Emotion Detection Using EEG Signals and Machine Learning: A Comparative Literature Review

Nimavat Hemali Mansukhbhai[1], Prof.Manali Surani[2]
[1]*M.E. Computer Engineering (Software Engineering) B H Gardi College of Engineering and Technology (GTU)Rajkot, Gujarat, India*
[2]*Assistant Professor Department of Computer EngineeringB H Gardi College of Engineering and Technology (GTU)Rajkot, Gujarat, India*

***Abstract*—Emotion recognition has become an important area of research as emotions strongly influence human behaviour, communication and decision-making. EEG signals provide a way to understand these emotions by analyzing changes in the brain's electrical activity. Machine learning techniques can be used to identify useful patterns from EEG signals and classify different emotional states. This work studies existing EEG-based emotion recognition methods, with a focus on signal features, classification techniques and their performance. The aim is to understand the strengths and limitations of current approaches and use these findings to develop an effective EEG-based emotion recognition system. EEG emotion recognition combines signal processing and machine learning to connect patterns of brain activity with emotional states.**

***Index Terms*—EEG, Emotion Recognition, Electroencephalography, Machine Learning, EEG Signal Processing, Emotional States, Feature Extraction, Classification.**

I. INTRODUCTION

It is no secret that emotions play an important role in the work of the human mind by influencing decisions, social interactions, memory, etc. The recognition of human emotions has become a topical scientific issue that has real-world applications in healthcare, human-computer interaction (HCI), educational and entertainment areas, and many others. Traditional techniques of emotions detection involved subjective self-reporting of emotions, facial expression analysis, tone of voice and body gestures. The above-mentioned external manifestations may be hidden or controlled by a person, and often cannot be used for detection of emotions of people who are not able to communicate their feelings in traditional way, e.g., hospitalized patients, disabled or nonverbal people. In this respect, electroencephalography (EEG) gives access to direct and objective manifestation of emotions through recording of electrical activity produced by the brain, which cannot be controlled or manipulated easily because of its involuntary nature.

Rapid development of inexpensive, portable and wireless EEG acquisition technology has brought emergence of emotion recognition from EEG as an active and constantly growing research field. Development of machine learning (ML) and deep learning (DL) techniques has become an important factor in this evolution, allowing moving from classical statistical analysis to developing systems capable of automatic detection of discriminative patterns from raw or slightly preprocessed EEG data. Classical studies in this area were using traditional machine learning techniques like SVM, k-NN, Naive Bayes and Decision Tree in conjunction with manually extracted features based on entropy measures, wavelet transform and statistics of EEG signal. More recently, deep learning methods like CNN, LSTM network, BiLSTM, as well as hybrid CNN-RNN structures have shown their potential for automatic extraction of spatial-temporal EEG features, outperforming classical machine learning algorithms.

With all the developments made in this field, it is still hard to compare various research works directly because of the differences between datasets used, types of emotional states classification performed (binary, ternary, quaternary, or multiple classification), preprocessing pipelines applied, methods of windowing chosen, as well as methods of validating the models. There are some public benchmarks for research work related to affect recognition using EEG

signals like DEAP, SEED, DREAMER, and AMIGOS; however, there is no standardization between the results presented due to the usage of subject-dependent approach (when data of one subject is used both for training and testing of a model). Therefore, the question arises whether the high accuracy achieved can be attributed to a real emotion recognition capability or it is rather due to memorization of subject-specific EEG pattern by the model.

This review analyzes recent research on latest developments in the field of EEG emotion recognition based on traditional machine learning and deep learning, as well as hybrid frameworks. The main purpose is to provide an overview of the data sets and methodologies employed in the reviewed studies, namely data pre-processing, features extraction algorithms, architectures and performance of the models used. Furthermore, there are several issues that are common in the existing literature, including subject-independent validation that is inconsistent and often mislabeled, inconsistency in the windowing and pre-processing procedures used, emotion recognition paradigms that make comparison between the studies difficult, small self-collected data sets used by researchers, high computational complexity of hybrid frameworks providing the best results and inconsistent reports in abstracts and body text.

II. LITERATURE REVIEW

2.1 Machine Learning Based Emotion Detection Approaches

Traditionally, machine learning-based approaches employ a pipeline of pre-processing, feature extraction, feature selection, and classification.

In this respect, Alhalaseh and Alasasfeh [2] have employed EMD/IMF and VMD decomposition to remove noise from DEAP EEG signal data, then extracted entropy and Higuchi Fractal Dimension (HFD) as features. The Naïve Bayes, k-NN, and Decision Tree classifiers were benchmarked against the CNN classifier, where the CNN approach achieved the highest accuracy of 95.20% and k-NN had an accuracy of 94.03%.

Various non-linear features such as Higuchi Fractal Dimension, sample entropy, and permutation entropy were calculated from 3 s EEG signal data obtained from DEAP and AMIGOS datasets in order to classify six different emotions as done by Rahman et al. [6]. Ten machine learning algorithms such as SVM, KNN, Random Forest, as well as various ensemble methods such as Bagging, AdaBoost, GBDT, XGBoost, LightGBM, and Stacking were applied in this research for comparison purposes. SMOTE was used to address class imbalance, while recursive feature elimination (RFE) was applied to select relevant EEG electrodes.

Dual Decomposition method was proposed by Kamble and Sengupta [4], where Discrete Wavelet Transform and Empirical Mode Decomposition techniques were used to acquire Intrinsic Mode Function from the DREAMER EEG data set. Thirty-one features were extracted but after applying ANOVA, the number of features was reduced to 26. Comparison of the following five classical classifiers – SVM, Logistic Regression, LDA, QDA, and KNN, was performed with the following five ensemble classifiers – Random Forest, Rotation Forest, Bagging, AdaBoost, and XGBoost. Bagging technique attained the highest accuracy of all the tested ensemble learning techniques, obtaining 95.81% and 95.53% for arousal and valence, respectively. In general, the five ensemble machine learning techniques attained an average accuracy which is 5.87% and 6.08% higher than the accuracy of the five traditional machine learning techniques for arousal and valence, respectively.

Patel et al. [13] evaluated the significance of Tsallis entropy for classification of cross-subject emotional states for the SEED database. The q-value varied for Tsallis entropy and had three different values (q = 2, 3, 4). Two classes (positive and negative emotions) were used in emotion classification process by means of k-NN (k=10) classifier. The best average accuracy and F-score were 79% and 0.81, respectively, using the gamma band with q = 3. The highest reported accuracy and F-score were 84% and 0.87, respectively.

Zhang, Yin & Chen [11] used k-means clustering for labeling valence/arousal instead of fixed threshold for wavelet-based and nonlinear-dynamics (entropy) features' evaluation for DEAP binary classification. Five different feature reduction approaches were tested along with four classifiers, and the best performance (nonlinear dynamics + KSR + Random Forest) was 94.85% (arousal) and 93.90% (valence), with gamma-band features proving the most efficient.

Waheed et al. [12] proposed a low-cost and single channel EEG headset (NeuroSky, mounted on Fp1) for

detecting four emotional states (meditation, boredom, happiness, frustration) using average band power features. Four types of classifiers such as Gaussian Naïve Bayes, k-NN, Decision Tree, and SVM were compared, whereby the SVM gave the best accuracy of 94.4%.

Prakash and Poulose [15] compared nine different supervised classifiers (four conventional and five ensemble) for three-class emotion recognition on the Feeling Emotions dataset (two subjects, four-channel Muse EEG). LightGBM produced the best 5-fold cross-validation accuracy of 99.30%, though it should be noted that two subject sample is limiting the generalization ability.

Observation:

Traditional and ensemble ML techniques are also capable of producing high levels of accuracy, as reported in some works (Rahman et al. [6], Prakash and Poulose [15]), where accuracy is more than 90% and even exceeds hybrid DL techniques accuracy. Yet, the presented numbers could be affected by dataset used for evaluation. Thus, in the work of Prakash and Poulose [15], a Feeling Emotions setup with just two subjects and four-channel Muse EEG is applied. This limits the generalization of the results presented in that work.

2.2 Methods for Emotion Recognition with Deep Learning

The advantage of deep learning algorithms is that they automatically learn the features from EEG signals without relying heavily on manually crafted features.

Chao and Dong [3] introduced the approach where six time-domain features (mean, variance, standard deviation, first order difference, second order difference and approximate entropy) from 32 DEAP channels were arranged into a 9×9×6 three-dimensional matrix based on electrode location. Specifically, they designed a novel CNN architecture consisting of the univariate convolution (1×1) layer and the multi-scale multivariate convolution layers (3×3, 5×5, 7×7) and exploited the features related to both channel specific and inter-channel correlations. The results were reported as 85.88% accuracy in binary classification (arousal) and 85.53% accuracy (valence) and 76.77% accuracy in four-class classification.

The idea behind the Differential Entropy Feature Matrix (DEFM) proposed by Wang et al. [10] was to map the differential entropy in four frequency bands computed over 0.5 seconds window intervals onto a 9×9 grid depending on the electrode locations. Then a hybrid 2D-CNN-LSTM network was applied which could extract both the spatial correlation between adjacent electrodes (CNN) and temporal dependency (LSTM). They evaluated the model on the DEAP dataset using 10-fold cross validation and obtained average accuracies of 91.92% and 92.31% for valence and arousal respectively.

Ju et al. [7] designed the Temporal Difference Minimizing Neural Network (TDMNN) method, leveraging the previous knowledge of emotional states being quite stable for short periods of time. The differential entropy was used to represent features in 4D format, and the Maximum Mean Discrepancy (MMD) loss was applied in the framework of multi-branch convolutional-recurrent network in order to minimize difference between temporally adjacent feature segments. The performance was tested using four different datasets (SEED, SEED-IV, DEAP, DREAMER), showing 97.20% state-of-the-art accuracy on SEED and 99.45%/99.51% on DREAMER; leave-one-subject-out cross validation resulted in only 50-70%.

Observation:The deep learning approaches usually perform better than traditional machine learning approaches if appropriate spatial and temporal data are properly utilized together. However, the results obtained using within-subject cross validation are always significantly higher compared to subject-independent validation, as shown in Ju et al.

2.3 Hybrid Methods for Emotion Recognition Based on Deep Learning

A hybrid method combines different architectures, usually CNN with Recurrent Networks for spatial and temporal modeling or multiple models/modalities by fusion/ensemble approach.

The multimodal approach introduced by Mutawa and Hassouneh [1] consisted of facial landmarks (using 10 facial markers through Lucas-Kanade optical flow) and EEG features (based on relative band power of theta, alpha, beta, and gamma bands) and was classified using a multi-layer LSTM. A unique approach called Log-Sync was proposed to synchronize a real-time human annotation of emotions

with EEG data. The multimodal method showed accuracy of 99.3%, whereas the unimodal methods reached 90.25% using only facial landmarks and 87.25% using only EEG data on a dataset collected by the authors containing 19 subjects.

Singh et al. [5] proposed an end-to-end hybrid architecture of a 25-layer 1D-CNN and BiLSTM, trained on raw, pre-segmented DEAP EEG signals without any feature extraction. The network performed binary, ternary, and quaternary (happy/relaxed/sad/anger) emotion recognition tasks, reaching accuracies of 91.31% (binary), 89.32% (ternary) and 88.19% (quaternary). The proposed quaternary model outperformed several other existing three and four class models.

Xu et al. [14] introduced MRBiL, which calculates baseline-subtracted differential entropy over four frequency bands followed by a multi-scale parallel CNN block, a residual block (to address the problem of model degradation) and a BiLSTM block. On the DEAP dataset, the algorithm reached an accuracy of 97.87%/97.89% for binary valence/arousal classification and 96.85% for four-class classification, which is higher by at least 4% compared to the results of eight benchmark models, ACRNN and 3D-feature CNN included.

The method by Dhara et al. [8] employs a fuzzy-rank-based ensemble of three deep learning models (CNN+LSTM, CNN+GRU, and 1D-CNN), combining their outputs using the Gompertz function. On the DEAP dataset, the model achieved 90.84% and 91.65% accuracy for valence and arousal, respectively. On the AMIGOS dataset, it achieved 98.73% and 98.39% accuracy for valence and arousal, respectively. The authors also reported accuracies of 99.38% and 98.66% for the subject-independent and subject-dependent setups, respectively.

Bardak, Seyman, and Temurtas [9] suggested a mixed model using multiple parallel ANFIS classifiers (one for each emotion class) that were then connected to an adaptive network stage, applying PSD features. The accuracy was 73.49% for DEAP (4-class, valence-arousal quadrants) and 95.97% for the Feeling Emotions (3-class) dataset.

Observation:

The best performance rates are still reported by the hybrid architectures, though the performance is highly dependent on the difficulty of the particular dataset (with Bardak et al.'s own model showing the accuracy drop from 95.97% for the 3-class Feeling Emotions to 73.49% for the 4-class DEAP dataset). It is yet another proof of Research Gap 3.3 (inconsistencies in classification make comparisons unreliable).

2.4 EEG Dataset and Data Acquisition

It is important to note that the success of an emotion detection system, based on the EEG, is highly dependent on the quality of the used dataset, including the size and variety of data samples. Several publicly available and collected EEG datasets have been used in previous works. Differences between those datasets include different devices used for recording, number of subjects, number of EEG channels, emotion-induction procedure and labels.

2.4.1 DEAP Dataset

The most frequently used dataset among the mentioned papers is the DEAP dataset. It comprises 32-channel EEG signals for 32 subjects during 40 one-minute music video clips, subjective ratings of valence, arousal, dominance and liking. The research papers which use DEAP dataset are Alhalaseh and Alasasfeh [2], Chao and Dong [3], Singh et al. [5], Rahman et al. [6], Ju et al. [7], Dhara et al. [8], Bardak et al. [9], Wang et al. [10], Zhang et al. [11], and Xu et al. [14]. The use of DEAP dataset is unique or accompanied by other datasets. Nevertheless, because of the subjective nature of the labels of valence/arousal and relatively small number of subjects it is impossible to compare different works.

2.4.2 SEED Dataset

SEED is the database of 62-channel EEG signals recorded from 15 subjects during the exposure to Chinese film clips and labeling of the emotions into three classes (positive, neutral, negative). The database was used in Patel et al. [13] and Ju et al. [7]. The small number of subjects and simple three class labeling (as opposed to continuous ratings in DEAP) makes it interesting for entropy-based and cross-subject studies, yet the problem of limited number of subjects remains

2.4.3 DREAMER Dataset

DREAMER is the wireless EEG dataset recorded with the use of 14-channel Emotiv EPOC device for 23 subjects under the influence of audio-visual film clips and nine discrete emotions. The database was used in

Kamble and Sengupta [4] and Ju et al. [7]. Its portable nature and low-cost acquisition device make it valuable for practical usage, yet the small number of channels makes its comparison to DEAP or SEED problematic.

2.4.4 AMIGOS Dataset
This dataset was used by Rahman et al. [6] and Dhara et al. [8] as either a primary or secondary validation dataset for affect, personality, and mood studies.

2.4.5 Self-Collected Datasets
Self-collected datasets were proposed by Mutawa and Hassouneh [1] and Waheed et al. [12] based on portable EEG devices (Emotiv Epoc and single-electrode NeuroSky respectively). The use of these devices allows customization of emotional induction protocol but these datasets have small sample sizes, non-standardized stimuli and lack reproducibility for comparison with benchmark datasets.

Observation:
Despite DEAP is still a major benchmark, it raises issues of comparability bias. Models trained on DEAP dataset reported to be having significantly lower accuracy than SEED and self-generated datasets with the same architecture because of the difference between the continuous rating scale of DEAP and discrete emotions classification elsewhere.

2.5 Preprocessing Techniques
In the reviewed papers, the standard preprocessing procedures for EEG signal include down sampling (to 128 Hz in case of DEAP-based research), band pass filtering (4-45 Hz range) and noise/artifact rejection followed by window segmentation.
Signal decomposition-based pre-processing methods were frequently used in machine learning-based studies. Specifically, Alhalaseh and Alasasfeh [2] utilized EMD/IMF and VMD decomposition technique as a two-step pre-filter for denoising before feature extraction. Similarly, Kamble and Sengupta [4] applied a combination of DWT and EMD decompositions for frequency band isolation and Intrinsic Mode Function extraction.
Windowing method was varied in all of the deep learning studies and proved to have an impact on the outcome. Thus, Wang et al. [10] considered window sizes 0.2, 0.5, 0.8 and 1.0 seconds, with the optimum one being 0.5 seconds (91.9%/92.3% accuracy vs. 85.4% for the other window sizes). The authors of Xu et al. [14] paper chose the size 6 seconds for window and utilized the baseline subtraction procedure when first 3 seconds of resting data EEG were subtracted from the trial. The same procedure is applied by Ju et al. [7].

Observation:
Both the baseline subtraction and window length modification techniques seem to considerably influence the accuracy results but window length varies greatly for different datasets and architectures.

2.6 Techniques for Feature Extraction
Manual feature extraction is still an important part of several deep learning workflows, particularly of Differential Entropy (DE), which was introduced in the studies conducted by Wang et al. [10], Xu et al. [14], and Ju et al. [7] as the dominant feature representation, organized in a spatial arrangement according to the location of electrodes prior to being fed into CNN architectures.
Entropy non-linear features were applied widely in the traditional ML research: Alhalaseh and Alasasfeh [2] used spectral entropy and Higuchi's Fractal Dimension; Rahman et al. [6] employed Higuchi Fractal Dimension, sample entropy, and permutation entropy; Kamble and Sengupta [4] applied 31 statistical features (downsampled using ANOVA); and Patel et al. [13] examined the application of Tsallis entropy in various parameters and frequency bands.
Time-domain statistical features (mean, variance, standard deviation, high-order differences) were applied by Chao and Dong [3], organized not as plain feature vectors but in the form of 3D spatial matrices, thus preserving correlation between channels that flat feature vectors ignore.

Observation
There is an evident trend of adopting DE-based spatial-matrix feature representations in more recent deep learning research (2023-2025) instead of previous plain statistical or only entropy-based features utilized in the traditional ML research.

2.7 Comparative Analysis

Table I. Comparative Summary of Eeg-Based Emotion Detection Studies (2020–2025)

Sr. No.	Author & Year	Method/Architecture	Dataset	Accuracy / F1-score
1	Mutawa & Hassouneh (2024)	Multimodal fusion: Facial + EEG + Log-Sync + LSTM	Self-collected (19 subjects, 14-ch EEG, HD camera, 6 emotions)	Multimodal: Acc 99.30%; Face: 90.25%; EEG: 87.25%
2	Alhalaseh & Alasasfeh (2020)	EMD+VMD + Entropy/HFD + ML/CNN	DEAP (32 subjects, 32 channels, 40 music videos)	CNN: Acc 95.20%; F1: Val 94.92, Aro 95.22
3	Chao & Dong (2021)	3D EEG Feature Matrix + CNN	DEAP (32 subjects, 32 channels, 12,800 samples from 6 s windows)	Arousal: Acc 85.88%, F1 87.80%; Valence: Acc 85.53%, F1 87.03%; 4-class: Acc 76.77%, F1 76.31%
4	Kamble & Sengupta (2022)	DWT+EMD + 31 Statistical Features + Ensemble ML	DREAMER (23 subjects, 14 channels, 9 emotions)	Aro: Acc 95.81%, F1 0.81; Val: Acc 95.53%, F1 0.80; Kappa 0.79/0.77; AUC 0.98
5	Singh et al. (2023)	1D-CNN + Bi-LSTM	DEAP (32 subjects, 32 channels)	Binary: 91.31%; Ternary: 89.32%; Quaternary: 88.19%
6	Rahman et al. (2022)	HFD+Sample Entropy+Permutation Entropy + Ensemble ML	DEAP (32 ch) and AMIGOS (14 ch)	DEAP: Acc 89.38%; AMIGOS: Acc 94.62%
7	Ju et al. (2024)	TDMNN + MMD + CNN-RNN	SEED, SEED-IV, DEAP, DREAMER	SEED: 97.20%; SEED-IV: 89.70%; DEAP: V 98.08%, A 98.25%; DREAMER: V 99.45%, A 99.51%
8	Dhara et al. (2024)	1D-CNN + LSTM/GRU + Fuzzy Ensemble	DEAP (32) and AMIGOS (40 subjects, subset used)	DEAP: V 90.84%, A 91.65%; AMIGOS: V 98.73%, A 98.39%
9	Bardak et al. (2024)	Hybrid ANFIS + PNN + RBFNN with PSD features	DEAP (32 subjects, 32 channels) Feeling Emotion dataset	DEAP: Acc 73.49%; Feeling Emotions: Acc 95.97%
10	Wang et al. (2024)	DEFM + 2D-CNN-LSTM	DEAP (32 subjects, 32 channels, 4800 samples/subject)	Val: Acc 91.92%, F1 91.85%; Aro: Acc 92.31%, F1 90.75%
11	Zhang, Yin & Chen (2023)	KSR/LPP/mRMR/Relieff/ PCA + RF/KNN/NB/SVM	DEAP (32 subjects, 32 channels)	Best (NDA+KSR+RF): Aro 94.85±1.77%, Val 93.90±2.23%
12	Waheed et al. (2021)	Single-electrode EEG + Band Power + SVM/KNN/GNB/DT	Self-collected, images as stimuli	SVM: Acc 94.4%, Specificity 98.18%, Sensitivity 94.44%, Precision 91.6%; KNN/GNB 88.8%, DT 83.34%
13	Patel et al. (2024)	Tsallis Entropy + KNN	SEED (15 subjects, 62 channels)	Max Acc 84%, F1 0.87; Avg Acc 79%, F1 0.81
14	Xu et al. (2025)	DE + Multi-scale CNN + ResBlock + BiLSTM	DEAP (32 subjects, 32 channels)	Val: Acc 97.87%, F1 97.52%; Aro: Acc 97.89%, F1 97.19%; 4-class: Acc 96.85%
15	A. Prakash and A. Poulose (2025)	LR, SVM, DT, KNN, RF, AdaBoost, LightGBM, XGBoost, CatBoost	Feeling Emotions – 2 subjects, 4-channel Muse EEG, 3 emotions	Acc 99.30%, F1 99.30% (LightGBM)

II. RESEARCH GAPS

Based on the reviewed studies, the following research gaps have been identified:

3.1. Insufficient and Ambiguous Cross-Subject Validation:

Although several studies describe their evaluation as subject-independent (Xu et al. [14], Dhara et al. [8],

Patel et al. [13]), their validation procedures do not consistently follow true leave-one-subject-out evaluation. Ju et al. [7] reported an average accuracy of 97.20% on SEED under their evaluation protocol. However, this result should not be directly compared with leave-one-subject-out performance unless the validation protocols and experimental settings are confirmed to be comparable.

3.2. Variation of Windowing and Preprocessing Approach:

There is an inconsistency in the optimal EEG windows: 0.5 s (Wang et al. [10]), 4 s (Zhang et al. [11]), 6 s (Xu et al. [14]), 15.75 s (Singh et al. [5]). The procedure of baseline subtraction is utilized by certain studies (Xu et al. [14], Ju et al. [7], Zhang et al. [11]) but not by others; therefore, comparison between different works becomes unreliable even on the same database (DEAP).

3.3 Inconsistent Classification Strategies for Emotions:

Research is inconsistent when it comes to classification strategies for emotions – binary (Wang et al. [10], Xu et al. [14]), ternary (Singh et al. [5]), quaternary (Singh et al. [5], Xu et al. [14]) or discrete multi-class (Mutawa: 6 classes [1]; Rahman et al.: 6 classes [6]; Waheed et al.: 4 classes [12]). Either valence-arousal thresholding or discrete emotion categories are used. This makes direct comparison of reported accuracy across studies difficult, since binary and six-class classification represent different classification tasks.

3.4 Self-Collected Small Datasets

Self-collected datasets, such as the dataset used by Mutawa et al. [1] involving 19 subjects, may limit generalizability and reproducibility because of sample size and differences in data-collection protocols. Bardak et al. [9] and Prakash and Poulose [15] both evaluated their proposed approaches using the Feeling Emotions dataset. Prakash and Poulose [15] reported 99.30% accuracy using the LightGBM algorithm with their experimental setup involving two participants and four-channel Muse EEG. The limited number of participants may affect the generalizability of their reported results.

3.5 High Computational Complexity of Hybrid Models:

Several hybrid models reporting high classification performance, such as MRBiL [14], TDMNN [7], and the fuzzy ensemble [8], combine multiple deep learning architectures (CNN + LSTM/GRU + attention/residual/ensemble techniques) which may increase computational time and resource requirements, making deployment on portable real-time EEG systems (as in work of Mutawa et al. [1] and Waheed et al. [12]) more challenging.

3.6 Inconsistencies in Reporting of Numbers:

Several publications show inconsistencies in reported performance values across the abstract, tables, and main body of the paper (Alhalaseh and Alasasfeh [2], Rahman et al. [6], Wang et al. [10]).

III. CONCLUSIONS

In this paper, we conducted a comparative literature review of fifteen studies devoted to EEG-based emotion detection in the period between 2020 and 2025. The reviewed works show the development of techniques for emotion recognition, from traditional machine learning to deep learning and hybrid approaches.

As observed, traditional machine learning techniques such as SVM, Random Forest, Decision Tree, and k-NN provide satisfactory results for EEG-based emotion classification when combined with effective feature extraction methods. At the same time, deep learning models such as CNN, LSTM, GRU, and hybrid CNN-RNN architectures have demonstrated strong classification capabilities through automatic feature learning.

There are several challenges in the field which have been observed in the majority of papers. First of all, small number of samples in existing databases can be considered as one of the major disadvantages. In addition, there is a problem of insufficiency of validation performed on an independent group of subjects. Moreover, the inconsistency of the protocols used in experiments prevents fair comparison of proposed solutions.

REFERENCES

[1] M. Mutawa and A. Hassouneh, "Multimodal real-time patient emotion recognition system using facial expressions and brain EEG signals based on machine learning and Log-Sync methods," *Biomedical Signal Processing and Control*, vol. 91, Art. no. 105942, 2024, doi: 10.1016/j.bspc.2023.105942.

[2] R. Alhalaseh and S. Alasasfeh, "Machine-learning-based emotion recognition system using EEG signals," *Computers*, vol. 9, no. 4, Art. no. 95, 2020, doi: 10.3390/computers9040095.

[3] H. Chao and L. Dong, "Emotion recognition using three-dimensional feature and convolutional neural network from multichannel EEG signals," *IEEE Sensors Journal*, vol. 21, no. 2, pp. 2024–2034, 2021, doi: 10.1109/JSEN.2020.3020828.

[4] K. S. Kamble and J. Sengupta, "Ensemble machine learning-based affective computing for emotion recognition using dual-decomposed EEG signals," *IEEE Sensors Journal*, vol. 22, no. 3, pp. 2496–2507, 2022, doi: 10.1109/JSEN.2021.3135953.

[5] K. Singh, M. K. Ahirwal, and M. Pandey, "Quaternary classification of emotions based on electroencephalogram signals using hybrid deep learning model," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, pp. 2429–2441, 2023, doi: 10.1007/s12652-022-04495-4.

[6] M. M. Rahman, A. K. Sarkar, M. A. Hossain, and M. A. Moni, "EEG-based emotion analysis using non-linear features and ensemble learning approaches," *Expert Systems with Applications*, vol. 207, Art. no. 118025, 2022, doi: 10.1016/j.eswa.2022.118025.

[7] X. Ju, M. Li, W. Tian, and D. Hu, "EEG-based emotion recognition using a temporal-difference minimizing neural network," *Cognitive Neurodynamics*, vol. 18, pp. 405–416, 2024, doi: 10.1007/s11571-023-10004-w.

[8] Dhara, R. Sarkar, S. K. Ghosh, and R. K. Tripathy, "A fuzzy ensemble-based deep learning model for EEG-based emotion recognition," *Cognitive Computation*, vol. 16, pp. 1364–1378, 2024, doi: 10.1007/s12559-023-10171-2.

[9] F. K. Bardak, M. N. Seyman, and F. Temurtas, "Adaptive neuro-fuzzy based hybrid classification model for emotion recognition from EEG signals," *Neural Computing and Applications*, vol. 36, pp. 9189–9202, 2024, doi: 10.1007/s00521-024-09573-6.

[10] T. Wang, X. Huang, Z. Xiao, W. Cai, and Y. Tai, "EEG emotion recognition based on differential entropy feature matrix through 2D-CNN-LSTM network," *EURASIP Journal on Advances in Signal Processing*, vol. 2024, Art. no. 49, 2024, doi: 10.1186/s13634-024-01146-y.

[11] J. Zhang, Z. Yin, and P. Chen, "EEG-based affect classification with machine learning algorithms," *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 11627–11632, 2023, doi: 10.1016/j.ifacol.2023.10.486.

[12] S. A. Waheed, S. Revathi, M. A. Matheen, A. K. Lodhi, M. Ashrafuddin, and G. S. Maboobatcha, "Processing of human motions using cost effective EEG sensor and machine learning approach," in *Proc. 2021 1st International Conference on Artificial Intelligence and Data Analytics (CAIDA)*, Chennai, India, 2021, pp. 1–6, doi: 10.1109/CAIDA51941.2021.9425088.

[13] P. Patel, S. Balasubramanian, and R. N. Annavarapu, "Cross subject emotion identification from multichannel EEG sub-bands using Tsallis entropy feature and KNN classifier," *Brain Informatics*, vol. 11, Art. no. 7, 2024, doi: 10.1186/s40708-024-00220-3.

[14] Y. Xu, Y. Gao, Z. Zhang, and S. Du, "Emotional recognition of EEG signals utilizing residual structure fusion in bi-directional LSTM," *Complex & Intelligent Systems*, vol. 11, Art. no. 62, 2025, doi: 10.1007/s40747-024-01682-y.

[15] Prakash and A. Poulose, "Electroencephalogram-based emotion recognition: A comparative analysis of supervised machine learning algorithms," *Data Science and Management*, vol. 8, pp. 342–360, 2025, doi: 10.1016/j.dsm.2024.12.004.

[16] K. R. Dharmendra, D. K. Verma, and S. Vats, "Neuro shield: Securing neural data with privacy-preserving algorithms for brain computer interface," *Journal of Information*

and Optimization Sciences, vol. 47, no. 6, pp. 2263–2279, 2026, doi: 10.47974/JIOS-2323.

[17] K. R. Dharmendra, D. K. Verma, and S. Vats, "Hybrid EEG-based eye state classification using LSTM, neural networks, and multivariate analysis," *Journal of Information Technology Management*, vol. 17, Special Issue on Intelligent Security and Management, pp. 16–31, 2025, doi: 10.22059/jitm.2025.102919.

[18] K. R. Dharmendra, D. K. Verma, and S. Vats, "Advancing brain-computer interfaces: A deep learning approach for enhanced neural signal processing," in *Proc. 2025 International Conference on Automation and Computation (AUTOCOM)*, Mar. 2025, pp. 1616–1621, doi: 10.1109/AUTOCOM64127.2025.10956305.