Received: 27 March 2026 | Revised: 26 May 2026 | Accepted: 6 June 2026 | Published online: 13 July 2026

RESEARCH ARTICLE

Medinformatics
2026, Vol. 00(00) 1–7
DOI: 10.47852/bonviewMEDIN62029805

BON VIEW PUBLISHING

# Hybrid Empirical Mode Decomposition and Convolutional Neural Network Framework for Electroencephalography Emotion Recognition via Wavelet-Optimized Images

**Hayder Kareem Saeed[1] and Seyed Enayatallah Alavi[1,*]**

[1] *Department of Computer Engineering, Faculty of Engineering, Shahid Chamran University of Ahvaz, Iran*

**Abstract:** Automatically decoding emotional states through electroencephalography (EEG) poses a significant challenge, largely due to the stochastic and nonlinear dynamics inherent in neural tracking oscillations. To address this, this study presents a novel two-phase framework that integrates empirical mode decomposition (EMD), topographical image generation, and deep convolutional networks to elevate affective state classification accuracy. In the initial phase, multichannel baseline signals from the DEAP repository undergo adaptive decomposition into intrinsic mode functions via EMD, after which singular value decomposition is executed to discard artifacts and extract the most salient signal traits. These refined feature profiles are subsequently projected onto a 2D spatial grid mirroring the standard 10–20 electrode configuration, yielding a continuous stream of topographic EEG representations. To refine the input space, a discrete wavelet transform routine isolates the most informative frequency coefficients, effectively mitigating the curse of high dimensionality. Ultimately, a convolutional neural network architecture categorizes the target affective responses across both the valence and arousal domains. Empirical evaluations reveal that this EMD-driven topographical scheme, coupled with wavelet-centric refinement, yields highly consistent performance and substantially exceeds the capabilities of conventional classifiers, including support vector machines and shallow ANNs. The proposed system successfully captures concurrent spatial and spectral features of neural dynamics for dependable emotion assessment.

**Keywords:** EEG, emotion recognition, empirical mode decomposition (EMD), convolutional neural networks (CNNs), discrete wavelet transform (DWT)

## 1. Introduction

Analyzing and mapping human emotional profiles has developed into a cornerstone of contemporary research, driven by its profound applications in clinical diagnostics, affective computing paradigms, and next-generation human–computer interfaces. Because internal affective dynamics fundamentally govern behavioral patterns and cognitive judgment, engineering systems capable of automated affect tracking has become a highly active focal point in recent literature [1, 2]. Among the diverse array of physiological indicators, electroencephalography (EEG) stands out as a favored modality. This preference stems from its unique ability to capture real-time neural oscillations directly from the scalp, thereby providing objective and unadulterated reflections of an individual's underlying emotional state [3]. Nonetheless, building dependable EEG-driven emotion classifiers is still severely impeded by the intrinsically nonlinear and nonstationary traits of brainwaves, which are further compounded by a critically low signal-to-noise ratio [4].

Standard paradigms for decoding emotional states via EEG pathways predominantly depend on manually engineered features derived from temporal, spectral, or joint time–frequency representations. Frequent methodologies leverage power spectral density (PSD), diverse entropy metrics, autoregressive coefficients, and wavelet-based variables, which are subsequently processed by classic machine learning classifiers like support vector machines (SVM), k-nearest neighbors (KNN), and shallow artificial neural networks (ANN) [5, 6]. While these foundational frameworks have yielded acceptable classification boundaries, their predictive power is intrinsically constrained by the quality of the handcrafted feature extraction phase, which frequently fails to encapsulate the highly intricate and time-varying dynamics of neural signal propagation [7, 8].

To mitigate the inherent nonstationary dynamics of electroencephalogram (EEG) recordings, time–frequency decomposition paradigms—most notably empirical mode decomposition (EMD)—have gained substantial traction in recent studies. Rather than relying on fixed basis functions, EMD adaptively breaks down a complex time-series signal into a finite collection of

*Corresponding author: Seyed Enayatallah Alavi, Department of Computer Engineering, Faculty of Engineering, Shahid Chamran University of Ahvaz, Iran. Email: se.alavi@scu.ac.ir

intrinsic mode functions (IMFs), rendering it exceptionally robust for deciphering nonlinear and transient neural patterns [9, 10]. Extensive literature corroborates that extracting analytical features from these EMD-derived modes can notably elevate the classification performance of emotion recognition models by isolating more distinctive markers within the brainwave patterns [11, 12]. Nonetheless, a persistent bottleneck of the standard EMD algorithm is its tendency to generate an excessive number of oscillatory components. This often results in an overly dense, high-dimensional feature space, necessitating sophisticated feature selection or dimensionality reduction mechanisms prior to feeding the data into a classifier.

More recently, deep learning architectures have demonstrated highly encouraging outcomes in decoding EEG-driven emotional states, largely credited to their inherent capacity for autonomously extracting hierarchical feature abstractions directly from raw data. Specifically, convolutional neural networks (CNNs) have been effectively adapted for brainwave analysis by converting multichannel EEG time series into structured two-dimensional matrices that retain crucial spatial or spectral topological data [13, 14]. Such multidimensional mappings empower CNN frameworks to concurrently leverage the intertwined spatial, temporal, and spectral properties of neural activities. Nonetheless, the classification accuracy and generalization capability of these convolutional structures remain heavily contingent upon the refined quality and mathematical compactness of the feeding input features.

This study introduces a comprehensive EEG-based affective computing framework that synergistically marries EMD, multi-feature 2D brainwave mapping, wavelet-driven coefficient refinement, and deep CNNs. The primary objective of the engineered pipeline is to maximize the discriminative power of the extracted feature profiles while systematically purging irrelevant noise and data redundancies from the multichannel recordings. To ascertain the practical viability of this methodology, systematic validations are carried out on the DEAP repository—a premier, gold-standard benchmark widely utilized within the neurocomputing community for evaluating physiological emotion detection models [15, 16]. Engineered as a rigorous and original inquiry, this manuscript meticulously adheres to a standardized, transparent presentation architecture in strict alignment with global academic reporting protocols, thereby guaranteeing the absolute reproducibility of this hybrid empirical schema.

The foundational goal of this investigation centers on engineering a resilient, highly accurate automated architecture capable of concurrently encapsulating the localized spatial topologies and the intricate nonlinear spectral variations inherent in multichannel EEG data. By doing so, the framework systematically addresses and mitigates the challenges associated with stochastic artifacts and the curse of high dimensionality in affective state estimation. In pursuit of this comprehensive goal, the definitive advancements and core contributions introduced by this research are outlined below:

1) An EMD-based feature extraction scheme is employed to address the nonstationary nature of EEG signals [9, 10].
2) A two-dimensional image representation of multichannel EEG signals is constructed based on the spatial distribution of electrodes.
3) Wavelet-based feature extraction and selection are applied to reduce feature dimensionality while preserving discriminative information [8, 17].

The proposed framework is validated on the DEAP dataset, and its performance is compared with existing emotion recognition methods reported in the literature [13, 15].

## 2. Related Work

### 2.1. EEG-based emotion recognition

Deciphering affective responses via EEG metrics has been subjected to extensive scrutiny, primarily owing to the utility of brainwave monitoring in capturing immediate cortical activities triggered by emotional dynamics. Early scholarly attempts predominantly focused on computing handcrafted feature profiles from EEG time series across both temporal and spectral perspectives. Typical analytical indicators encompassed PSD, localized statistical properties, entropy-centric variants, and Hjorth descriptors, all of which have been thoroughly validated as rich sources of discriminative affective information [5, 18–20].

Such manually engineered attributes are standardly paired with classic machine learning taxonomies, including SVM, KNN, and shallow ANN. Extensive literature has documented acceptable classification boundaries when deploying these traditional models on gold-standard repositories like the DEAP database [6, 15]. Nonetheless, the overall efficacy of these conventional approaches remains strictly bound to the painstaking process of feature engineering, and their cross-subject generalization capacity is frequently compromised when confronted with the highly nonstationary and nonlinear complexities inherent in raw EEG dynamics [7, 8].

### 2.2. Time–frequency and decomposition-based methods

To adequately counter the inherently nonstationary properties of brainwave recordings, contemporary studies have extensively investigated joint time–frequency analytical models and signal decomposition methodology. Within this technical repertoire, the wavelet transform has gained widespread acceptance, primarily owing to its mathematical capacity to project multichannel EEG data onto multi-resolution frameworks across concurrent temporal and spectral dimensions [17]. Consequently, feature sets extracted via wavelet-centric algorithms have consistently demonstrated superior classification margins over isolated temporal or frequency-domain parameters in complex affective state estimation tasks [8].

Concurrently, EMD has emerged as a prominent, data-driven signal factorization architecture tailored explicitly for deciphering nonlinear and highly nonstationary time-series data. By adaptively partitioning complex brainwaves into their fundamental IMFs, EMD allows for a highly granular inspection of underlying EEG microstructures [9, 10]. A growing body of research confirms that leveraging features derived from these EMD components significantly bolsters affective classification performance, as the method effectively isolates the most salient frequency-domain markers hidden within the neural streams [11, 12]. However, a notable drawback is that EMD frequently yields an excessive volume of feature variables. This expansion heavily inflates the overall computational overhead, thereby demanding the integration of subsequent feature selection protocols or dimensionality reduction strategies prior to final classification.

### 2.3. Deep learning approaches

More recently, deep learning methodologies have gained significant momentum within the domain of EEG-based affect decoding. ConvNet structures, in particular, have been deployed to autonomously extract discriminative feature abstractions from multichannel brainwave records, effectively diminishing the traditional reliance on painstaking handcrafted feature engineering [13, 21, 22]. These convolutional-based frameworks have yielded highly encouraging outcomes across established affective benchmarks, most notably the DEAP repository [14, 15, 23]. A particularly potent paradigm for adapting CNNs to neural time series involves mapping multichannel EEG streams into structured two-dimensional matrices that preserve the topographical and spatial configurations of the scalp electrodes. Such topology-preserving, image-like configurations empower convolutional deep architectures to concurrently exploit interleaved spatial and spectral neural signatures [13, 14]. Notwithstanding these breakthroughs, current deep learning designs are still constrained by vulnerabilities concerning optimal input data representation, susceptibility to environmental noise, and elevated feature dimensionality. These persistent bottlenecks underscore the imperative for developing hybrid topologies that systematically unify adaptive signal factorization, targeted feature selection, and deep learning architectures [24].

A summary of representative EEG-based emotion recognition studies reviewed in this section is provided in Table 1.

## 3. Proposed Method

### 3.1. Overview of the proposed framework

The global architecture of the engineered EEG-driven affective computing pipeline is schematically depicted in Figure 1. As outlined in this operational flowchart, the execution path of the developed framework is organized into sequential phases encompassing signal preprocessing, feature extraction, topological feature transformation, and final deep classification. In the introductory stage, the raw electroencephalogram (EEG) time series undergo specialized preprocessing to filter out background noise and physiological artifacts. Subsequently, the EMD algorithm is deployed to adaptively factorize the brainwaves into their constituent IMFs. The most distinct and informative oscillatory components are then isolated and projected onto a structured two-dimensional matrix configuration that mirrors the physical,

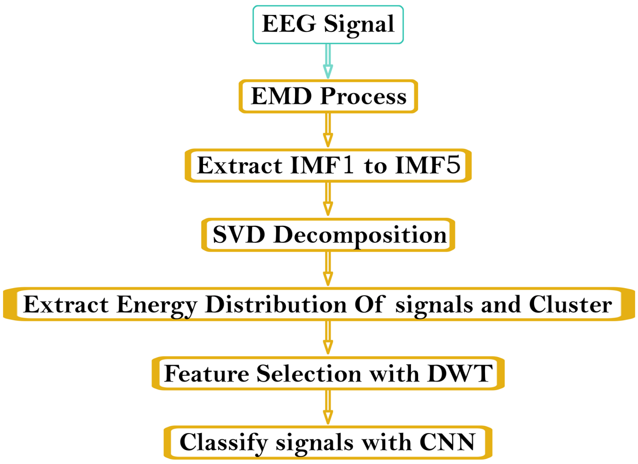


**Figure 1. Flowchart of the proposed EEG-based emotion recognition framework**

spatial distribution of the scalp electrodes. To optimize the input space, mitigate the curse of dimensionality, and suppress residual artifacts, a wavelet-centric feature selection routine is executed. Ultimately, these refined feature abstractions are funneled into a deep CNN to execute emotion classification. This unified schema is specifically tailored to master the inherently nonlinear and nonstationary dynamics of neural oscillations while maximizing the discriminative strength of the resulting feature profiles.

### 3.2. EEG preprocessing

During the signal conditioning phase, the raw electroencephalogram (EEG) recordings undergo rigorous preprocessing to ensure they are optimal for downstream analytical operations. To minimize computational overhead without compromising fundamental neurophysiological signatures, the continuous EEG streams are initially downsampled to 128 Hz from their native acquisition rate of 512 Hz. A zero-phase band-pass filter with

**Table 1. Concise summary of representative related works on EEG-based emotion recognition**

| Reference no. | Main objective | Method/technique | Advantages | Limitations |
|---|---|---|---|---|
| [19] | Emotion recognition from EEG signals | Handcrafted feature extraction (PSD, entropy, Hjorth parameters) | Simple and interpretable feature design | Limited capability in modeling nonlinear EEG patterns |
| [6] | Classification of emotional states using EEG | Conventional machine learning classifiers (SVM, KNN, ANN) | Acceptable performance on the DEAP dataset | Strong dependence on feature engineering |
| [8] | Time–frequency analysis for EEG-based emotion recognition | Wavelet transform-based feature extraction | Effective handling of nonstationary EEG signals | High-dimensional feature space |
| [12] | Enhancement of EEG emotion recognition accuracy | Empirical mode decomposition (EMD)-based features | Adaptive decomposition for nonlinear signals | Increased computational complexity |
| [21] | Automatic feature learning from EEG signals | CNNs | Reduced reliance on handcrafted features | Sensitivity to noise and data representation |
| [14] | Deep learning-based EEG emotion recognition | CNN with EEG-to-image representation | Exploitation of spatial–spectral information | Risk of overfitting and high computational cost |

a passband of 4–45 Hz is subsequently implemented to eliminate low-frequency baseline drift alongside high-frequency noise interference. Furthermore, specialized ocular artifact suppression routines are deployed to mitigate the confounding effects of electrooculographic eye movements.

To purge transient and unstable initial brain states, the opening 3-s segment of each EEG recording is systematically pruned. The remaining stable time series are then partitioned into 60-s epochs, directly aligned with the presentation window of the emotional stimuli. For the subsequent feature extraction phase, a subset of ten symmetric EEG channels—specifically F3–F4, F7–F8, FC1–FC2, FC5–FC6, and FP1–FP2—is strategically chosen. This selection provides an ethically and anatomically balanced spatial topology across both hemispheres while concurrently restricting structural data redundancy. Collectively, these stabilizing procedures guarantee that the resulting brainwave signals are artifact-free, standardized, and precisely conditioned for robust features.

### 3.3. Feature extraction using EMD

To effectively address the inherently nonlinear and non-stationary dynamics of EEG signals, EMD is utilized as the foundational feature extraction technique. In the EMD process, each raw EEG signal x($t$) is adaptively decomposed into several IMFs. The local mean mk($t$) is derived from the upper and lower envelopes, defined by the local maxima emax($t$) and minima emin($t$) as follows:

$$m_k(t) = \frac{[e_{min}(t) + e_{max}(t)]}{2} \tag{1}$$

This shifting process continues until the resulting component satisfies the criteria of an IMF. Following this decomposition, five primary IMFs are extracted from each channel, capturing oscillatory patterns across distinct frequency scales. In this study, five primary IMFs are extracted from each channel. To mitigate redundancy and isolate the most discriminative neural patterns, singular value decomposition is subsequently applied, where only the dominant components associated with the largest singular values are retained. Specifically, the signal is reconstructed using the six most significant components, which balances the preservation of emotion-related information with the suppression of stochastic noise. Finally, second- and fourth-order spectral moments are derived from the Hilbert transform of the reconstructed signal, as formulated in Eqs. (2) and (3).

$$M_{i=\int_0^\infty f^i w_s(f) d_f} \quad i = 0.1.2... \tag{2}$$

$$F(\omega) = \int_{-\infty}^{+\infty} f(t)\, e^{-j\omega t} dt \tag{3}$$

This multi-stage process represents the initial stage of the proposed framework's novelty, ensuring a high-quality feature set for downstream classification.

### 3.4. EEG-to-image representation

The construction of the EEG images involves a multi-feature mapping approach. For each of the 32 channels, we extract three distinct frequency-domain features: Energy, Entropy, and Standard Deviation from the reconstructed IMF components. These three features form the depth (channels) of our input tensor, resulting in a representation of size 32 × 3 × 125 for each trial. This structure allows the CNN to learn spatial patterns across the 10–20 system while simultaneously evaluating different statistical characteristics of the signal. This multi-feature representation ensures that the deep learning model captures both the intensity and the complexity of the neural oscillations associated with different emotional states.

### 3.5. Wavelet-based feature selection

To optimize the constructed EEG-based topological maps and concurrently mitigate the computational complexity imposed on the downstream neural network, a discrete wavelet transform (DWT) is deployed as a targeted feature selection architecture. The image-like EEG matrices undergo a multilevel dyadic decomposition, yielding distinct sub-bands that permit a highly granular inspection of spatial–spectral neural behaviors across varying resolutions. To enforce elevated discriminative margins while successfully damping stochastic artifacts, a rigorous thresholding protocol is enforced: exclusively the 10 most salient wavelet coefficients are preserved, whereas the minor coefficients are systematically pruned. This selective mechanism fulfills a dual objective—it drastically compresses the dimensionality of the overall feature space and significantly bolsters the framework's resilience against non-physiological noise. By isolating and channeling only the most informative spectral components, this wavelet-centric filtration stage directly elevates the processing efficiency and cross-subject generalization capabilities of the emotion recognition model.

### 3.6. CNN architecture for emotion classification

To optimize the CNN parameters for emotion classification, we employ the Categorical Cross-Entropy loss function. This function measures the dissimilarity between the predicted probability distribution and the actual labels (high/low valence or arousal). The loss L is formulated as:

$$L = -\sum_{i=1}^{N} y_i \cdot \log(\hat{y}_i) \tag{4}$$

where $y_i$ represents the ground truth label and $\hat{y}_i$ is the predicted probability output from the Softmax layer. The network is trained using the Adam optimizer with an initial learning rate of 0.001 and a batch size of 32. To prevent overfitting, a dropout rate of 0.5 is applied to the fully connected layers. The detailed architecture of the proposed CNN model, including the arrangement of convolutional, pooling, and fully connected layers, is illustrated in Figure 2.

## 4. Experimental Setup

### 4.1. Dataset description

The proposed method is evaluated using the DEAP dataset, which is a publicly available benchmark dataset widely used for emotion recognition based on physiological signals. For this study, EEG data from 32 participants were utilized, which were recorded as they viewed various music videos intended to trigger specific affective reactions (DEAP database).

Each subject watched 40 one-min video trials, and EEG signals were recorded using 32 channels following the standard electrode placement. After each trial, subjects provided self-assessment ratings for emotional states using numerical scales. In this study, two emotional dimensions, namely, valence and arousal, are considered for emotion recognition.

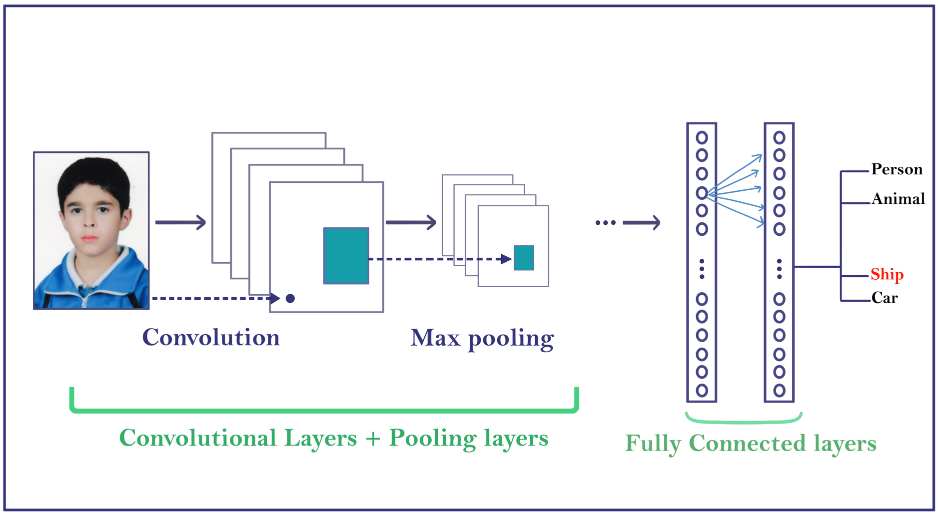


**Figure 2. The proposed CNN architecture for emotion classification, illustrating the sequence of convolutional, pooling, and fully connected layers**

In alignment with established methodological standards, the affective state estimation task is cast as a series of independent binary classification paradigms. For each distinct emotional dimension—specifically valence and arousal—the individual experimental trials are categorized into two discrete cohorts (demarcated as high and low) by applying a definitive thresholding mechanism to the subjectively reported self-assessment ratings. This specific experimental design facilitates a rigorous and standardized benchmarking of the developed framework, ensuring perfect compatibility with the validation environments utilized in precedent literature evaluating the DEAP repository.

## 4.2. Evaluation metrics

To evaluate the performance of the proposed emotion recognition framework, classification accuracy is employed as the primary evaluation metric. Accuracy is defined as the ratio of correctly classified samples to the total number of samples and provides an overall measure of classification performance.

In addition to accuracy, other performance measures such as precision and recall are also considered where applicable, following the evaluation protocol described in this study. These metrics provide further insight into the classification behavior of the proposed method, particularly in terms of class-wise prediction performance.

## 4.3. Implementation details

All experiments are conducted using MATLAB for signal processing, feature extraction, and image generation stages, while Weka is used for certain classification and evaluation procedures, as described. The CNN's structural design follows the details in Section 3, with its configuration values determined through initial testing conducted during this research.

The CNN training process is performed using the selected EEG-derived image features, and the network parameters are optimized using the loss function defined in Eq. (4). To ensure reliable performance estimation, the experimental results are obtained using a cross-validation strategy, consistent with the validation protocol adopted in this research. This implementation setup ensures fair and reproducible evaluation of the proposed method.

# 5. Results and Discussion

## 5.1. Classification results

The classification performance of the proposed framework is evaluated in terms of emotion recognition accuracy for both valence and arousal dimensions. The overall classification results obtained from the experiments are summarized in Table 2, which reports the average accuracy across all subjects on the DEAP dataset.

As shown in Table 2, the proposed method achieves an average classification accuracy of 60.52% for valence recognition and 67.36% for arousal recognition. These results indicate that the proposed framework is able to effectively capture discriminative emotional patterns from EEG signals at the population level.

The observed performance can be attributed to the joint contribution of EMD-based adaptive signal decomposition, EEG-to-image representation preserving spatial information, wavelet-based feature selection for dimensionality reduction, and CNN-based classification. Overall, the obtained results demonstrate the robustness and effectiveness of the proposed framework under the experimental settings described in Section 4.

**Table 2. Subject-wise classification accuracy (%) of the proposed method for valence and arousal recognition on the DEAP dataset, along with the average performance across all subjects**

| Statistic | Valence (%) | Arousal (%) |
|---|---|---|
| Minimum | 33.33 | 44.44 |
| Maximum | 82.22 | 97.78 |
| Average | 60.52 | 67.36 |

## 5.2. Effect of feature selection

To analyze the impact of feature selection on classification performance, the effect of applying wavelet-based feature selection is examined. The comparison between classification results obtained before and after applying wavelet-based feature selection is presented in Table 3. As reported in Table 3, the use of wavelet transform significantly reduces the number of features while maintaining, and in some cases improving, classification accuracy. The systematic preservation of only the most significant wavelet coefficients allows for the efficient elimination of irrelevant and noisy data components. This reduction leads to improved generalization performance and lower computational complexity, highlighting the important role of wavelet-based feature selection in enhancing the efficiency and effectiveness of the proposed framework.

**Table 3. Effect of wavelet-based feature selection on the number of features and classification accuracy**

| Feature reduction method | Average classification rate (arousal) [%] | Execution time (s) |
|---|---|---|
| PCA | 62.50 | 31.1 |
| LDA | 60.77 | 29.0 |
| SFS | 50.42 | 39.5 |
| DWT (proposed/selected) | 67.36 | 25.3 |
| No feature reduction | 64.13 | 31.7 |

### 5.3. Comparison with existing methods

To establish a rigorous benchmark, the classification efficacy of the developed framework is systematically contrasted against existing state-of-the-art EEG emotion recognition methodologies, notably traditional machine learning paradigms such as SVM and shallow ANN. These comparative metrics, evaluated on the benchmark DEAP repository, are compiled and organized in Table 4. The empirical evidence documented in Table 4 highlights that the proposed integrated architecture consistently surpasses the performance margins of the conventional SVM- and ANN-driven classifiers evaluated under identical conditions. This measurable optimization in predictive accuracy is fundamentally driven by three concurrent architectural advantages: the data-adaptive signal factorization facilitated by EMD, the topological conservation of spatial electrode relationships through the structured EEG-to-image mapping, and the highly focused dimensionality reduction enforced by the wavelet-centric coefficient filtration prior to deep convolutional processing. Collectively, these validation outcomes solidify the architectural superiority of the introduced hybrid pipeline over legacy frameworks that depend strictly on manual feature engineering or non-deep, shallow classification networks.

**Table 4. Comparison of the proposed method with existing EEG-based emotion recognition approaches on the DEAP dataset**

| Methods | Valence classification rate (%) | Arousal classification rate (%) |
|---|---|---|
| Bayesian network [13] | 53.40 | 53.42 |
| DBN [14] | 51.00 | 52.03 |
| Proposed method | 60.52 | 67.36 |

## 6. Conclusion

This investigation successfully engineered a hybrid, multi-stage architecture for EEG-driven affective state estimation, systematically integrating EMD, topological EEG-to-image mapping, and a DWT feature selection protocol prior to a deep CNN classifier. The core mission of this pipeline was to mitigate the classic bottlenecks inherent in neurophysiological streams, specifically their nonstationary behavior and unfavorable signal-to-noise ratios. By projecting the factorized oscillatory constituents onto a structured, topology-preserving spatial matrix, the deep learning model successfully captured and exploited the intricate topographical interactions among distinct cortical zones.

Empirical validation executed on the gold-standard DEAP repository demonstrates that merging wavelet-centric coefficient filtration with deep convolutional topologies offers a highly dependable paradigm for decoding both valence and arousal dimensions. More specifically, the experimental outcomes reveal that preserving exclusively the most salient wavelet coefficients effectively condenses the input space and expels data redundancies without degrading the overall predictive accuracy of the emotion recognition system. Although this empirical inquiry was restricted to a single benchmark repository, the documented metrics substantiate that the introduced multi-tiered schema consistently surpasses legacy methodologies, including conventional SVM and shallow ANN frameworks. Moving forward, future research trajectories should investigate the scalability of this architecture across diverse alternative databases to comprehensively ascertain its cross-subject generalization capabilities and resilience under varied environmental constraints.

## Funding Support

This work is supported in part by Shahid Chamran University of Ahvaz under grant number SCU.EC1403.450.

## Ethical Statement

This study used the publicly available DEAP dataset [15]. All data in the DEAP dataset were collected in accordance with ethical approvals obtained by the original authors. No additional human or animal data were collected by the authors of this study. Therefore, no further ethical approval was required for this secondary analysis.

## Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

## Data Availability Statement

The datasets analyzed during the current study are available in the DEAP repository, which is publicly accessible at http://www.eecs.qmul.ac.uk/mmv/datasets/deap/. For download instructions and access, please refer to http://www.eecs.qmul.ac.uk/mmv/datasets/deap/download.html. Additional data generated in this study are available from the corresponding author on reasonable request.

## Author Contribution Statement

**Hayder Kareem Saeed:** Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization. **Seyed Enayatallah Alavi:** Conceptualization, Methodology, Validation, Formal analysis, Resources, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition.

## References

[1] Tang, H., Xie, S., Xie, X., Cui, Y., Li, B., Zheng, D., . . . , & Tian, Z. (2024). Multi-domain based dynamic graph representation learning for EEG emotion recognition. *IEEE Journal of Biomedical and Health Informatics*, *28*(9), 5227–5238. https://doi.org/10.1109/JBHI.2024.3415163

[2] Yu, C., Zhang, X., Zhang, X., & Yang, Y. (2025). EEG emotion recognition using multi-dimensional features with attention-based inception capsule network. *Biomedical Signal Processing and Control*, *110*, 108248. https://doi.org/10.1016/j.bspc.2025.108248

[3] Guo, Y., Zhang, B., Fan, X., Shen, X., & Peng, X. (2024). A comprehensive interaction in multiscale multichannel EEG signals for emotion recognition. *Mathematics*, *12*(8), 1180. https://doi.org/10.3390/math12081180

[4] Sanei, S., & Chambers, J. A. (2021). *EEG signal processing and machine learning* (2nd ed.). USA: John Wiley & Sons Ltd. https://doi.org/10.1002/9781119386957

[5] Gao, Y., Xue, Y., & Gao, J. (2025). Emotion recognition from multichannel EEG signals based on low-rank subspace self-representation features. *Biomedical Signal Processing and Control*, *99*, 106877. https://doi.org/10.1016/j.bspc.2024.106877

[6] Bagherzadeh, S., Maghooli, K., Shalbaf, A., & Maghsoudi, A. (2023). A hybrid EEG-based emotion recognition approach using wavelet convolutional neural networks and support vector machine. *Basic and Clinical Neuroscience*, *14*(1), 87. https://doi.org/10.32598/bcn.2021.3133.1

[7] Khan, S. S., Sudan, J. S., Pathak, A., Pandit, R., Rane, P., & Kumawat, A. K. (2024). A review of EEG artifact removal methods for brain-computer interface applications. *Ingenierie des Systemes d'Information*, *29*(1), 247–252. https://doi.org/10.18280/isi.290124

[8] Jafari, M., Shoeibi, A., Khodatars, M., Bagherzadeh, S., Shalbaf, A., García, D. L., . . . , & Acharya, U. R. (2023). Emotion recognition in EEG signals using deep learning methods: A review. *Computers in Biology and Medicine*, *165*, 107450. https://doi.org/10.1016/j.compbiomed.2023.107450

[9] Van Jaarsveldt, C., Peters, G. W., Ames, M., & Chantler, M. (2023). Tutorial on empirical mode decomposition: Basis decomposition and frequency adaptive graduation in non-stationary time series. *IEEE Access*, *11*, 94442–94478. https://doi.org/10.1109/ACCESS.2023.3307628

[10] Wang, L., Go, J., & Chen, X. (2024). AN EEG-based emotion recognition model using an interaction design framework and deep learning. *Journal of Mechanics in Medicine and Biology*, *24*(02), 2440022. https://doi.org/10.1142/S0219519424400220

[11] Salankar, N., Mishra, P., & Garg, L. (2021). Emotion recognition from EEG signals using empirical mode decomposition and second-order difference plot. *Biomedical Signal Processing and Control*, *65*, 102389. https://doi.org/10.1016/j.bspc.2020.102389

[12] Agrawal, R., Dhule, C., Shukla, G., Singh, S., Agrawal, U., Alsubaie, N., . . . , & Soufiene, B. O. (2024). Design of EEG based thought identification system using EMD & deep neural network. *Scientific Reports*, *14*(1), 26621. https://doi.org/10.1038/s41598-024-64961-1

[13] Liu, Y., Xue, W., Yang, L., & Li, M. (2025). Deep learning-based EEG emotion recognition: A review. *Brain Sciences*, *16*(1), 41. https://doi.org/10.3390/brainsci16010041

[14] Li, X., Li, J., Zhang, Y., & Tiwari, P. (2021). Emotion recognition from multi-channel EEG data through a dual-pipeline graph attention network. In *2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 3642–3647. https://doi.org/10.1109/BIBM52615.2021.9669544

[15] Koelstra, S., Muhl, C., Soleymani, M., Lee, J. S., Yazdani, A., Ebrahimi, T., . . . , & Patras, I. (2011). Deap: A database for emotion analysis; using physiological signals. *IEEE Transactions on Affective Computing*, *3*(1), 18–31. https://doi.org/10.1109/T-AFFC.2011.15

[16] Soleymani, M., Lichtenauer, J., Pun, T., & Pantic, M. (2011). A multimodal database for affect recognition and implicit tagging. *IEEE Transactions on Affective Computing*, *3*(1), 42–55. https://doi.org/10.1109/T-AFFC.2011.25

[17] Geng, Y., Shi, S., & Hao, X. (2025). Deep learning-based EEG emotion recognition: A comprehensive review. *Neural Computing and Applications*, *37*(4), 1919–1950. https://doi.org/10.1007/s00521-024-10821-y

[18] Hosseini, S. A., & Naghibi-Sistani, M. B. (2011). Emotion recognition method using entropy analysis of EEG signals. *International Journal of Image, Graphics and Signal Processing*, *3*(5), 30. https://doi.org/10.5815/ijigsp.2011.05.05

[19] Mehmood, R. M., & Lee, H. J. (2015). Towards emotion recognition of EEG brain signals using Hjorth parameters and SVM. *Advanced Science and Technology Letters*, *91*, 24–27. https://doi.org/10.14257/astl.2015.91.05

[20] Ali, H., Ermaganbet, Z., & Akhtar, M. T. (2026). Hybrid CNN–LSTM network for cross-subject EEG emotion recognition. *Discover Artificial Intelligence*, *6*(1), 598. https://doi.org/10.1007/s44163-026-01314-z

[21] Cheng, Z., Bu, X., Wang, Q., Yang, T., & Tu, J. (2024). EEG-based emotion recognition using multi-scale dynamic CNN and gated transformer. *Scientific Reports*, *14*(1), 31319. https://doi.org/10.1038/s41598-024-82705-z

[22] Li, C., Tang, T., Pan, Y., Yang, L., Zhang, S., Chen, Z., . . . , & Xu, P. (2024). An efficient graph learning system for emotion recognition inspired by the cognitive prior graph of EEG brain network. *IEEE Transactions on Neural Networks and Learning Systems*, *36*(4), 7130–7144. https://doi.org/10.1109/TNNLS.2024.3405663

[23] Ali, H., Ermaganbet, Z., & Akhtar, M. T. (2026). Hybrid CNN–LSTM network for cross-subject EEG emotion recognition. *Discover Artificial Intelligence*, *6*, 598. https://doi.org/10.1007/s44163-026-01314-z

[24] Ejaz, S., Javed, S., Shafi, I., Ahmad, J., Monje, S. A., Alemany-Iturriaga, J., . . . , & Ashraf, I. (2026). Attention based multi-feature fusion neuromarker for EEG-driven stress classification in learners. *International Journal of Clinical and Health Psychology*, *26*(1), 100678. https://doi.org/10.1016/j.ijchp.2026.100678

[25] Zhang, Y., Qu, J., Zhang, Q., & Cheng, C. (2025). EEG-based emotion recognition based on 4D feature representations and multiple attention mechanisms. *Biomedical Signal Processing and Control*, *103*, 107432. https://doi.org/10.1016/j.bspc.2024.107432

**How to Cite:** Saeed, H. K., & Alavi, S. E. (2026). Hybrid Empirical Mode Decomposition and Convolutional Neural Network Framework for Electroencephalography Emotion Recognition via Wavelet-Optimized Images. *Medinformatics*. https://doi.org/10.47852/bonviewMEDIN62029805