Volume 19, Issue 2, 192529 (1-8)

Majlesi Journal of Electrical Engineering (MJEE)

https://doi.org/10.57647/j.mjee.2025.1902.29

Hybrid attention-based deep learning network for emotion recognition by ECG signal

Mahtab Vaezi[1], Mehdi Nasri[2,3,*], Farhad Azimifar[1], Mahdi Mosleh[4]

[1] *Department of Biomedical Engineering, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran.*
[2] *Department of Biomedical Engineering Khomeinishahr Branch, Islamic Azad University, Isfahan, Iran.*
[3] *Efficiency and Smartization of Energy Systems Research Center, Khomeinishahr Branch, Islamic Azad University, Isfahan, Iran.*
[4] *Department of Computer Engineering, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran.*

[*] **Corresponding author:** nasri_me@iaukhsh.ac.ir

Original Research

Received:
12 March 2025
Revised:
2 April 2025
Accepted:
26 April 2025
Published online:
1 June 2025

Abstract:

Emotions play an important role in our daily activities, decision-making, and artificial intelligence needs to identify emotions to interact constructively with its audience. In this paper, an intelligent method for two-dimensional emotion recognition is proposed. The ECG signal available in the DREAMER database has been used to recognize emotions because of the high correlation of this signal with emotions and easy recording. First step for valence and arousal recognition, the ECG signal is entered into the deep learning network, which is a combination of CNN and LSTM. CNN performs feature extraction and LSTM performs data classification. The attention mechanism aims to optimize the weights and improve the performance of the network, overseeing the proposed deep learning network. Using the proposed method, valence and emanation were identified with 95% and 94% accuracy, respectively. The proposed hybrid network is very suitable for high-dimensional data, and the use of the attention mechanism helps to improve the performance of the network by preventing overfit and getting stuck in local optimal.

Keywords: Emotion; CNN; LSTM; ECG; Attention mechanism

1. Introduction

Emotions are one of the critical factors in creating human desires, choices, and behaviors [1]. Therefore, recognizing emotions plays a significant role in improving people's mental quality and human-computer interaction [2]. Using smart devices, robots, and artificial intelligence without recognizing human emotions and feelings is practically impossible. In the future, intelligent systems will not work without emotional interaction with humans [3]. By creating systems that understand and analyze human emotions well, we can hope for efficient human-computer interaction [2].
Emotions are identified in two classical and dimensional ways [4]. In the classic system, a limited number of emotions are labeled, such as anger, anger, disgust, etc., and a person's emotions are placed in several different groups. Due to the complexities of emotions, the classical method cannot be helpful. In dimensional diagnosis, various dimensions of a person's emotions are examined. One of the most famous dimensional models is the two-dimensional model that Russel introduced in 1980 [5], which explores the valance and arousal state for emotions and examines the momentary feeling of a person according to the low or high level of each [6]. Fig. 1 shows the two-dimensional diagram of emotions. The complexity of emotional states is more challenged in dimensional models. It is believed that in the discrete model, due to the oversimplification of emotions, some emotional experiences are ignored, and some are lost [6, 7].
Valence concerning an emotional state expresses the degree of pleasantness of the feeling [8]. Arousal expresses the intensity of an emotional state. According to figure 2, in the two-dimensional model, four states express the person's feelings: high valance-high arousal (HVHA), high valance-low arousal (HVLA), low valance-high arousal (LVHA) and low valance-low arousal (LVLA). Suppose a person is in a situation that is high in terms of valance and high in terms of arousal. In that case, he is probably experiencing a positive and intense emotional state, such as excitement.

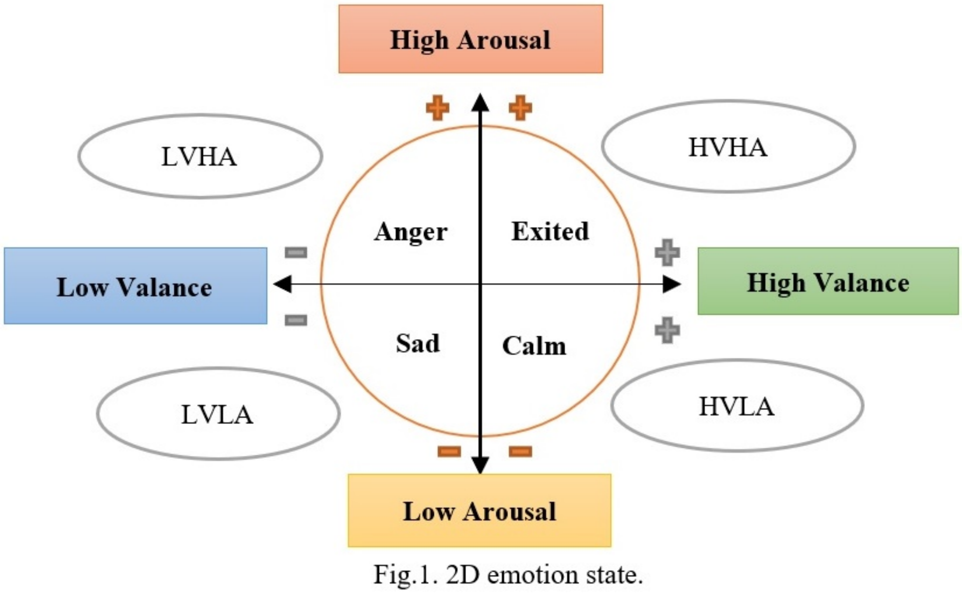

Figure 1. 2D emotion state.

Our primary goal in the proposed method is to identify these four emotional regions.
Artificial intelligence must identify people's emotions to interact efficiently with its audience. For this reason, much research has been presented in recent years to provide intelligent emotion-recognition systems [9–11]. Intelligent methods of identifying emotions are classified into two categories: non-contact methods (using photos (facial images) [12, 13], writing and text [14], audio, visual modalities [15], eye movement [16], and speech [17]) and contact methods (using physiological signals such as EEG [18], ECG [19], GSR [20]). Among the mentioned methods, physiological signals are the best option because a person can hide his feelings in facial images or sounds. Still, they cannot interfere with the physiology of their body [21]. EEG and ECG are the most popular for identifying emotions among the physiological signals. However, EEG is complex to record, and because of its valuable information, it is only helpful for clinical purposes [22]. The ECG signal offers less information than the EEG, but its easy recording can be beneficial for many purposes. There is also a special relationship between emotion and ECG signal [23]. Therefore, the ECG signal is also used in the proposed method.
In line with the automatic identification of emotions using physiological signals, the studies are divided into two categories: classical machine learning methods and deep learning methods. The first categories of signal processing and emotion recognition operations are usually performed using manual feature extraction and feature classification using classic classifiers [24]. Dujaili et al. [25] used SVM and KNN classifiers to identify emotions. The extracted features were classified using speech features such as Fourier, SVM, and KNN. Liu et al. [26] Identified people's feelings using the optimized combination of genetic algorithm and SVM classifier. Hasnul et al. [27] Identified emotional states using the ECG signal and the K-nearest neighbor classifier. Gao et al. [28] recognition of three emotional states, happy, neutral, and sad, using wavelets feature extraction and SVM classifier. By manually extracting the features, it is possible to lose important information, and selecting the appropriate features and classification is challenging. In this article, an attempt has been made to automatically perform the steps manually using the machine learning method.
Researchers have very welcomed the use of deep learning

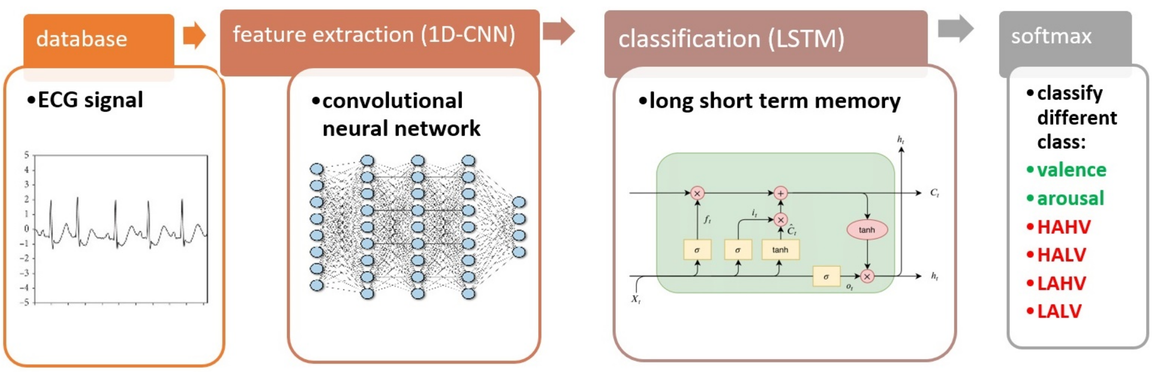

Figure 2. Flowchart of the proposed method.

networks in recent years. One of the advantages of these networks is the automatic performance of operations related to signal processing. Fan et al. [23] identified arousal and valance in emotions using ECG signals and deep learning networks. Their proposed method uses the attention mechanism in the convolutional deep learning network, improving work performance. Wang et al. [29] a combination of EEG and ECG signals was used to study emotions. They fed the raw signal into a deep LSTM network and discretely identified emotional states. Wang et al. [30] Emotions are continuously identified using a deep convolutional network and a combination of EEG signals and facial expressions. EEG signals are complex to record and cannot be used daily. In this article, an attempt has been made to use a signal related to emotions, which has a more straightforward structure than EEG. The classical method is highly dependent on parameter setting. It is also challenging to find the appropriate feature, and such information may need to be recovered during manual feature extraction. Deep learning methods have solved these two defects of the classical method. Deep learning methods automatically perform processing operations and provide the desired output. In addition, deep learning networks are resistant to noise [31].

Due to being sensitive to noise and high sensitivity to setting parameters and features, machine learning methods cannot be suitable for identifying emotions. On the other hand, deep learning networks such as CNN, which were recently used to identify emotions, do not have high accuracy in classification and cannot be a suitable option for everyday use. In this method, we are trying to combine machine learning and deep learning methods to provide a deep multitasking network using the capabilities of any deep learning network. In this paper, a deep learning hybrid network is used to detect the dimensions of emotions using an ECG signal. Since each deep learning network has its Advantages, we have tried to achieve the desired results in this study by using their combination. For this reason, a CNN-LSTM hybrid network is used. The convolutional neural network has a high ability to identify the feature map and learn the model. Due to its memory structure and previous information review, the LSTM network can highly classify data among deep learning networks. Using this ability of LSTM and CNN's high ability in preparing feature maps leads to improved classification. ECG information. To optimize the weights and improve the performance of the CNN-LSTM model, the attention mechanism is used in the proposed method, which updates the weights during all processing stages to reach the optimal weight. The proposed attention-based network is very suitable for high-dimensional and large-volume data. In high-dimensional data, it is essential to adjust the weight so the network can be trained well (not stuck in an optimal place or overfit).

Contributions in this paper include:

- Using a hybrid deep learning network that uses two separate deep networks for automatic feature map extraction and classification. LSTM has a high ability in classification, and CNN has a high ability to prepare feature maps. In this article, the capabilities of these two networks are used to design a hybrid deep-learning network.
- Optimizing the weights of the combined network using the attention mechanism to achieve the highest accuracy of the CNN-LSTM network.
- ECG signals are used to identify emotions with easy registration and high interaction with emotions.

The article's sections are as follows: in the second part, the data used, the CNN and LSTM network, and the attention mechanism are discussed. In the third part, we examine the results of network attention-based CNN-LSTM simulation, and in the last part, the discussion and conclusions are discussed.

2. Materials and methods

This article uses a hybrid deep learning method using CNN and LSTM networks based on the attention mechanism. The single-channel ECG signal is entered into the proposed network to identify the dimensions of emotions. Figure 2 shows the block diagram of the proposed method. 1D-CNN prepares the feature map, which goes to LSTM for classification. Finally, the desired classes are classified in the softmax layer. The proposed method uses the DREAMER database along with the CNN-LSTM network. Single-channel ECG is considered to be the input of the deep network. The outputs of the proposed network are examined in two stages. In the first stage, high and low valence and arousal values are calculated and discussed in 2 classes. Then, HAHV, HALV, LAHV, LALV modes are calculated and analyzed in a 4-class classification. First, the DREAMER database is examined.

2.1 Database

The DREAMER database was introduced in 2018 and includes ECG and EEG signals from 23 people (For each person, we have 18 signal recordings using different stimuli) [32]. Audio and visual stimuli stimulate people's feelings, and then each person's feelings are labeled by experts in terms of capacity and arousal. People's EEG and ECG signals are recorded by wearable and portable systems that allow recording in different situations of daily activities. This database considers the numbers 1-5 for each valance and arousal state. Numbers 4 and larger are considered "high," and numbers smaller than 4 are considered "low" [33]. Table 1 shows the additional information from the database.

2.2 Proposed network (attention-based CNN-LSTM)

The proposed network comprises CNN, LSTM, and an attention mechanism. We will examine each of these parts below.

1. Convolutional Neural Network

Convolutional neural networks (CNN) are a type of MLP artificial neural network with more hidden layers and neurons, which can learn complex patterns using their hidden layers [34]. Each CNN consists of three essential parts: the convolutional, pooling, and Fully connected layers. The

Table 1. Information on the DREAMER database.

Class	Full class name	Class distribution	Number of ECG sample
HAHV	High arousal-high valence	18%	76
HALV	High arousal-low valence	26%	105
LAHV	Low arousal-high valence	21%	87
LALV	Low arousal-low valence	35%	146

convolutional layer uses network vision by using the relationship of its neurons and recognizes the pattern between the input data [35]. The pooling layer is placed after each convolutional layer and performs the feature dimension reduction operation [34]. The Fully Connected layer is placed at the end of the network architecture and performs data classification operations.
In the proposed method, five conventional layers are used, and the filter dimensions of the layers are 8, 16, 32, 32, and 64, respectively. The ReLu activator function is added after each conventional layer to achieve a unified feature map. The 1D-CNN network's Full connection layer has been moved to the end of the network architecture. Before using the fully connected layer, the LSTM network was used, and we will examine this network in the following.

2. Long Short-Term Memory Network

An advanced, recurrent neural network (RNN) type is long short-term memory (LSTM), first introduced in 1997. LSTM networks have a strong memory, and according to their memory, they can consider before and after the time series and make the best choice for data classification [36]. These networks also have a forgetting gate and can remove redundant data over time to make the best decision.
Due to their high classification ability, these networks have been used for data classification in the proposed method. The LSTM classifier is placed before the softmax layer, and combining LSTM and softmax provides the best sentiment classification results.

3. Attention mechanism

LSTM network is used for high-dimensional data, gradient distortion occurs due to the large volume of data, leading to the network's inefficiency. For this reason, an attention mechanism has been introduced to optimize the weights and provide network training with the highest classification accuracy [37].
Fig. 3 shows the layout of the proposed deep-learning network. As shown in figure 3, the proposed network consists of 5 convolutional layers, one LSTM layer, one attention mechanism layer, and softmax. To get the best number of convolutional layers, several recent articles were reviewed in this regard [2, 23, 38], and then five layers were selected through different iterations. The size of the 5-layer convolutional filter three and the number of filters are 16, 32, 32, 64, and 64, respectively. Also, the LSTM used in the proposed method has 128 hidden units (neurons). Hyperparameter information of the proposed network is shown in Table 2.

Table 2. Deep learning model parameters.

	Parameters	Value
1	Number of convolutional layers	5
2	Number of LSTM layer	1
3	Maximum epoch	400
4	Mini Batch Size	128
5	Initial Learn Rate	0.95
6	Learn Rate Drop Factor	0.2
7	L2 Regularization	0.01

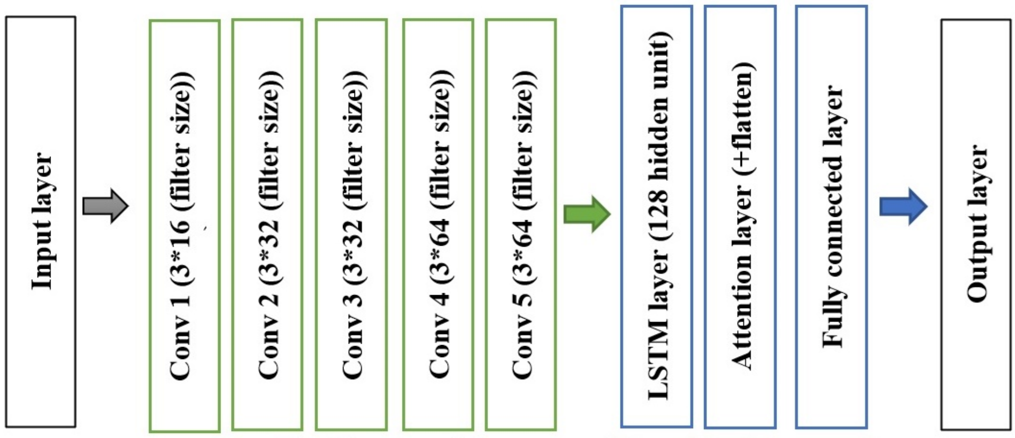

Figure 3. Layout of the proposed deep-learning network.

3. Results

In this article, A network is presented to identify emotions using a combination of two networks, 1D-CNN and LSTM; it performs ECG signal processing and provides the output of different emotional states regarding valance and arousal. The simulations of the proposed method have been done in MATLAB 2023b. The k-fold method is used in the proposed evaluation method, and all the values are presented after the review. The black curve in Fig. 4 shows the accuracy values after the evaluation.

Table 3 shows the accuracy, sensitivity, and specificity values for different classifications of emotions. In identifying continuous emotions, we face multi-class problems; first, high and low states for valence and arousal are identified, then in a four-class (HAHV, HALV, LAHV, LALV) classification.

Fig. 4 shows the proposed network's convergence and reduction of classification error for different classifications of Valance and Arousal.

Fig. 5 shows the confusion matrix resulting from the proposed method's simulation. According to figure 5, the proposed method can identify the low valence and low arousal classes.

Much research has been done on two-dimensional identification of emotions in recent years. Table 4 shows the results of some research compared to the proposed method. The ECG signal is one of the best options for examining emotions due to its high relationship with emotions and easier recording than EEG. However, due to the limited information on ECG, recent research providing different machine learning and deep learning methods could not provide high accuracy to the identification of emotions using this signal, and this issue makes it impossible to use these methods in the clinic in the future. As shown in Table 4, the results of the proposed method improve the overall performance of emotion recognition.

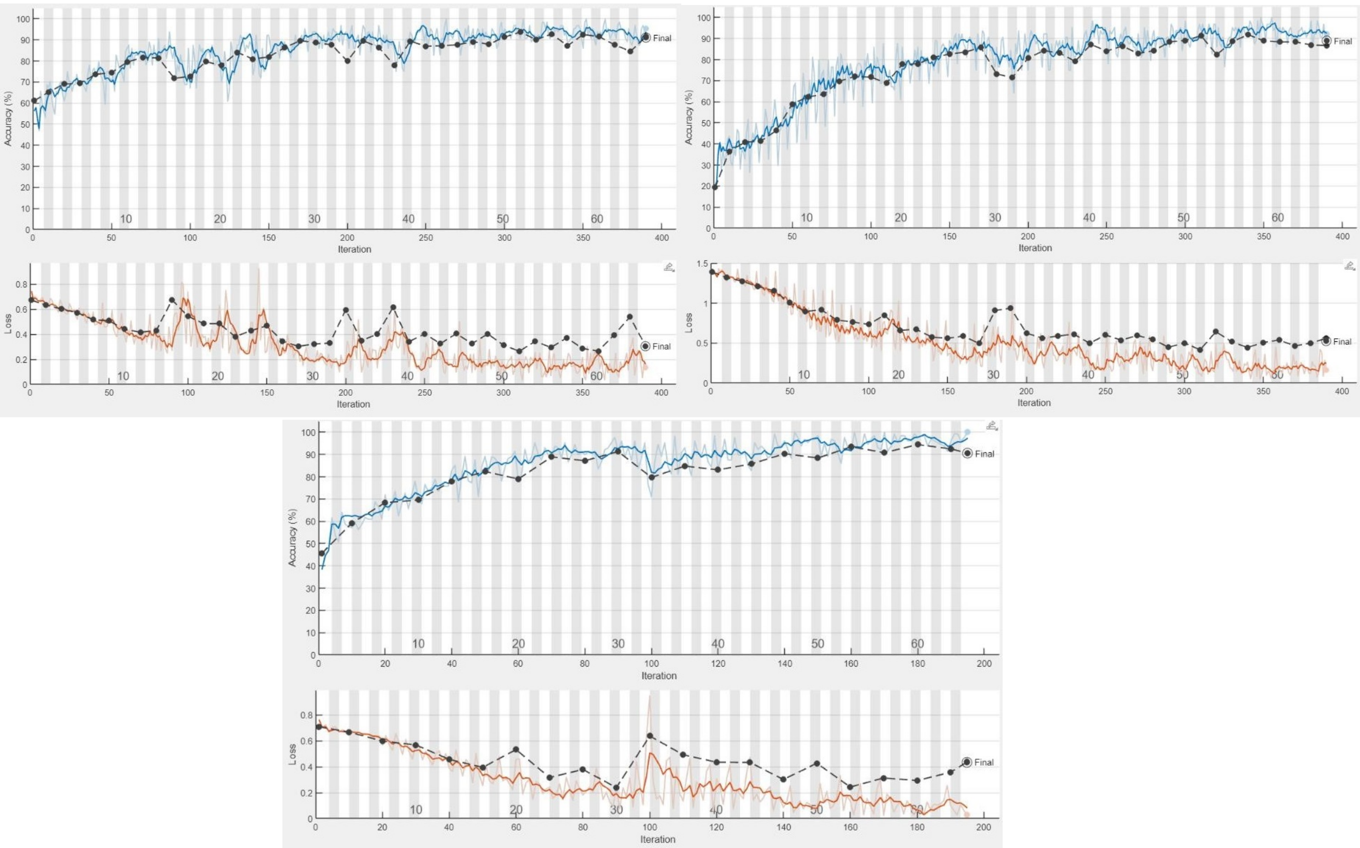

Figure 4. Convergence and reduction of classification error of the proposed network. (a): High-low valance, (b): High-low arousal, (c) 4 Class (HAHV, HALV, LAHV, LALV).

Table 3. Accuracy, sensitivity, and specificity for different classifications of emotions.

Emotion class	Sensitivity	Specificity	Accuracy
Low-high valence	0.89	0.91	0.95
Low-high arousal	0.925	0.92	0.94
4 Class	0.89	0.87	0.92

135	2	3	4
2	65	5	5
1	9	51	3
0	10	3	53

Figure 5. Confusion matrix of simulation of the proposed method.

4. Conclusion

A hybrid deep-learning method for valence and arousal detection has been presented. The deep network combines CNN and LSTM, where CNN prepares the feature map from the ECG signal, and the data is classified by LSTM. The performance of the hybrid deep network should be such that it considers the advantage of each deep learning network. CNN has a high ability to extract features, LSTM has a high ability to classify data, and the combination of these two networks provides good results in identifying emotions. One of the strengths of the proposed method is the use of the attention mechanism in the deep network, which, with its performance, optimizes the weights and increases the classification accuracy. The proposed hybrid network is very suitable for high-dimensional data (such as emotion data).

In future research to identify emotions, it is better to determine the domain dimension in addition to valence and arousal to understand all feelings better. It is also possible to increase the accuracy of emotion classification by providing other deep hybrid networks to the point where it is suitable for daily use.

Authors contributions
Authors have contributed equally in preparing and writing the manuscript.

Availability of data and materials
The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflict of interests
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Table 4. Results of some research compared to the proposed method.

Ref.	Year	Database	Signal	Method	Accuracy
Hsu [39]	2017	Self-record	ECG	LS-SVM	Valance: 82% Arousal: 72%
Nita [40]	2022	Dreamer	ECG	CNN	Valance: 95% Arousal: 85%
Cheng [21]	2020	Dreamer	EEG	Deep Forest	Valance: 89% Arousal: 90%
Khan [41]	2023	Dreamer	ECG	Adaboost ML	Valance: 74% Arousal: 75%
Proposed method	2024	Dreamer	ECG	Attention-based CNN-LSTM	Valance: 95% Arousal: 94%

References

[1] T. Baltrušaitis, C. Ahuja, and L.-P. Morency. "**Multimodal machine learning: A survey and taxonomy**.". 41(2):423–443, 2018. DOI: https://doi.org/10.1109/TPAMI.2018.2798607.

[2] N. Ahmed, Z. Al Aghbari, and S. Girija. "**A systematic survey on multimodal emotion recognition using learning algorithms**.". *Intelligent Systems with Applications*, 17:200171, 2023. DOI: https://doi.org/10.1016/j.iswa.2022.200171.

[3] S.-W. Byun, J.-H. Kim, and S.-P. Lee. "**Multi-modal emotion recognition using speech features and text-embedding**.". *Applied Sciences*, 11(17):7967, 2021. DOI: https://doi.org/10.3390/app11177967.

[4] Z. Lian, Y. Li, J.-H. Tao, J. Huang, and M.-Y. Niu. "**Expression analysis based on face regions in real-world conditions**.". *International Journal of Automation Computing*, 17:96–107, 2020. DOI: https://doi.org/10.1007/s11633-019-1176-9.

[5] J. A. Russell. "**A circumplex model of affect**.". *Journal of Personality Social Psychology*, 39(6):1161, 1980. DOI: https://doi.org/10.1037/h0077714.

[6] A. Geetha, T. Mala, D. Priyanka, and E. Uma. "**Multimodal Emotion Recognition with deep learning: advancements, challenges, and future directions**.". *Information Fusion*, 105:102218, 2024. DOI: https://doi.org/10.1016/j.inffus.2023.102218.

[7] C. E. Izard. "**The many meanings/aspects of emotion: Definitions, functions, activation, and regulation**.". *Emotion Review*, 2(4):363–370, 2010. DOI: https://doi.org/10.1177/1754073910374661.

[8] C. Kauschke, D. Bahn, M. Vesker, and G. Schwarzer. "**The role of emotional valence for the processing of facial and verbal stimuli—positivity or negativity bias?**.". *Frontiers in Psychology*, 10: 444703, 2019. DOI: https://doi.org/10.3389/fpsyg.2019.01654.

[9] S. Kim and S.-P. Lee. "**A BiLSTM–Transformer and 2D CNN Architecture for Emotion Recognition from Speech**.". *Electronics*, 12(19):4034, 2023. DOI: https://doi.org/10.3390/electronics12194034.

[10] D. Avola et al. "**Spatio-Temporal Image-Based Encoded Atlases for EEG Emotion Recognition**.". *International Journal of Neural Systems*, 34(5):2450024, 2024. DOI: https://doi.org/10.1142/S0129065724500242.

[11] Y. Chen, H. Zhang, J. Long, and Y. Xie. "**Temporal shift residual network for EEG-based emotion recognition: a 3D feature image sequence approach**.". *Multimedia Tools Applications*, 83(15):45739–45759, 2024. DOI: https://doi.org/10.1007/s11042-023-17142-7.

[12] N. Kumari and R. Bhatia. "**Saliency map and deep learning based efficient facial emotion recognition technique for facial images**.". *Multimedia Tools Applications*, 83(12):36841–36864, 2024. DOI: https://doi.org/10.1007/s11042-023-16220-0.

[13] M. K. Chowdary, T. N. Nguyen, and D. J. Hemanth. "**Deep learning-based facial emotion recognition for human–computer interaction applications**.". *Neural Computing Applications*, 35(32):23311–23328, 2023. DOI: https://doi.org/10.1007/s00521-021-06012-8.

[14] K. Korovai, D. Zhelezniakov, O. Yakovchuk, O. Radyvonenko, N. Sakhnenko, and I. Deriuga. "**Handwriting Enhancement: Recognition-Based and Recognition-Independent Approaches for On-device Online Handwritten Text Alignment**.". *IEEE Access*, 2024. DOI: https://doi.org/10.1109/ACCESS.2024.3412433.

[15] S. Zhang, Y. Yang, C. Chen, X. Zhang, Q. Leng, and X. Zhao. "**Deep learning-based multimodal emotion recognition from audio, visual, and text modalities: A systematic review of recent advancements and future prospects**.". *Expert Systems with Applications*, 237:121692, 2024. DOI: https://doi.org/10.1016/j.eswa.2023.121692.

[16] P. Chakraborty, S. Ahmed, M. A. Yousuf, A. Azad, S. A. Alyami, and M. A. Moni. "**A human-robot interaction system calculating visual focus of human's attention level**.". *IEEE Access*, 9:93409–93421, 2021. DOI: https://doi.org/10.1109/ACCESS.2021.3091642.

[17] S. Mishra and N. Bhatnagar. "**Speech emotion recognition and classification using hybrid deep CNN and BiLSTM model**.". *Multimedia Tools Applications*, 83(13):37603–37620, 2024. DOI: https://doi.org/10.1007/s11042-023-16849-x.

[18] M. Aslan, M. Baykara, and T. B. Alakuş. "**Analysis of brain areas in emotion recognition from EEG signals with deep learning methods**.". *Multimedia Tools Applications*, 83(11):32423–32452, 2024. DOI: https://doi.org/10.1007/s11042-023-16696-w.

[19] M. Vaezi, M. Nasri, F. Azimifar, and M. Mosleh. "**BMPA-DSL: Binary Marine Predators Algorithm to Identify Driver's Different Levels of Stress**.". *14th International Conference on Information and Knowledge Technology (IKT)*, 2023. DOI: https://doi.org/10.1109/IKT62039.2023.10433031.

[20] A. Dessai and H. Virani. "**Multimodal and Multidomain Feature Fusion for Emotion Classification Based on Electrocardiogram and Galvanic Skin Response Signals**.". *Sci.*, 6(1):10, 2024. DOI: https://doi.org/10.3390/sci6010010.

[21] J. Cheng et al. "**Emotion recognition from multi-channel EEG via deep forest**.". *IEEE Journal of Biomedical Health Informatics*, 25(2):453–464, 2020. DOI: https://doi.org/10.1109/JBHI.2020.2995767.

[22] M. Vaezi, M. Nasri, F. Azimifar, and M. Mosleh. "**Driver Assistance System for Stress Recognition by Handcrafted Feature Extraction and Convolutional Neural Network**.". *20th CSI International Symposium on Artificial Intelligence and Signal Processing (AISP)*, pages 1–5, 2024. DOI: https://doi.org/10.1109/AISP61396.2024.10475225.

[23] T. Fan et al. "**A new deep convolutional neural network incorporating attentional mechanisms for ECG emotion recognition**.". *Computers in Biology Medicine*, 159:106938, 2023. DOI: https://doi.org/10.1016/j.compbiomed.2023.106938.

[24] P. Chakraborty, M. A. Yousuf, and S. Rahman. "**Predicting level of visual focus of human's attention using machine learning approaches**.". *International Conference on Trends in Computational and Cognitive Engineering: Proceedings of TCCE 2020*, pages 683–694, 2021. DOI: https://doi.org/10.1007/978-981-33-4673-4_56.

[25] M. J. Al Dujaili, A. Ebrahimi-Moghadam, and A. Fatlawi. "**Speech emotion recognition based on SVM and KNN classifications fusion**.". *International Journal of Electrical Computer Engineering*, 11(2):1259, 2021. DOI: https://doi.org/10.11591/ijece.v11i2.

[26] X. Liu, X. Cheng, and K. Lee. "**GA-SVM-based facial emotion recognition using facial geometric features**.". *IEEE Sensors Journal*, 21(10):11532–11542, 2020. DOI: https://doi.org/10.1109/JSEN.2020.3028075.

[27] M. A. Hasnul, N. A. Ab Aziz, and A. Abd Aziz. "**Augmenting ECG Data with Multiple Filters for a Better Emotion Recognition System**.". *Arabian Journal for Science Engineering*, pages 1–22, 2023. DOI: https://doi.org/10.1007/s13369-022-07585-9.

[28] Q. Gao, C.-H. Wang, Z. Wang, X.-L. Song, E.-Z. Dong, and Y. Song. "**EEG based emotion recognition using fusion feature extraction method**.". *Processes*, 79(37):27057–27074, 2020. DOI: https://doi.org/10.1007/s11042-020-09354-y.

[29] X. Wang, J. Zhang, C. He, H. Wu, and L. Cheng. "**A novel emotion recognition method based on the feature fusion of single-lead EEG and ECG signals**.". *IEEE Internet of Things Journal*, 2023. DOI: https://doi.org/10.1109/JIOT.2023.3320269.

[30] S. Wang, J. Qu, Y. Zhang, and Y. Zhang. "**Multimodal emotion recognition from EEG signals and facial expressions**.". *IEEE Access*, 11:33061–33068, 2023. DOI: https://doi.org/10.1109/ACCESS.2023.3263670.

[31] M. Vaezi and M. Nasri. "**AS3-SAE: Automatic Sleep Stages Scoring using Stacked Autoencoders**.". *Frontiers in Biomedical Technologies*, pages 400–416, 2020. DOI: https://doi.org/10.18502/fbt.v10i4.13722.

[32] S. Katsigiannis and N. Ramzan. "**DREAMER: A database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-shelf devices**.". *IEEE Journal of Biomedical Health Informatics*, 22(1):98–107, 2017. DOI: https://doi.org/10.1109/JBHI.2017.2688239.

[33] S. N. M. S. Ismail, N. A. A. Aziz, and S. Z. Ibrahim. "**A comparison of emotion recognition system using electrocardiogram (ECG) and photoplethysmogram (PPG)**.". *Journal of King Saud University-Computer Information Sciences*, 34(6):3539–3558, 2022. DOI: https://doi.org/10.1016/j.jksuci.2022.04.012.

[34] M. U. Hossain, M. A. Rahman, M. M. Islam, A. Akhter, M. A. Uddin, and B. K. Paul. "**Automatic driver distraction detection using deep convolutional neural networks**.". *Intelligent Systems with Applications*, 14:200075, 2022. DOI: https://doi.org/10.1016/j.iswa.2022.200075.

[35] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi. "**Convolutional neural networks: an overview and application in radiology**.". *Insights into Imaging*, 9:611–629, 2018. DOI: https://doi.org/10.1007/s13244-018-0639-9.

[36] Q. Zhang, T. Gao, X. Liu, and Y. Zheng. "**Public environment emotion prediction model using LSTM network**.". *Sustainability*, 12(4):1665, 2020. DOI: https://doi.org/10.3390/su12041665.

[37] Z. Liu, W. Zhou, and H. Li. "**AB-LSTM: Attention-based bidirectional LSTM model for scene text detection**.". *ACM Transactions on Multimedia Computing, Communications, Applications*, 15(4): 1–23, 2019. DOI: https://doi.org/10.1145/3356728.

[38] S. M. Alarcao and M. J. Fonseca. "**Emotions recognition using EEG signals: A survey**.". *IEEE Transactions on Affective Computing*, 10(3):374–393, 2017. DOI: https://doi.org/10.1109/TAFFC.2017.2714671.

[39] Y.-L. Hsu, J.-S. Wang, W.-C. Chiang, and C.-H. Hung. "**Automatic ECG-based emotion recognition in music listening**.". *IEEE Transactions on Affective Computing*, 11(1):85–99, 2017. DOI: https://doi.org/10.1109/TAFFC.2017.2781732.

[40] S. Nita, S. Bitam, M. Heidet, and A. Mellouk. "**A new data augmentation convolutional neural network for human emotion recognition based on ECG signals**.". *Biomedical Signal Processing Control*, 75:103580, 2022. DOI: https://doi.org/10.1016/j.bspc.2022.103580.

[41] C. M. T. Khan, N. A. Ab Aziz, J. E. Raja, S. W. B. Nawawi, and P. Rani. "**Evaluation of machine learning algorithms for emotions recognition using electrocardiogram**.". *Emerging Science Journal, vol. 7, no. 1, pp. , 2022.*, 7(1):147–161, 2022. DOI: https://doi.org/10.28991/ESJ-2023-07-01-011.