can be considered the starting point for the optimal range because further increases in window size do not lead to statistically significant improvements in performance metrics across the analyzed datasets. This indicated the importance of conducting preliminary experiments to determine the optimal window size for specific datasets and tasks.

2) LIMITATION

The analysis also revealed a trade-off between memory usage and running time during image generation using different encoding methods. Some variants of AWRP methods are memory-intensive, but expedite image generation. This tradeoff is crucial for applications constrained by memory or computational time. In memory-constrained scenarios, a tight encoding method for memory usage is prudent even if it requires longer running times. Conversely, a method that quickly generates images with sufficient memory in time-constrained applications is more appropriate. This discussion provides a granular analysis of the noise sensitivity and entropy value trends, the relationship between windowing size and performance, and the tradeoff between memory usage and running time for RP_T, RP_U, and the proposed AWRP encoding method. The insights gained from this analysis are essential for practitioners and researchers alike, guiding them in selecting an encoding method that is congruent with the requirements and constraints of their specific EEG-based emotion recognition tasks.

G. COMPARISON WITH RELATED WORKS

Table 9 shows an accuracy comparison between the proposed AWRP and the state-of-the-art methods. Notably, all referenced methods, including this study, adhere to subject-dependent validation, emphasizing the personalized aspect of EEG data analysis. The current study employed an innovative approach to EEG-based emotion recognition using image representations derived from the proposed AWRP and LeNet-5 for automated feature extraction and classification. The results were compared with those of existing frameworks, and showed improvements. For the DEAP dataset, Wirawan et al. [36] used a time-series signal input with differential entropy for feature extraction and a continuous capsnet for classification, achieving a 91.35% accuracy. By contrast, the proposed method utilizes image-based input formulations (AWRP), which provide a rich set of features for higher accuracy. Liu et al. [37] also used time-series signals from the DEAP dataset but introduced GLFANet for feature extraction and a fully-connected neural network (FCNN) for classification, resulting in 92.92% accuracy. However, the method used in this study outperformed the method with 99.84% accuracy. Jiang et al. [19] used an AM-MSFFN for time-series signal feature extraction on the DEAP dataset, achieving 99.32% accuracy. Although impressive, the current method outperformed it with 99.84% accuracy. For the SEED dataset, Song et al. [20] achieved 95.30% accuracy by using a convolutional transformer (conformer) for feature extraction from time-series signals. Gong et al. [21] used an attention-based convolutional transformer neural network (ACTNN) to achieve an accuracy of 98.47%. Li et al. [18] improved it to 99.11% with a STFCGAT. However, the method in this study outperformed all the other methods with 99.69% accuracy using the AWRP input formulation and LeNet-5 classifier. The methodological choices of the study, including AWRP, automated feature extraction, and LeNet-5 classification, contributed to the framework's effectiveness in emotion recognition tasks on both DEAP and SEED datasets, offering a new perspective for future research.

V. CONCLUSION

In conclusion, this study presented a novel methodology employing AWRP and CNN for EEG-based emotion

recognition, demonstrating superior performance on the DEAP and SEED datasets. The proposed approach shows significant resilience against noise and enhanced capability to capture complex EEG signal details, thereby facilitating accurate and reliable emotion classification. However, there is a discernible trade-off between computational efficiency and performance, where increased window sizes in AWRP, while improving classification accuracy, also increase the computational cost in terms of encoding running time. To address this, future work could explore the parallelization of AWRP encoding computations. Given that the calculations for multiple windows or segments in the AWRP are independent, they are conducive to parallel processing. Implementing such parallel computations could potentially mitigate running time costs and enhance the scalability and efficiency of the method, particularly in scenarios involving larger datasets or more sophisticated emotion recognition tasks. Through continuous refinement and optimization, the proposed methodology holds promise for advancing the field of EEG-based emotion recognition, contributing valuable insights and tools for real-world applications and research.

ACKNOWLEDGMENT

This article has benefited from the use of artificial intelligence tools for grammar enhancement and editing purposes. Specifically, OpenAI's ChatGPT-4 was utilized to refine grammar and improve the clarity of the text. The application of this AI tool was strictly limited to language editing to ensure the readability of the manuscript and did not influence the research outcomes or development of ideas.

REFERENCES

[1] B. D. Fulcher and N. S. Jones, "Highly comparative feature-based time-series classification," *IEEE Trans. Knowl. Data Eng.*, vol. 26, no. 12, pp. 3026–3037, Dec. 2014.

[2] W. K. Wang, I. Chen, L. Hershkovich, J. Yang, A. Shetty, G. Singh, Y. Jiang, A. Kotla, J. Z. Shang, R. Yerrabelli, A. R. Roghanizad, M. M. H. Shandhi, and J. Dunn, "A systematic review of time series classification techniques used in biomedical applications," *Sensors*, vol. 22, no. 20, p. 8016, Oct. 2022.

[3] X. Hu, J. Chen, F. Wang, and D. Zhang, "Ten challenges for EEG-based affective computing," *Brain Sci. Adv.*, vol. 5, no. 1, pp. 1–20, Mar. 2019.

[4] D. W. Prabowo, H. A. Nugroho, N. A. Setiawan, and J. Debayle, "A systematic literature review of emotion recognition using EEG signals," *Cogn. Syst. Res.*, vol. 82, Dec. 2023, Art. no. 101152.

[5] U. Sharma, K. N. Faisal, R. R. Sharma, and K. V. Arya, "Facial landmark-based human emotion recognition technique for oriented viewpoints in the presence of facial attributes," *Social Netw. Comput. Sci.*, vol. 4, no. 3, Mar. 2023, Art. no. 273.

[6] A. Oker, N. Glas, F. Pecune, and C. Pelachaud, "An embodied virtual agent platform for emotional stroop effect experiments: A proof of concept," *Biologically Inspired Cogn. Archit.*, vol. 24, pp. 107–114, Apr. 2018.

[7] D. Glowinski, A. Camurri, G. Volpe, N. Dael, and K. Scherer, "Technique for automatic emotion recognition by body gesture analysis," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops*, Jun. 2008, pp. 1–6.

[8] S. Piana, A. Staglianò, F. Odone, and A. Camurri, "Adaptive body gesture representation for automatic emotion recognition," *ACM Trans. Interact. Intell. Syst.*, vol. 6, no. 1, pp. 1–31, Mar. 2016.

[9] S. Mariooryad and C. Busso, "Compensating for speaker or lexical variabilities in speech for emotion recognition," *Speech Commun.*, vol. 57, pp. 1–12, Feb. 2014.

[10] Y. Liu and G. Fu, "Emotion recognition by deeply learned multi-channel textual and EEG features," *Future Gener. Comput. Syst.*, vol. 119, pp. 1–6, Jun. 2021.

[11] V. V. Yerigeri and L. K. Ragha, "Speech stress recognition using semi-eager learning," *Cogn. Syst. Res.*, vol. 65, pp. 79–97, Jan. 2021.

[12] T. Fan, S. Qiu, Z. Wang, H. Zhao, J. Jiang, Y. Wang, J. Xu, T. Sun, and N. Jiang, "A new deep convolutional neural network incorporating attentional mechanisms for ECG emotion recognition," *Comput. Biol. Med.*, vol. 159, Jun. 2023, Art. no. 106938.

[13] M. A. Hasnul, N. A. A. Aziz, and A. A. Aziz, "Augmenting ECG data with multiple filters for a better emotion recognition system," *Arabian J. Sci. Eng.*, vol. 48, no. 8, pp. 10313–10334, Jan. 2023.

[14] J. A. Domínguez-Jiménez, K. C. Campo-Landines, J. C. Martínez-Santos, E. J. Delahoz, and S. H. Contreras-Ortiz, "A machine learning model for emotion recognition from physiological signals," *Biomed. Signal Process. Control*, vol. 55, Jan. 2020, Art. no. 101646.

[15] V. K. Patil, V. R. Pawar, S. Randive, R. R. Bankar, D. Yende, and A. K. Patil, "From face detection to emotion recognition on the framework of Raspberry pi and galvanic skin response sensor for visual and physiological biosignals," *J. Electr. Syst. Inf. Technol.*, vol. 10, no. 1, Apr. 2023, Art. no. 24.

[16] T. Paus, P. K. Sipila, and A. P. Strafella, "Synchronization of neuronal activity in the human primary motor cortex by transcranial magnetic stimulation: An EEG study," *J. Neurophysiol.*, vol. 86, no. 4, pp. 1983–1990, Oct. 2001.

[17] A. Iyer, S. S. Das, R. Teotia, S. Maheshwari, and R. R. Sharma, "CNN and LSTM based ensemble learning for human emotion recognition using EEG recordings," *Multimedia Tools Appl.*, vol. 82, no. 4, pp. 4883–4896, Apr. 2022.

[18] Z. Li, G. Zhang, L. Wang, J. Wei, and J. Dang, "Emotion recognition using spatial–temporal EEG features through convolutional graph attention network," *J. Neural Eng.*, vol. 20, no. 1, Feb. 2023, Art. no. 016046.

[19] Y. Jiang, S. Xie, X. Xie, Y. Cui, and H. Tang, "Emotion recognition via multiscale feature fusion network and attention mechanism," *IEEE Sensors J.*, vol. 23, no. 10, pp. 10790–10800, May 2023.

[20] Y. Song, Q. Zheng, B. Liu, and X. Gao, "EEG conformer: Convolutional transformer for EEG decoding and visualization," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 710–719, 2023.

[21] L. Gong, M. Li, T. Zhang, and W. Chen, "EEG emotion recognition using attention-based convolutional transformer neural network," *Biomed. Signal Process. Control*, vol. 84, Jul. 2023, Art. no. 104835.

[22] W. Zhang, F. Wang, Y. Jiang, Z. Xu, S. Wu, and Y. Zhang, "Cross-subject EEG-based emotion recognition with deep domain confusion," in *Intelligent Robotics and Applications* (Lecture Notes in Computer Science: Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 11740. Cham, Switzerland: Springer, 2019, pp. 558–570.

[23] A. Topic and M. Russo, "Emotion recognition based on EEG feature maps through deep learning network," *Eng. Sci. Technol., Int. J.*, vol. 24, no. 6, pp. 1442–1454, Dec. 2021.

[24] L. Farokhah, R. Sarno, and C. Fatichah, "Cross-subject channel selection using modified relief and simplified CNN-based deep learning for EEG-based emotion recognition," *IEEE Access*, vol. 11, pp. 110136–110150, 2023.

[25] H. Liu, Y. Zhang, Y. Li, and X. Kong, "Review on emotion recognition based on electroencephalography," *Frontiers Comput. Neurosci.*, vol. 15, Oct. 2021, Art. no. 84.

[26] X. Wang, Y. Ren, Z. Luo, W. He, J. Hong, and Y. Huang, "Deep learning-based EEG emotion recognition: Current trends and future perspectives," *Frontiers Psychol.*, vol. 14, Feb. 2023, Art. no. 1126994.

[27] J.-P. Eckmann, S. O. Kamphorst, and D. Ruelle, "Recurrence plots of dynamical systems," *Europhys. Lett.*, vol. 4, no. 9, pp. 973–977, Nov. 1987.

[28] N. Marwan and J. Kurths, "Nonlinear analysis of bivariate data with cross recurrence plots," *Phys. Lett. A*, vol. 302, nos. 5–6, pp. 299–307, Sep. 2002.

[29] N. Marwan, M. Carmenromano, M. Thiel, and J. Kurths, "Recurrence plots for the analysis of complex systems," *Phys. Rep.*, vol. 438, nos. 5–6, pp. 237–329, Jan. 2007.

[30] W. Huang, G. Yan, W. Chang, Y. Zhang, and Y. Yuan, "EEG-based classification combining Bayesian convolutional neural networks with recurrence plot for motor movement/imagery," *Pattern Recognit.*, vol. 144, Dec. 2023, Art. no. 109838.

[31] Q. He, F. Yu, J. Chang, and C. Ouyang, "Fuzzy granular recurrence plot and quantification analysis: A novel method for classification," *Pattern Recognit.*, vol. 139, Jul. 2023, Art. no. 109456.
[32] Y. Jeon and S. J. Kang, "Multi-slice nested recurrence plot (MsNRP): A robust approach for person identification using daily ECG or PPG signals," *Eng. Appl. Artif. Intell.*, vol. 126, Nov. 2023, Art. no. 106799.
[33] Y. Zhang, Y. Hou, K. OuYang, and S. Zhou, "Multi-scale signed recurrence plot based time series classification using inception architectural networks," *Pattern Recognit.*, vol. 123, Mar. 2022, Art. no. 108385.
[34] S. Koelstra, C. Mühl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "DEAP: A database for emotion analysis ;Using physiological signals," *IEEE Trans. Affect. Comput.*, vol. 3, no. 1, pp. 18–31, Jan. 2012.
[35] W.-L. Zheng and B.-L. Lu, "Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks," *IEEE Trans. Auto. Mental Develop.*, vol. 7, no. 3, pp. 162–175, Sep. 2015.
[36] I. M. A. Wirawan, R. Wardoyo, D. Lelono, and S. Kusrohmaniah, "Continuous capsule network method for improving electroencephalogram-based emotion recognition," *Emerg. Sci. J.*, vol. 7, no. 1, pp. 116–134, Nov. 2022.
[37] S. Liu, Y. Zhao, Y. An, J. Zhao, S.-H. Wang, and J. Yan, "GLFANet: A global to local feature aggregation network for EEG emotion recognition," *Biomed. Signal Process. Control*, vol. 85, Aug. 2023, Art. no. 104799.
[38] J. Shen, Y. Zhang, H. Liang, Z. Zhao, K. Zhu, K. Qian, Q. Dong, X. Zhang, and B. Hu, "Depression recognition from EEG signals using an adaptive channel fusion method via improved focal loss," *IEEE J. Biomed. Health Informat.*, vol. 27, no. 7, pp. 3234–3245, Jul. 2023.
[39] J. Shen, Y. Zhang, H. Liang, Z. Zhao, Q. Dong, K. Qian, X. Zhang, and B. Hu, "Exploring the intrinsic features of EEG signals via empirical mode decomposition for depression recognition," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 356–365, 2023.
[40] J. Chen, H. Li, L. Ma, and F. Soong, "DEEMD-SPP: A novel framework for emotion recognition based on EEG signals," *Frontiers Psychiatry*, vol. 13, Apr. 2022, Art. no. 885120.
[41] M. A. Asghar, M. J. Khan, F. Khan, Y. Amin, M. Rizwan, M. Rahman, S. Badnava, and S. S. Mirjavadi, "EEG-based multi-modal emotion recognition using bag of deep features: An optimal feature selection approach," *Sensors*, vol. 19, no. 23, p. 5218, Nov. 2019.
[42] J. Debayle, N. Hatami, and Y. Gavet, "Classification of time-series images using deep convolutional neural networks," *Proc. SPIE*, vol. 10696, p. 23, Apr. 2018.
[43] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.
[44] K. Swersky, J. Snoek, and R. P. Adams, "Multi-task Bayesian optimization," in *Advances in Neural Information Processing Systems*, vol. 26, C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Weinberger, Eds. Red Hook, NY, USA: Curran Associates, 2013.

DWI WAHYU PRABOWO (Graduate Student Member, IEEE) received the B.S. degree from the Department of Mathematics, University of Indonesia, Indonesia, in 2011, and the M. Eng. degree in biomedical engineering from Universitas Gadjah Mada, Indonesia, in 2014, where he is currently pursuing the Ph.D. degree with the Department of Electrical and Information Engineering. He is also a Lecturer and a Researcher with the Department of Information Systems, Universitas Darwan Ali, Indonesia. His current research interests include biomedical signals and image processing and analysis.

NOOR AKHMAD SETIAWAN (Member, IEEE) received the bachelor's and master's degrees in electrical engineering from Universitas Gadjah Mada, in 1998 and 2003, respectively, and the Ph.D. degree in electrical and electronics engineering from Universiti Teknologi PETRONAS, in 2009. He is currently an Associate Professor with the Department of Electrical and Information Engineering, Universitas Gadjah Mada. His research interests include soft computing and machine learning, with applications in electrical and biomedical engineering.

JOHAN DEBAYLE (Senior Member, IEEE) received the M.Sc. degree in images from Université Jean Monnet, Saint-Étienne, France, in 2002, and the Ph.D. degree in image, vision and signal from École Nationale Supérieure des Mines, Saint-Étienne, France, in 2005. In 2006, he joined the French National Institute for Research in Computer Science and Control (INRIA) as a Postdoctoral Fellow in the field of biomedical image analysis. He is currently a Full Professor with École Nationale Supérieure des Mines de Saint-Étienne (ENSM-SE), France, within the SPIN Center and the LGF Laboratory, UMR CNRS 5307. His research interests include image processing and analysis, pattern recognition, and stochastic geometry.

HANUNG ADI NUGROHO (Senior Member, IEEE) received the bachelor's degree in electrical engineering from Universitas Gadjah Mada, Indonesia, in 2001, the M.Eng. degree in biomedical engineering from The University of Queensland, Australia, in 2005, and the Ph.D. degree in electrical engineering from Universiti Teknologi PETRONAS, Malaysia, in 2012. He is currently a Professor with the Department of Electrical and Information Engineering, Faculty of Engineering, Universitas Gadjah Mada. His current research interests include biomedical signal and image processing and analysis, computer vision, medical instrumentation, and pattern recognition.