

Stress Level Based Emotion Classification Using Hybrid Deep Learning Algorithm

Sivasankaran Pichandi^{1*}, Gomathy Balasubramanian², and Venkatesh Chakrapani³

¹Electronics and Communication Engineering, Sengunthar Engineering College,
Tiruchengode, Tamilnadu, India
[e-mail: sivaarm05@gmail.com]

²Computer Science and Engineering, Dr. NGP Institute of Technology,
Coimbatore, Tamilnadu, India
[e-mail: bgomramesh@gmail.com]

³Electronics and Communication Engineering, Builders Engineering College,
Kangayam, Tamilnadu, India
[e-mail: prof.c.venkatesh@gmail.com]

*Corresponding author: Sivasankaran P

*Received May 5, 2023; revised June 29, 2023; revised September 12, 2023; accepted October 9, 2023;
published November 30, 2023*

Abstract

The present fast-moving era brings a serious stress issue that affects elders and youngsters. Everyone has undergone stress factors at least once in their lifetime. Stress is more among youngsters as they are new to the working environment. whereas the stress factors for elders affect the individual and overall performance in an organization. Electroencephalogram (EEG) based stress level classification is one of the widely used methodologies for stress detection. However, the signal processing methods evolved so far have limitations as most of the stress classification models compute the stress level in a predefined environment to detect individual stress factors. Specifically, machine learning based stress classification models requires additional algorithm for feature extraction which increases the computation cost. Also due to the limited feature learning characteristics of machine learning algorithms, the classification performance reduces and inaccurate sometimes. It is evident from numerous research works that deep learning models outperforms machine learning techniques. Thus, to classify all the emotions based on stress level in this research work a hybrid deep learning algorithm is presented. Compared to conventional deep learning models, hybrid models outperforms in feature handing. Better feature extraction and selection can be made through deep learning models. Adding machine learning classifiers in deep learning architecture will enhance the classification performances. Thus, a hybrid convolutional neural network model was presented which extracts the features using CNN and classifies them through machine learning support vector machine. Simulation analysis of benchmark datasets demonstrates the proposed model performances. Finally, existing methods are comparatively analyzed to demonstrate the better performance of the proposed model as a result of the proposed hybrid combination.

Keywords: Stress analysis; Emotion analysis; Deep learning; Classification; Convolutional neural network.

A preliminary version of this paper appeared in IEEE ICC 2020. June 15-19, Dresden, Germany. This version

1. Introduction

Stress is a major factor that leads to chronic disorders and makes people more uncomfortable. Stress affects problem-solving power, and decision-making intelligence disturbs imagination and memory factor. Stress affects the prefrontal cortex activities of people which directly reflects in their daily lives. Activity changes in cerebral, immune, and endocrine systems are possible due to stress. When the physical or psychological tensions are processed by the brain, it triggers hormones and antibodies secretion. Long-time exposure to stress levels affects health and weakens the immune system. Stress-related corticosteroid hormones segregate due to stress which affects the immune system and makes people more susceptible to infection. Long-time stress factors lead to heart diseases, diabetes, etc., and Depression due to stress reduces the activities of a person. Thus, it is essential to take precautionary measures to analyze the stress factors and their levels to secure individual health as well as society's welfare.

In our daily life communication and interaction, emotion plays a major role. Emotions are high level psychological experience which is faced by everyone in their daily life. Emotions occurs due to complex feelings, unease, comfort, discomfort, and pleasure. Numerous studies on emotion classification and recognition evolved in recent times in education, robotics, driving assistance, education, and health domains. Interpreting human feelings through emotion analysis gained more attention in recent times. Generally, emotion analysis models are described as discrete and dimensional models. In the discrete models, the emotions are categorized into fear, happiness, sadness, etc., whereas dimensional models categorize the emotions as valence and arousal. **Fig. 1** depicts an illustration of basic emotions in valence-arousal plane.

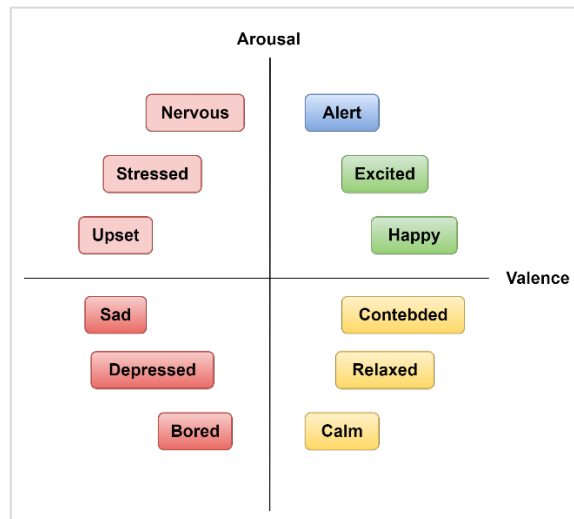


Fig. 1. Basic emotions in arousal -valance plane

The emotion states given in four quadrants categorize different types of emotions. To classify these emotions, several modalities like speech, facial expression and body gestures are used. However, these analyses are accurate at sometimes as the emotions are subjective to individual willing and sensitive to social masking. To overcome these limitations, emotions are analyzed thorough physiological signals where electroencephalogram (EEG), skin temperature, electrocardiogram (ECG), galvanic skin response are used to define the emotions.

True emotions can be detected from the physiological responses.

In the medical domain, stress analysis is considered as one of the complex processes specifically in oncology or chronic disease management. Generally based on human responses and perceptual factors the stress factors are measured. To describe the stress factors and mental state, continuous measurements are required. The brain-related physiological information can be analyzed through functional near-infrared spectroscopy images and Electroencephalography (EEG) signals. At other times, based on heart rate variability, the stress factors are analyzed. In addition to heart rate variability analysis, a few other measurements like skin temperature, blood pressure, and galvanic skin response are considered to define the stress factors.

Earlier for stress analysis, machine learning models are widely used [1-2]. Familiar machine learning models like decision tree, support vector machine, random forest, and logistic regression-based stress analysis models are evolved. However, when the deep learning models proved better performance in all domains over machine learning models recent research works provide numerous deep learning-based stress analysis processes. The major advantage of the deep learning model is its automatic feature abstraction and maps the extracted features to the target. Convolutional neural network is one of the familiar deep learning models which is widely used in object detection, image classification, computer vision applications, etc. [3] However, the classification performance of the convolutional neural network model can be improved if the traditional classifier layer is replaced. Thus, to attain better classification performance in stress-based emotion analysis, a hybrid deep learning model is presented in this research work. The contributions made in this research work are presented as follows.

1. Emotion classification from EEG data is proposed by incorporating the features benefits of deep learning algorithm. For accurate and enhanced emotion classification a hybrid deep learning model is presented which analyzes different states of emotions.
2. Emotion classification model includes convolutional neural network with machine learning support vector machine algorithm as a hybrid model to define two emotion plane arousal and valance. The novel classification model has unique architecture which replaces the classification layer of CNN with support vector machine for enhanced classification performances.
3. An intense experimental analysis is provided to verify the proposed model performance using a benchmark DEAP and EEG brainwave datasets considering the metrics recall, precision, f1-score, specificity, Mathew correlation coefficient, and accuracy.
4. A comparative analysis of the proposed hybrid model and existing models like support vector machine, convolutional neural network, bidirectional LSTM, and artificial neural networks is presented for better validation.

The remaining discussions in the article are arranged in the following order. A brief literature review is presented in section 2, the proposed hybrid deep learning model is presented in section 3, experimental results and comparative analysis are presented in section 4, and a summary of the research work is presented in the last section.

2. Related Works

A brief literature review of existing stress analysis models is presented in this section. The existing research works are analyzed based on the methodology and its features. The articles

are selected based on machine learning techniques and deep learning techniques. The initial discussion covers the machine learning based stress and emotion recognition, prediction models. The later part of the discussion covers the deep learning based techniques on stress and emotion classification.

2.1 Machine learning based emotion classification models

A Gradient boosting model presented in [4] analyzes the categorizes different types of emotions based on the noise signals in traffic and the EEG responses of people. Emotions like stress, excitement, relaxation, engagement, and interest are categorized through the boosting model, and observed that loudness and spectral content produces impacts in excitement and engagement indices. However, when the loudness and spectral content increase beyond the limit then the excitement and engagement turn into a stress factor. A similar gradient-boosted tree classifiers-based emotion classification model is presented in [5] incorporates filtering to preprocess the input and employed random forest to extract the features from the frequency and time domain. Further classifiers like k-NN, SVM, and boosted tree classifiers are employed for emotion classification. Experimentation results provide that the performance of boosted gradient tree model-based emotion classification is better than other models.

Machine learning approaches for emotion classification are presented in [6] using a benchmark emotion dataset. The presented approach extracts the wavelet features from EEG signals and classifies them through k-nearest neighbor(kNN), fine-tuned kNN, and support vector machine. Experimentation confirms that the performance of kNN is better than the other models. The emotion analysis model presented in [7] extracts root means square successive difference and heart rate variability from the ECG and classifies the happy and angry emotions. The presented approach processes the ECG signals and classifies them using a support vector machine. Due to the utilization of audio and visual stimulus in the classification process presented approach attained better performance than traditional approaches. An emotion prediction model presented in [8] analyzes the EEG signals of brain damage patients and extracts the emotions induced while using audio-visual stimuli. Once the data is collected it is normalized and filtered using butter worth band pass filter to cut off the frequencies. The frequency bands are localized using wavelet packet transform and classified using kNN and probabilistic neural network models.

An emotion classification model presented in [9] developed a federated learning classifier that secures user information and analyzes emotions without accessing local data. The presented classification model considers the physiological data streams and neglects the other information thus it provides better privacy and security for user information in real-time emotion analysis. A hierarchical extreme puzzle learning machine is presented in [10] for emotion classification. The presented approach initially preprocesses the input using a smoothing filter and employed wavelet scattering to convert the signal into an image. For feature extraction ResNet and Inception-v3 models are incorporated. The extracted features are combined using a puzzle optimization algorithm. To reduce the feature dimensionality hierarchical extreme learning machine is used which further improves the classification accuracy by reducing the feature dimensionality. The presented approach attained better classification performances than existing machine learning-based approaches.

The stress recognition model presented in [11] has two stage feature subset selection procedure in which a multi-cluster feature selection is employed in the first stage to identify the relevant features and to minimize the feature space. Biogeography and particle swarm optimization algorithms are incorporated in the second stage to reduce the feature dimensions and effectively differentiate different emotions. Experimentations on the benchmark dataset

confirm that the presented approach attained better classification results compared to existing methods. Stress recognition based on photoelectric information is presented in [12] which determines the stress level based on the physical stress information and emotional stress information. The presented Eulerian magnification canonical correlation model performs signal amplification and correlation analysis. Further sparse coding and canonical correlation are employed for fusing amplified features and original signals. Experimentations provide better classification accuracy which is better than the existing methods.

An emotional speech recognition model presented in [13] reduces the combinational features using wavelet coefficients. The presented approach utilizes spectral features, Mel frequency cepstral coefficients and extracts the information as emotion features. Further using the radial basis function network, the emotional features are classified and attained better performance over traditional methods. An emotion recognition model presented in [14] employed a stationary wavelet entropy model for initial-level feature extraction. A single hidden layer feedforward neural network is used to classify the features. To avoid local optimal solutions in the training process, Jaya optimization algorithm is incorporated which improves the classification performances.

Detecting and labeling human stress levels for stress identification is focused on [15] and to attain the desired objective a finite state machine-based classification model is presented which identifies the relationship between induced stress factors and temporal signals. Results categorize stress into three levels as high, low, and medium, and identify stressful situations effectively. The stress analysis model presented in [16] incorporates fuzzy logic for stress detection based on physiological signals like breath, heart rate, and skin response. The R peaks of ECG signals are initially enhanced through wavelet processing and decompose the signals into two components as average value and its variations. Final classification effectively detects stress changes based on frequent composition and provides better results than existing methods.

A hybrid optimization model for emotion recognition from speech signals is presented in [17] using a particle swarm optimization algorithm with a feed-forward network. The presented adaptive learning architecture learns the speech features. Similarly, the features from the particle swarm optimization algorithm are obtained and combined with network features to train the system. This combined feature selection and classification procedure improves the recognition performances compared to conventional methods. A multimodal emotion classification model presented in [18] provides a hybrid feature extraction strategy and adaptive decision fusion model. Initially, the statistical and even related features are extracted from the physiological signals to describe the context and statistical dependent characteristics. Further an adaptive decision fusion model is presented to classify the emotions into different categories better than traditional approaches. The machine learning model presented in [19] generates multi-modal emotion features to recognize different types of emotions from videos. The limitations like constrained interactive utilization and conflict in modal features are overcome by the presented multi-model feature classification model. The presented approaches select multiple modalities and assigns different weights based on attention network. meanwhile the complementary relationship between the modalities is defined and used in the feature reconstruction process. Finally, the features are fused to attain a complete multi-modal feature in emotion recognition process and attained better performances than traditional methods.

Though the performance of machine learning based emotion and stress classification models performs well. However, these techniques require optimal features to process the features in minimum time. Otherwise, the models process all the features and takes more time for computation. Moreover, the attained accuracy seems to be high, but the real fact is it is

inaccurate due to improper feature selection and processing in machine learning based techniques. compared to machine learning deep learning can extract more relevant optimal features for classification process so that better classification or detection performance can be attained when using deep learning techniques in emotion and stress classification.

2.2 Deep learning based emotion classification models

The deep learning model presented in [20][21] fuses the multi-modal information with heterogenous signals to improve the stress detection performance. The collected video data, respiration, and ECG data are preprocessed and fed into a deep neural network along with facial feature sequences. The deep neural network model processes the signal and fuse the information based on feature and based on decision. The results of the deep neural network model are a three-class classification of stress factors which is better than traditional classification models. The deep learning-based emotion recognition model presented in [22] performs two stages of learning for understanding the ECG representations and classifying the emotions. Initially, the high-level abstracts are obtained from the unlabeled ECG data. Further to convert the signals into pretext tasks, six different signal transformations are applied. This transformation is used to enhance the spatiotemporal representations and generalize the data to classify emotions into different categories. Experimentations provided that the multi-task structure performs better than the single-task structure and provides optimal solutions for the pretext of self-supervised tasks.

The facial emotion recognition model presented in [23] includes a graph neural network to process the facial expressions from human visual perception. The presented approach initially divides the face into six different parts and selects the key facial feature from each part as graph nodes. The selected features are then combined to learn the facial expressions. Experimentation confirms that the presented approach has better generalization and characterization ability compared to conventional facial recognition models. An automatic emotion recognition model presented in [24] incorporates VGGNet to learn the input image facial features. The presented approach extracts the emotional features based on tracking the facial nerves. Deep learning-based feature extraction and processing provides better classification accuracy and effectively categorizes the fear, anger, and disgust emotions. A similar emotion recognition model presented in [25] employs a deep neural network for the complex classification of emotions based on facial landmarks. The presented approach incorporates a sampling procedure to avoid padding in the process of extracting facial landmarks. Further, the randomized classifier performance is compared with the non-randomized classifier to validate the performances in terms of training time and accuracy. A deep convolution graph network-based emotion classification model presented in [26] combines the feature merits of graph network and convolution features to attain better performances. The presented approach overcomes the interference limitations in EEG channels and establishes a better relationship between brain regions to analyze emotions. Experimentations on benchmark datasets confirmed better performance when compared to existing machine learning-based approaches. From the literature analysis, it is observed that deep learning models are recently utilized in emotion analysis. However, the performance of the traditional deep learning models-based emotion analysis can be improved if hybrid models are employed. Thus, a hybrid deep learning model is presented for emotion analysis using a convolutional neural network and support vector machine algorithms.

3. Proposed Work

The proposed stress level-based emotion classification model using a hybrid deep learning model is presented in this section. The hybrid model incorporates convolutional neural network architecture and support vector machine for accurate classification. The classifier layer in the traditional CNN architecture is replaced with a support vector machine to enhance the classification accuracy. Fig. 2 depicts a simple overview of the proposed emotion classification model. The initial preprocessing in the classification model removes the noise artifacts using a bandpass filter. Followed by filtering, the essential features are extracted using deep convolution layers in the CNN model and classified through a support vector machine. The final classification results provide the status of emotion based on the features extracted from the input data.

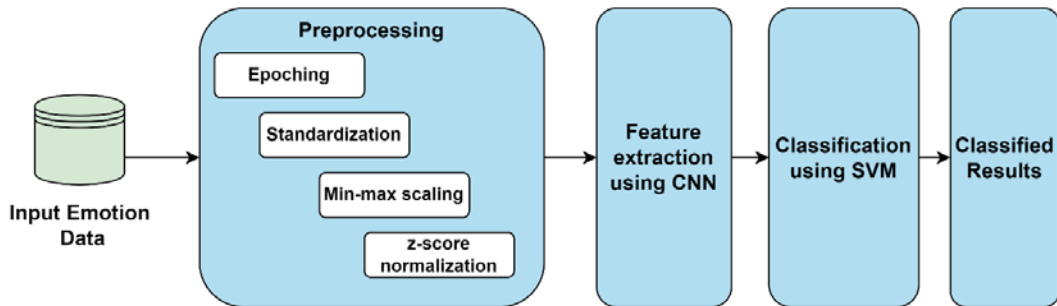


Fig. 2. Proposed model overview

The preprocessing steps in the proposed model include Epoching, standardization, min-max scaling, and z-score normalization. In the epoching, a specific time window from the continuous EEG signal is extracted. For example, if the task is performed to extract 2 seconds wavelength of each signal for 2 second time step, then 31 signals for the channel can be obtained with 256 data points for each epoch. All these data are provided based on the benchmark dataset used in the analysis. Then standardization in the preprocessing step employs two methods like min-max scaling and z-score normalization. This standardization prevents overfitting and also improves the model's accuracy. Min-max scaling is a standardization method that scales the data into a fixed range from 0 to 1 and it is mathematically formulated as

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

where x_{min} is the minimum range and x_{max} is the maximum range. Further, the standardized results are normalized using z-score normalization in which the features are rescaled so that the features will have the standard normal distribution properties. Mathematically z-score normalization is formulated as

$$z = \frac{x - \mu}{\sigma} \quad (2)$$

where the mean is indicated as μ and the standard deviation is indicated as σ . Further, the normalized data is fed into the deep learning model. The proposed deep learning model includes a one-dimensional convolution layer and pooling layers since all the input data are time series data and a one-dimensional network is very powerful in processing these data. The input layer has the normalized data and the convolution layer selects the optimal features from the input and max pooling is employed to reduce the feature dimensions. Four convolution

layers and max pooling layers are included in the architecture. For each convolution layer, an activation function is used such as ReLU (Rectified linear unit), and a dropout function is employed after max pooling to avoid data overfitting. Finally, after the fourth pooling layer, the features are flattened and fed into the machine learning classifier to classify different types of emotions. Fig. 3 depicts the complete architecture details of the proposed hybrid deep learning model.

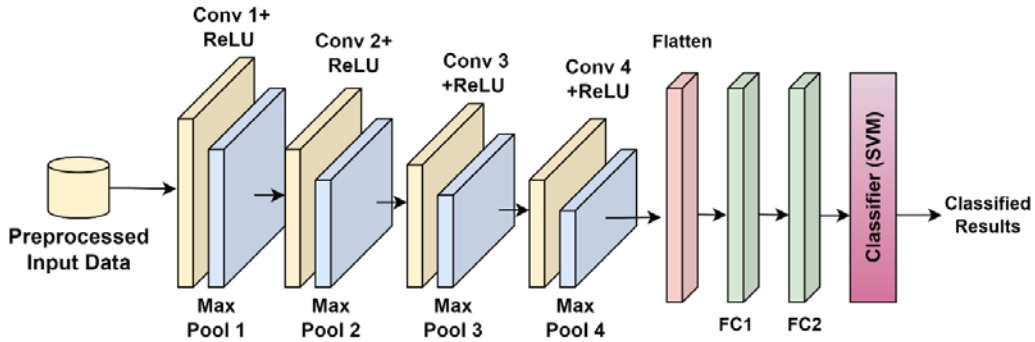


Fig. 3. Proposed hybrid deep learning model

As mentioned earlier the architecture of CNN includes a convolution layer, pooling layer, activation layer, and classifier layer. The basic architecture of CNN follows forward and backpropagation and the forward propagation algorithm can be optimized using backpropagation algorithms. Multiple convolution layers are included in the forward propagation and extract different features from the input. The low-level features are extracted in the first layer and the complex features are extracted in the subsequent convolution layers. Additionally, the features obtained in the previous layers are mapped to the specific feature of the data. Mathematically the convolution process is formulated as

$$y^m = \sum_i y_i^{m-1} * w_{ij}^m + c_i^m \quad (3)$$

where the input data is represented as y_i^{m-1} and it is also termed as previous convolution layer outputs. The weight matrix is represented as w_{ij}^m and the bias vector is represented as c_i^m . After convolution batch normalization is used to overcome the gradient loss in the training process. The abnormal distribution is data will be under control after batch normalization. However, in the proposed model preprocessing all the normalizations are performed thus it is removed from the proposed architecture. Generally, batch normalization is defined as

$$n_c = \frac{1}{l} \sum_{i=1}^l x_i \quad (4)$$

$$v_c = \frac{1}{l} \sum_{i=1}^l (x_i - n_c)^2 \quad (5)$$

$$\hat{x}_i = \frac{x_i - n_c}{\sqrt{v_c^2 + \epsilon}} \quad (6)$$

$$y_i^m = d\hat{x}_i + s \quad (7)$$

where the input is represented as x_i and batch mean is represented as n_c and its variance is represented as v_c . The parameters d and s are shift factors and learnable parameters which are used to adjust to optimal value in the training process. An activation function is used in the architecture after convolution, instead of other activation functions like sigmoid, tangent, and liner function, Rectified Linear Unit (ReLU) is used, and it is mathematically formulated as

$$r_i^m = \max(0, y_i^m) \quad (8)$$

where r_i^m indicates the ReLU layer output and it neglects the negative values in the input data. The pooling layer in the architecture reduces the input size from the convolution layer and due to this the computation speed of the deep learning model increases. The forward propagation model includes the max pooling function which is formulated as

$$p_i^m = \text{maximum}\{r_i^m\} \quad (9)$$

Similar properties are combined in the pooling layer and it reduces the properties matrix size in the convolution layers. Finally, after convolution and pooling operations, the features are flattened and then fed into a fully connected layer which enables data transmission between the classification layer and previous layers. The final classifier layer includes the SoftMax function in the classification process. However, to improve the classification performances instead of the SoftMax function, a support vector machine is employed in the proposed architecture. Compared to SoftMax, SVM can able to produce stable results and trains faster while SoftMax lags in performance due to multiple calculations. The training time of softmax is high if the data has multiple labels. Thus, the proposed model includes support vector machine for final calculation. The results of the support vector machine provide the emotion status based on the features. a detailed experimental analysis of the proposed hybrid deep learning model is presented in the following section.

4. Results and Discussion

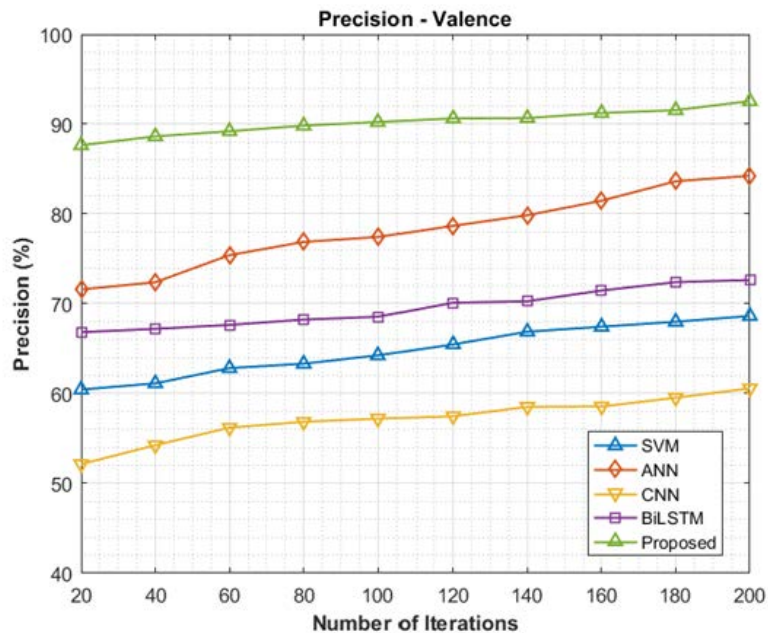
The proposed hybrid deep learning model performance is verified through simulation analysis performed in MATLAB and the benchmark DEAP [27] and Brainwave [28] datasets are used to evaluate the performances. The benchmark dataset DEAP database is an emotion database that has EEG and peripheral recordings. DEAP dataset has physiological recordings of 32 participants and measured their emotions while watching a video. Based on the observation, the user provided ratings for the video as arousal, valence, dominance, liking, and familiarity. Participants of different age groups and different genders are involved to provide their feedback. A few participants' face frontal videos are also included in the dataset. EEG channels are sampled at the 512Hz range and electrodes are used to record their responses. For peripheral signals, blood volume, respiration amplitude, electrocardiogram, skin temperature, electrooculogram, trapezius muscles, and zygomaticus are used. The recent version of the dataset has signals which are down-sampled at 128Hz and the raw data is segmented into 32 files. Each file has arrays in which one is used as a data array which has video, channel, and data and the other is a label array which is video and label information. The label indicates the rating for different emotions like arousal, valence, etc., The initial preprocessing step includes noise removal followed by filtering. A bandpass filter is used in the range of 4-45Hz to allow only the desired range of signals. The second dataset EEG brainwave dataset categorizes emotions into three types as negative, positive and neutral. The participant emotions are measured by measuring the micro voltage obtained from the electrodes placed over the head. Six film clips of 12 minutes are played to the participants to measure the brain activity. Per day three minutes of data are collected from the participants to avoid interferences. The signals are sampled to 150Hz so that the final dataset has 324000 data points. Performance metrics like accuracy, precision, f1-score, recall, specificity, and Mathew's correlation coefficient are considered for performance evaluation. The performance metrics of the proposed model are depicted in **Table 1** for valence and arousal labels.

Table 1. Proposed work performance metrics

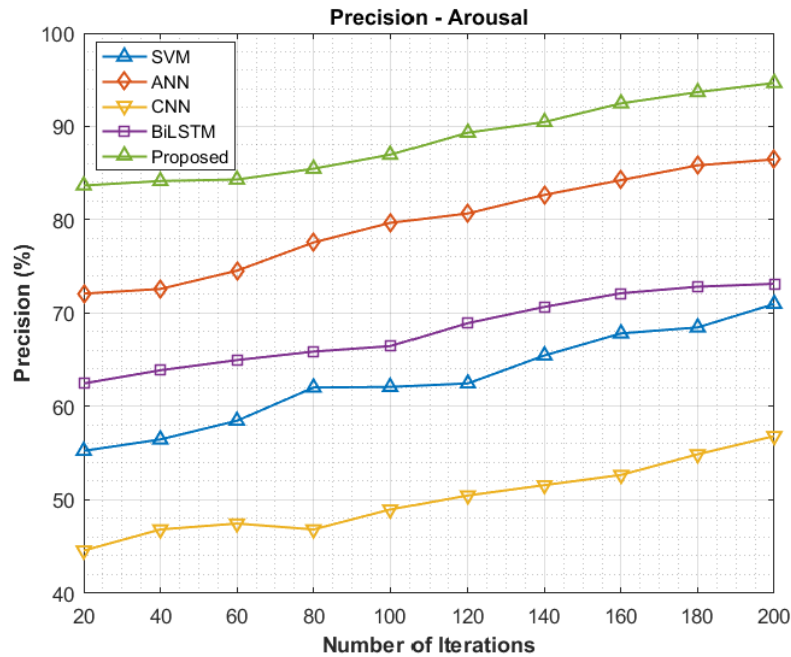
S. No	Metrics	Labels	
		Valence	Arousal
1.	Accuracy	93.45	95.68
2.	Precision	92.56	94.66
3.	F1-Score	92.07	93.66
4.	Recall	91.58	92.68
5.	Specificity	90.56	91.68
6.	Matthews Correlation	92.41	94.48

Further to validate the performance of the proposed model, existing methods that are evolved to detect different types of emotions are used for comparative analysis. Methods like SVM [29], BiLSTM [30], CNN [31], and ANN [32] are used to compare with the proposed model for all the metrics. The total number of iterations considered for experimentation is 300 and the proposed model does not show much improvement after 200 iterations. Thus, the performance of all the models is considered for 200 iterations and the results are plotted as graphs.

The precision analysis of the proposed model for valence and arousal labels are presented in Fig. 4(a) and (b) respectively. The maximum precision value attained by the proposed model for valence is 92.56% and for arousal is 94.66% which is superior to the existing methods like SVM, ANN, CNN, and BiLSTM. The precision value attained by the SVM is 68.64% and 70.99% for the valence and arousal labels respectively which is 24% lesser than the proposed model. The precision value attained by the CNN model is 60.58% for the valence and 56.84% for arousal labels which is 32% and 38% lesser than the proposed model. Similarly, the BiLSTM model attains a 72.64% of precision value for the valence and 73.15% for arousal labels which is approximately 21% lesser than the proposed model. The precision value attained by the ANN model is 84.26% for the valence and 86.48% for arousal which is approximately 8% lesser than the proposed model.



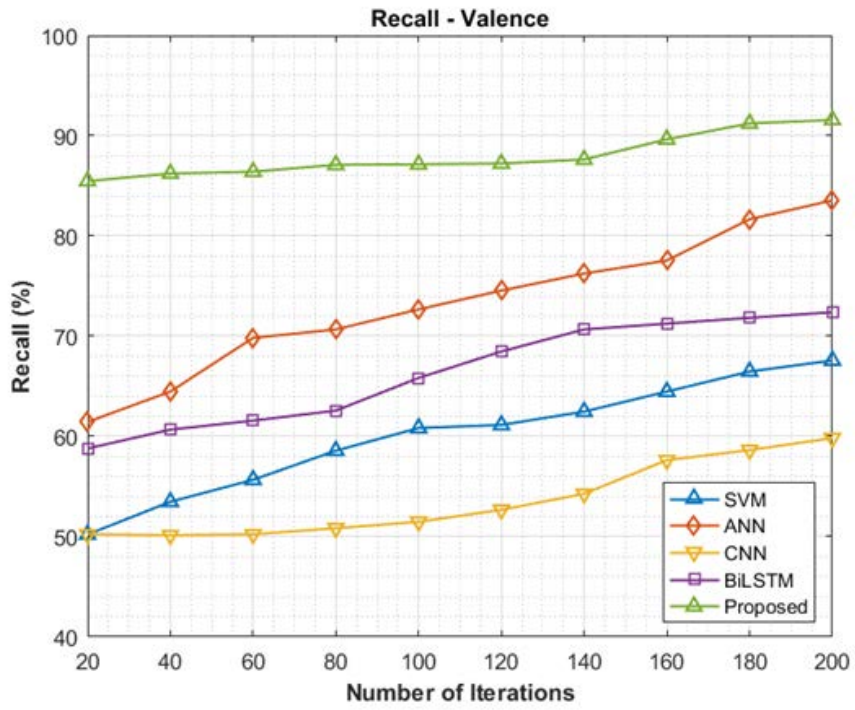
(a) Valence



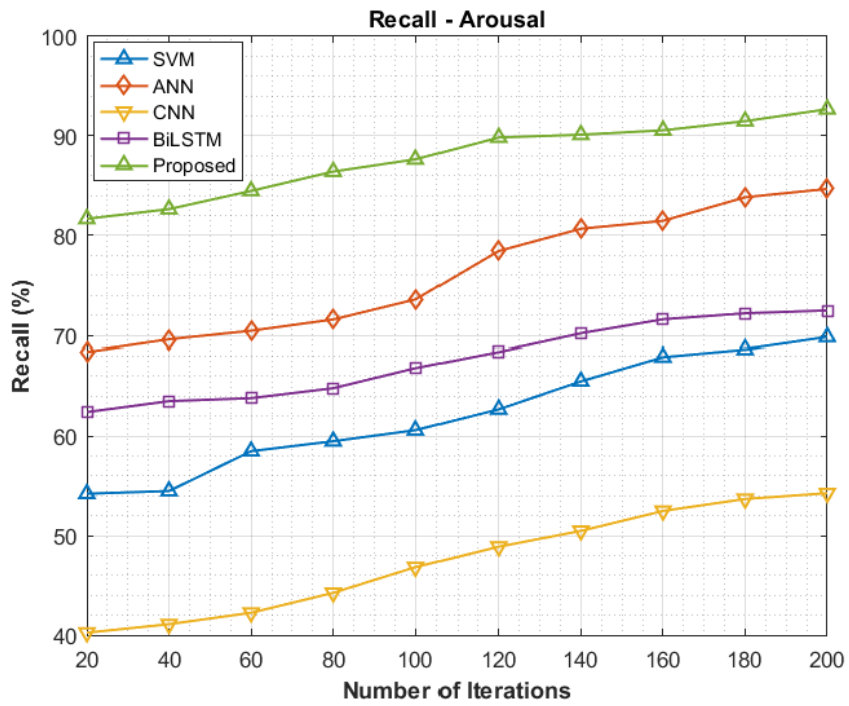
(b) Arousal

Fig. 4. Precision analysis (a) Valance (b) Arousal

The recall analysis of the proposed model for valance and arousal labels is presented in **Fig. 5(a)** and **(b)** respectively. The maximum recall attained by the proposed model for valance is 91.58% and for arousal is 92.68% which is superior to the existing methods like SVM, ANN, CNN, and BiLSTM. The recall value attained by the SVM is 67.58% and 69.84% for the valance and arousal labels respectively which is 23% lesser than the proposed model. The recall value attained by the CNN model is 59.84% for the valance and 54.26% for arousal labels which is 32% and 38% lesser than the proposed model. Similarly, the BiLSTM model attains a 72.4% of recall value for the valance and 72.48% for arousal labels which is approximately 20% lesser than the proposed model. The recall value attained by the ANN model is 83.56% for the valance and 84.68% for arousal which is approximately 8% lesser than the proposed model. The f1-score analysis of the proposed model for valance and arousal labels is presented in **Fig. 6(a)** and **(b)** respectively. The maximum f1-score attained by the proposed model for valance is 92.07% and for arousal is 93.66% which is superior to the existing methods like SVM, ANN, CNN, and BiLSTM. The f1-score attained by the SVM is 68.11% and 70.41% for the valance and arousal labels respectively which is 24% lesser than the proposed model. The f1-score attained by the CNN model is 60.21% for the valance and 55.52% for arousal labels which is 32% and 38% lesser than the proposed model. Similarly, the BiLSTM model attains 72.52% of the f1-score for the valance and 72.581% for arousal labels which is approximately 20% lesser than the proposed model. The f1-score attained by the ANN model is 83.91% for the valance and 85.57% for arousal which is approximately 8% lesser than the proposed model.

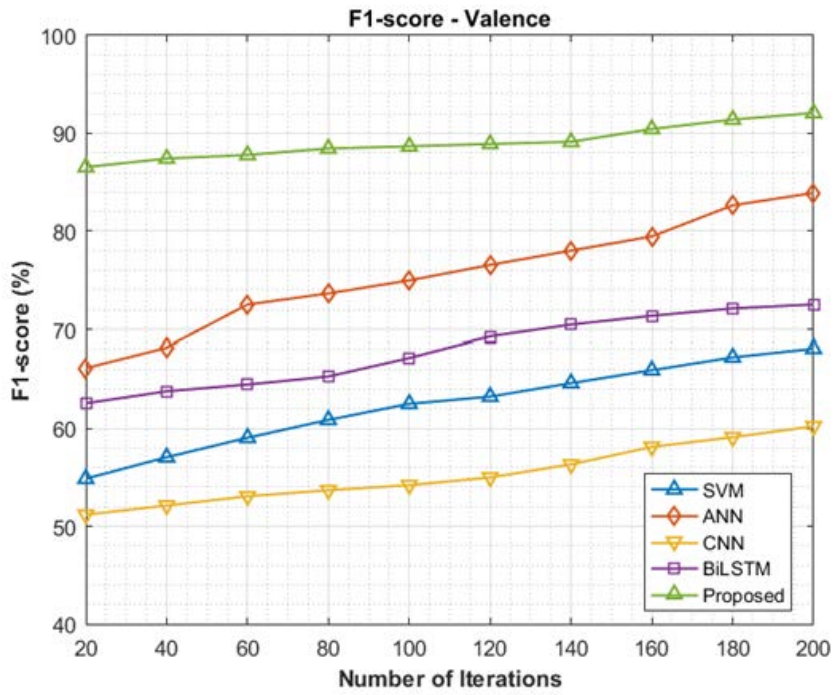


(a) Valence

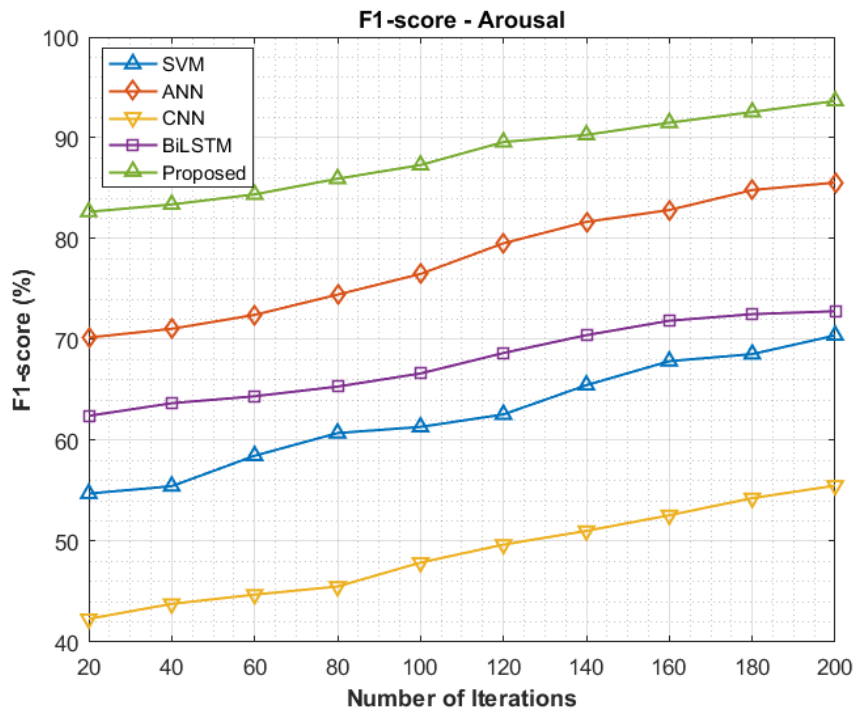


(b) Arousal

Fig. 5. Recall analysis (a) Valence (b) Arousal

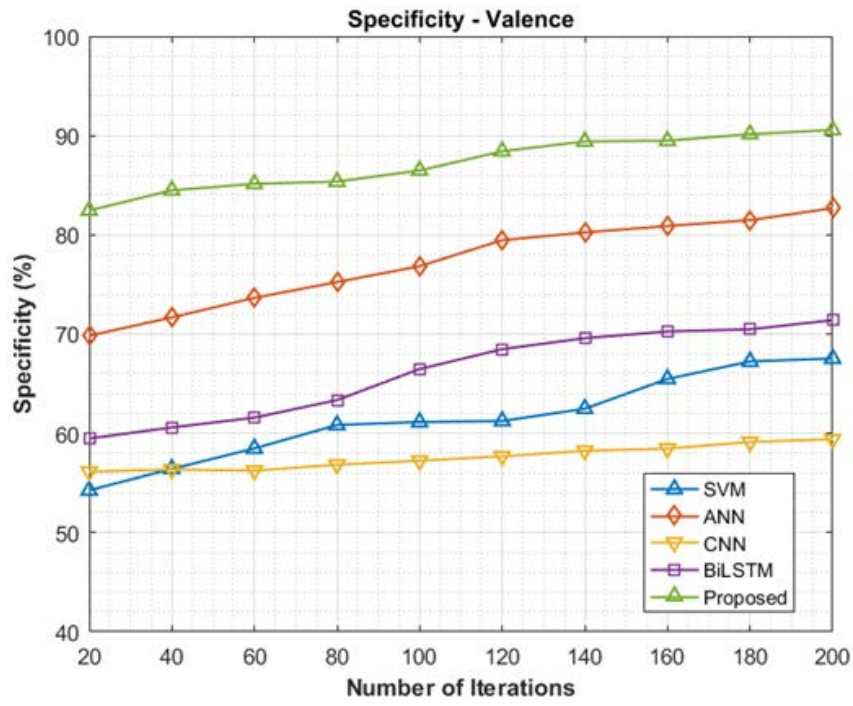


(a) Valance

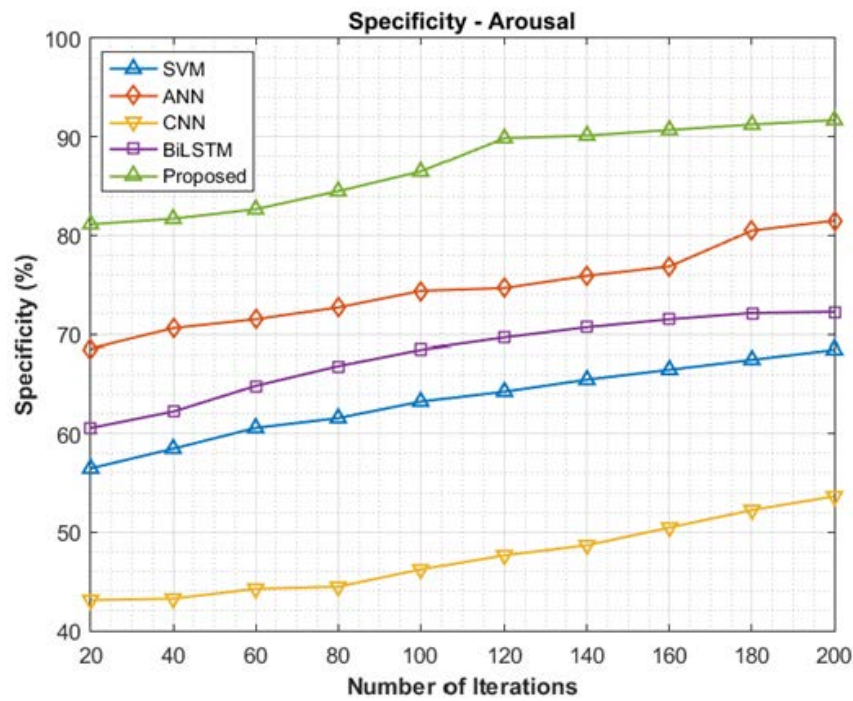


(b) Arousal

Fig. 6. F1-Score analysis (a) Valance (b) Arousal



(a) Valence



(b) Arousal

Fig. 7. Specificity analysis (a) Valence (b) Arousal

The specificity analysis of the proposed model for valance and arousal labels is presented in Fig. 7(a) and (b) respectively. The maximum specificity attained by the proposed model for valance is 90.56% and for arousal is 91.68% which is superior to the existing methods like SVM, ANN, CNN, and BiLSTM. The specificity attained by the SVM is 67.54% and 68.48% for the valance and arousal labels respectively which is 23% lesser than the proposed model. The specificity attained by the CNN model is 59.41% for the valance and 53.65% for arousal labels which is 31% and 38% lesser than the proposed model. Similarly, the BiLSTM model attains 71.4% of specificity for the valance and 72.24% for arousal labels which is approximately 19% lesser than the proposed model. The specificity attained by the ANN model is 82.68% for the valance and 81.48% for arousal which is approximately 9% lesser than the proposed model.

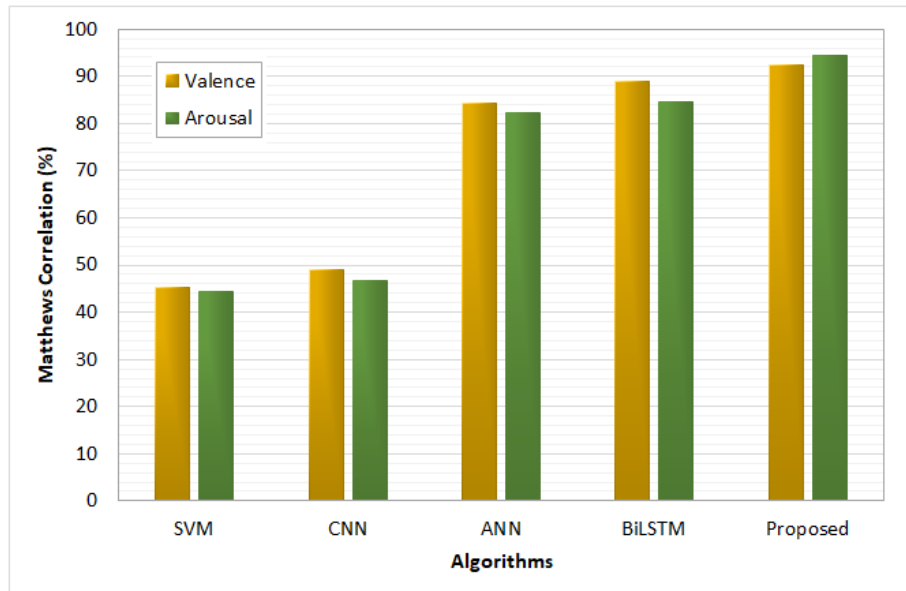


Fig. 8. Mathew correlation coefficient

Fig. 8 depicts the comparative analysis of the Mathew correlation coefficient attained by the proposed model and existing models for the valance and arousal labels. From the figure, it can be observed that the proposed model attains a maximum correlation coefficient of 92.41% for the valance and 94.48% for arousal labels which is higher than the existing methods. As the Mathew correlation coefficient is a statistical measure that produces high scores only all the elements in the confusion matrix produce better results. Thus, it is depicted as a bar chart in the above figure and it can be visible that the proposed model has a maximum score which indicates that the confusion elements are better compared to other methods.

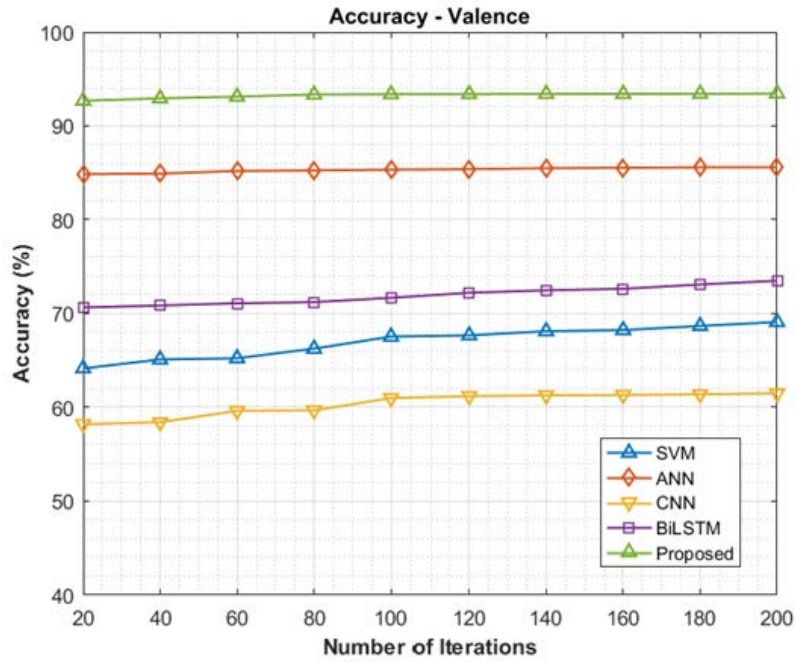


Fig. 9. Accuracy analysis (a) Valence

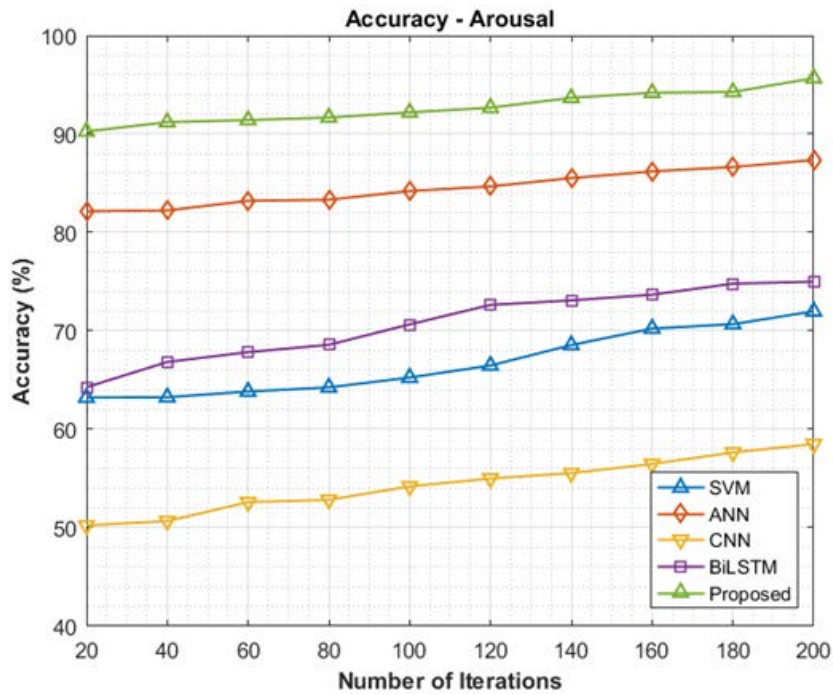


Fig. 9. Accuracy analysis (b) Arousal

The accuracy analysis of the proposed model for valence and arousal labels is presented in **Fig. 9(a)** and **(b)** respectively. The comparative analysis clearly presents the better performance of the proposed model. The maximum accuracy attained by the proposed model

for valance is 93.45% and for arousal is 95.68% which is much superior to the existing methods like SVM, ANN, CNN, and BiLSTM. The accuracy attained by the SVM is 69.1% and 71.99 for the valance and arousal labels respectively which is 24% lesser than the proposed model. In the proposed method SVM is used as a classifier, however due to optimal feature processing using deep learning model, the classification accuracy of SVM used in hybrid model is increased. since traditional method handles all the features which are not optimal the accuracy is reduced but in case of proposed model, due to optimal features, the classification accuracy is increased.

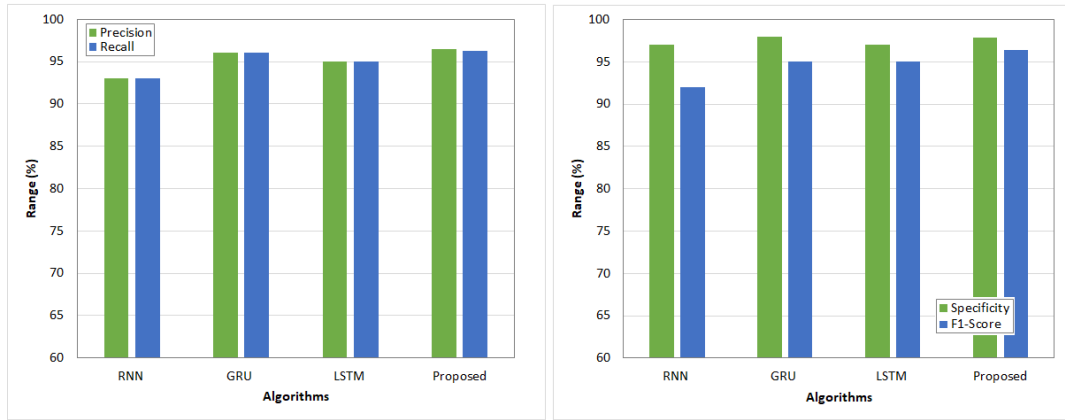
The accuracy attained by the CNN model is 61.5% for the valance and 58.5% for arousal labels which is 32% and 37% lesser than the proposed model. Similarly, the BiLSTM model attains 73.5% of accuracy for the valance and 75% for arousal labels which is approximately 20% lesser than the proposed model. The performance of ANN is better than other models, however, when it is compared with the proposed model, the accuracy attained by the ANN model is 8% lesser for both labels. Compared to the proposed model the performance of traditional CNN is low because the classifier used in the proposed model provides maximum classification accuracy than the traditional classifier used in the conventional CNN architecture.

Table 2. Comparative analysis with existing research works

Reference	Methods	Accuracy (%)	
		Valence	Arousal
Ning Zhuang. et.al., (2017)	Support vector machine (SVM)	69.10	71.99
Vaishali.et.al., (2022)	Bidirectional LSTM (BiLSTM)	73.5	75
Chen. et.al., (2019)	Hierarchical Bi-GRU	67.9	66.5
Pallavi.et.al., (2021)	Convolutional Neural Network	61.50	58.50
Hao Chao.et.al., (2020)	Stacked Autoencoder with LSTM	81.10	74.38
Rama Chaudhary.et.al., (2021)	Artificial Neural Network	85.60	87.36
Proposed	Hybrid CNN	93.45	95.68

To validate the proposed model performance further, few other research works like Bidirectional GRU [33], and stacked autoencoder with LSTM models are comparatively analyzed and the results are presented in Table 2. The result presents that the performance of the proposed model is much better than the existing methods for both valence and arousal labels.

Further to validate the proposed model performance for EEG brainwave dataset existing methods like Long Short-Term Memory (LSTM) Network, Recurrent Neural Network (RNN), and Gated Recurrent Unit (GRU) models are considered for analysis. The results of existing works are obtained from Kalpana et.al., 2022 [34] research work. The average values of proposed model and existing methods are considered for comparative analysis.



(a) Precision and Recall

(b) Specificity and F1-Score

Fig. 10. Performance metrics Comparative analysis

The parameters like precision, recall, specificity, and F1-score are comparatively analyzed for brain wave dataset using the proposed model and presented in **Fig. 10 (a) and (b)**. The maximum value attained by the proposed model for all the metrics is clearly visible from the **Fig. 10(a)**. The precision and recall attained by the proposed model is 96% whereas existing models like LSTM attained 95 % and RNN attained 93% which is lesser than the proposed model. Similarly, the maximum specificity and f1-score attained by the proposed model is clearly indicated in **Fig. 10(b)**. The specificity attained by the proposed model is 97.8% and F1-score is 96.3% whereas existing methods specificity and f1-score are comparatively lesser than the proposed model.

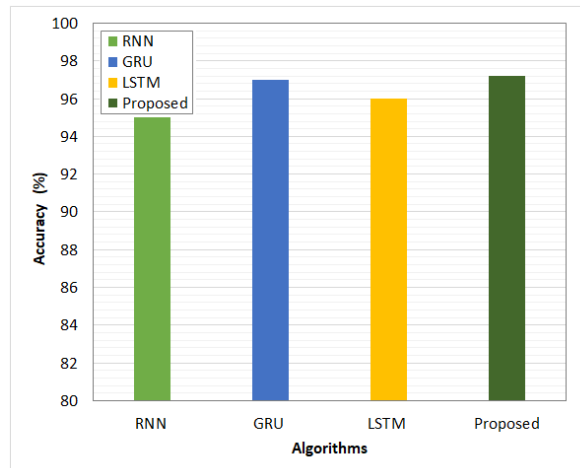
**Fig. 11.** Accuracy Comparative analysis

Fig. 11 depicts the accuracy comparative analysis of proposed model and existing models for brain wave dataset. The maximum accuracy attained by the proposed model is 97. Though GRU exhibits similar accuracy it is inaccurate due to poor feature handling characteristics. The overall comparative analysis of proposed model and existing methods for all the metrics is presented in **Table 3**. From the numerical results it can be observed that the performance of the proposed model is much better than the existing methods.

Table 3. Performance comparative analysis with existing methods for EEG Brainwave Dataset

S.No	Metrics/Methods	RNN	GRU	LSTM	Proposed
1	Accuracy	95	97	96	97
2	Precision	93	96	95	96
3	Recall	93	96	95	96
4	Specificity	97	98	97	98
5	F1-Score	92	95	95	96

From the above results, it is observed that the proposed hybrid model can effectively categorize the emotions by extracting the optimal features from the input. Using this hybrid approach, emotions and its related stress factors can be effectively classified.

5. Conclusion

A hybrid deep learning model for emotion classification based on stress level is presented in this research work. The hybrid deep learning model incorporates a convolutional neural network and support vector machine to attain better classification accuracy. The proposed model is experimented with a benchmark DEAP and Brainwave dataset and classifies the emotions. Performance metrics like precision, recall, f1-score, specificity, Mathew correlation coefficient, and accuracy are considered for analysis. Existing works like support vector machine, artificial neural network, RNN, GRU and LSTM models are used to compare the performance of the proposed model. From the experimental results, the performance of the proposed model is much better than the existing methods. Though the attained classification accuracy is better it can be improved further if deep feature extraction is performed through multiple deep learning algorithms and this could be the future scope of this research work to attain better classification accuracy.

Declaration

- Funding – The author did not receive support from any organization for the submitted work.
- Conflicts of Interest - The author has no relevant financial or non-financial interests to disclose.
- Ethics Approval – The paper is an original contribution of research and is not published elsewhere in any form or language.
- Consent Statement – All authors mentioned have contributed towards the research work, drafting of the paper as well as have given consent for publishing of this article.
- Availability of Data & Material – The author hereby declare that no specific data sets are utilized in the proposed work. The have also agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.
- Consent to publication – all authors listed above have consented to get their data and image published.
- Author's contribution – P.sivasankaran– Research proposal – construction of the work flow and model – Final Drafting – Dr.B.Gomathy – Survey of Existing works –

Improvisation of the proposed model; Dr.C.Venkatesh– Initial Drafting of the paper – Collection of datasets and choice of their suitability –Formulation of pseudocode

- Code Availability – Since, future works are based on the custom codes developed in this work, the code may not be available from the author.

References

- [1] Lan-lan Chen, Yu Zhao, and Jun-Zhong Zou, "Detecting driving stress in physiological signals based on multimodal feature analysis and kernel classifiers," *Expert Systems with Applications*, vol.85, pp.279-291, Nov 2017. [Article\(CrossRef Link\)](#)
- [2] Shaoling Jing, Xia Mao, and Lijiang Chen, "Prominence features: Effective emotional features for speech emotion recognition," *Digital Signal Processing*, vol.72, pp.216-231, Jan 2018. [Article\(CrossRef Link\)](#)
- [3] V Rekha, J Samuel Manoharan, R Hemalatha, and D Saravanan, "Deep Learning Models for Multiple Face Mask Detection under a Complex Big Data Environment," *Procedia Computer Science*, vol. 215, pp. 706 – 712, April 2022. [Article\(CrossRef Link\)](#)
- [4] E Manish Manohare, Rajasekar, and Manoranjan Parida, "Electroencephalography based classification of emotions associated with road traffic noise using Gradient boosting algorithm," *Applied Acoustics*, vol. 206, pp. 1-18, Jan 2023. [Article\(CrossRef Link\)](#)
- [5] Amna Khan, and Shahzad Rasool, "Game-induced emotion analysis using electroencephalography," *Computers in Biology and Medicine*, vol.145, pp. 1-19, June 2022. [Article \(CrossRef Link\)](#)
- [6] G S Shashi Kumar, Niranjana Sampathila, and Tanishq Tanmay, "Wavelet-based machine learning models for classification of human emotions using EEG signal," *Measurement: Sensors*, vol. 24, pp. 1-17, Dec 2022. [Article\(CrossRef Link\)](#)
- [7] Khairun Nisa Minhad, Sawal Hamid Md Ali, and Mamun Bin Ibne Reaz, "Happy-anger emotions classifications from electrocardiogram signal for automobile driving safety and awareness," *Journal of Transport & Health*, vol.7, pp.75-89, Dec 2017. [Article\(CrossRef Link\)](#)
- [8] Siao Zheng Bong, Khairunizam Wan, and Khairiyah Mohamad, "Implementation of wavelet packet transform and nonlinear analysis for emotion classification in stroke patient using brain signals," *Biomedical Signal Processing and Control*, vol.36, pp.102-112, July 2017. [Article\(CrossRef Link\)](#)
- [9] X. Xu and J. Sun, "Study on the influence of Alpha wave music on working memory based on EEG," *KSIIT Transactions on Internet and Information Systems*, vol. 16, no. 2, pp. 467-479, Feb 2022. [Article \(CrossRef Link\)](#)
- [10] Anushka Pradhan, and Subodh Srivastava, "Hierarchical extreme puzzle learning machine-based emotion recognition using multimodal physiological signals," *Biomedical Signal Processing and Control*, vol. 83, pp. 1-18, May 2023. [Article\(CrossRef Link\)](#)
- [11] Yogesh C.K, M. Hariharan, and Kemal Polat, "Hybrid BBO_PSO and higher order spectral features for emotion and stress recognition from natural speech," *Applied Soft Computing*, vol.56, pp. 217-232, July 2017. [Article\(CrossRef Link\)](#)
- [12] Kan Hong, Guodong Liu, and Sheng Hong, "Classification of the emotional stress and physical stress using signal magnification and canonical correlation analysis," *Pattern Recognition*, vol.77, pp.140-149, May 2018. [Article\(CrossRef Link\)](#)
- [13] Hemanta Kumar Palo, and Mihir Narayan Mohanty, "Wavelet-based feature combination for recognition of emotions," *Ain Shams Engineering Journal*, vol.9, no.4, pp.1799-1806, Dec 2018. [Article\(CrossRef Link\)](#)
- [14] Shui-Hua Wang, Preetha Phillips, and Yu-Dong Zhang, "Intelligent facial emotion recognition based on stationary wavelet entropy and Jaya algorithm," *Neurocomputing*, vol.272, pp.668-676, July 2018. [Article\(CrossRef Link\)](#)

- [15] R Martinez, E Irigoyen, and J Muguerza, "A real-time stress classification system based on arousal analysis of the nervous system by an F-state machine," *Computer Methods and Programs in Biomedicine*, vol.148, pp.81-90, Sep. 2017. [Article\(CrossRef Link\)](#)
- [16] Salazar-Ramirez. A, E. Irigoyen, and U. Zalabarria, "An enhanced fuzzy algorithm based on advanced signal processing for identification of stress," *Neurocomputing*, vol.271, pp. 48-57, Jan 2018. [Article\(CrossRef Link\)](#)
- [17] Raviraj Vishwambhar Darekar, and Ashwinikumar Panjabrao Dhande, "Emotion recognition from Marathi speech database using adaptive artificial neural network," *Biologically Inspired Cognitive Architectures*, vol.23, pp.35-42, Jan 2018. [Article\(CrossRef Link\)](#)
- [18] MaoSong Yan, Zhen Deng, BingWei He, ChengSheng Zou, Jie Wu, and ZhaoJu Zhu, "Emotion classification with multichannel physiological signals using hybrid feature and adaptive decision fusion," *Biomedical Signal Processing and Control*, vol.71, pp. 1-18, Jan 2022. [Article\(CrossRef Link\)](#)
- [19] Shuai Liu, Peng Gao b, Yating Li, Weina Fu, and Weiping Ding, "Multi-modal fusion network with complementarity and importance for emotion recognition," *Information Sciences*, vol. 619, pp. 679-694, Jan 2023. [Article\(CrossRef Link\)](#)
- [20] Wonju Seo, Namho Kim, Cheolsoo Park, and Sung-Min Park, "Deep Learning Approach for Detecting Work-Related Stress Using Multimodal Signals," *IEEE Sensors Journal*, vol. 22, no. 12, pp. 11892-11902, June 2022. [Article\(CrossRef Link\)](#)
- [21] J. He, D. Li, S. Bo and L. Yu, "Facial Action Unit Detection with Multilayer Fused Multi-Task and Multi-Label Deep Learning Network," *KSII Transactions on Internet and Information Systems*, vol. 13, no. 11, pp. 5546-5559, Nov 2019. [Article \(CrossRef Link\)](#)
- [22] Pritam Sarkar; and Ali Etemad, "Self-Supervised ECG Representation Learning for Emotion Recognition," *IEEE Transactions on Affective Computing*, vol. 13, no. 3, pp. 1541-1554, July 2022. [Article\(CrossRef Link\)](#)
- [23] Shuai Liu, Shichen Huang, Weina Fu and Jerry Chun-Wei Lin, "A descriptive human visual cognitive strategy using graph neural network for facial expression recognition," *International Journal of Machine Learning and Cybernetics*, pp.1-18, October 2022. [Article\(CrossRef Link\)](#)
- [24] Cuiting Xu, Chunchuan Yan, Mingzhe Jiang, Fayadh Alenezi, Adi Alhudhaif, Norah Alnaim, Kemal Polat, and Wanqing Wu, "A novel facial emotion recognition method for stress inference of facial nerve paralysis patients," *Expert Systems with Applications*, vol.197, pp. 1-16, July 2022. [Article\(CrossRef Link\)](#)
- [25] Francesco Di Luzio, Antonello Rosato, and Massimo Panella, "A randomized deep neural network for emotion recognition with landmarks detection," *Biomedical Signal Processing and Control*, vol.81, pp. 1-22, March 2023. [Article\(CrossRef Link\)](#)
- [26] Xuefen Lin, Jieli Chen, Weifeng Ma, Wei Tang, and Yuchen Wang, "EEG emotion recognition using improved graph neural network with channel selection," *Computer Methods and Programs in Biomedicine*, vol.231, pp.1-19, April 2023. [Article\(CrossRef Link\)](#)
- [27] <https://www.eecs.qmul.ac.uk/mmv/datasets/deap/download.html>
- [28] <https://www.kaggle.com/datasets/birdy654/eeg-brainwave-dataset-feeling-emotions>
- [29] Ning Zhuang, Ying Zeng, Li Tong, Chi Zhang, Hanming Zhang, and Bin Yan "Emotion Recognition from EEG Signals Using Multidimensional Information in EMD Domain," *BioMed Research International*, vol 2017. pp. 1-9, May 2017. [Article\(CrossRef Link\)](#)
- [30] Vaishali M. Joshi, and Rajesh B. Ghongade, "IDEA: Intellect database for emotion analysis using EEG signal," *Journal of King Saud University - Computer and Information Sciences*, vol.34, no.7, pp. 4433-4447, July 2022. [Article\(CrossRef Link\)](#)
- [31] Pallavi Pandey, and K. R. Seeja, "Subject independent emotion recognition system for people with facial deformity: an EEG based approach," *Journal of Ambient Intelligence and Humanized Computing*, vol.12, pp. 2311–2320, July 2021. [Article\(CrossRef Link\)](#)
- [32] Hao Chao, Liang Dong, Yongli Liu, and Baoyun Lu, "Improved Deep Feature Learning by Synchronization Measurements for Multi-Channel EEG Emotion Recognition," *Complexity*, vol. 2020. pp. 1-15, June 2020. [Article\(CrossRef Link\)](#)

- [33] J X Chen, D. M. Jiang, and Y. N. Zhang, "A Hierarchical Bidirectional GRU Model with Attention for EEG-Based Emotion Classification," *IEEE Access*, vol. 7, pp. 118530-118540, Aug 2019. [Article\(CrossRef Link\)](#)
- [34] Kalpana Chowdary.M, J. Anitha, Jude Hemanth.D, "Emotion Recognition from EEG Signals Using Recurrent Neural Networks," *Electronics*, vol.11, pp. 1-20, 2022. [Article\(CrossRef Link\)](#)



Mr.P.Sivasankaran is currently working as a Assistant Professor of Department of Electronics and Communication Engineering in Sengunthar Engineering College and he is doing Ph.D. in Anna University in the field of Artificial Intelligence. He has 13 years of teaching experience. He is having life membership at Institution of Engineers India (IEI).



Dr.B.Gomathy is currently working as professor of Computer Science and Business Systems department at Dr NGP Institute of Technology, Coimbatore, Tamil Nadu, India. She received his Ph.D. degree in computer Science in the year 2014. His area of interest includes Data Analytics, Security, Image Processing and wireless networks. Presently she is guiding research scholars in the field of Big Data and Wireless networks. She is having a lifetime membership of Indian Society for Technical Education (ISTE) and IAENG.



Dr.C.Venkatesh is graduated With B.E degree in Electronics and Communication Engineering in the year 1988,M.E in Applied Electronics in 1989 and Ph.D.in the year 2007 from Jawaharlal Nehru Technological University. Currently he is working as Chief Executive Officer at Builders engineering college, Kangayam, Tirupur. He has 30 years of teaching experiences and has published more than 145 papers in the leading journals and conferences. His areas of interest are Wireless Networks, Image Processing, and Machine Learning. He is an Active member of IEEE, ISTE, IE and IETE.