A NOVEL MULTIMODAL APPROACH FOR EMOTION RECOGNITION IN DIFFERENTLY ABLED INDIVIDUALS

JAYACHITRA V P[1], RAJESH G[1], VIGNESH SIVA P[1], ABINESH V[1]

[a]*Department of Computer Technology, Anna University, MIT Campus, Chennai, Tamilnadu, India*

Abstract

Emotion detection refers to the process of identifying and interpreting emotional states in individuals. This can involve analysing a variety of cues such as facial expressions, vocal intonation, physiological responses, and behavioural patterns. Emotion detection finds application in various fields such as healthcare, education, marketing, and customer service. There are many techniques for emotion analysis like facial expression, voice analysis, brain imaging, behaviour analysis, etc. Detecting emotions in differently abled people like paralyzed people can be challenging for several reasons, including limited or absent motor functions, communication difficulties, physiological changes, and psychological factors. Hence general methods of emotion detection cannot be used for paralyzed people. For emotion detection, this proposed model of using Electroencephalogram (EEG) and Electrocardiogram (ECG) signals will be more efficient and reliable than the existing methods of using facial expressions, Galvanic Skin Response (GSR), Heart Rate Variability (HRV) and so on.

Keywords: EEG, ECG, emotion recognition, deep learning, hyper parameter tuning

1. Introduction

Emotions encompass intricate psychological and physiological states that emerge as a reaction to stimuli originating from either within oneself or from the surrounding environment. They play a significant role in human experience, influencing thoughts, behaviour, and overall well-being. Some common emotions include happiness, sadness, fear, anger, surprise and so on. Emotions are subjective and can vary in intensity and duration depending on individual experiences, cultural influences, and personal interpretations. Recognition the emotions of humans has various applications across different fields like mental-health monitoring, education, training, entertainment, etc. Emotion recognition involves the task of recognizing and comprehending the emotions conveyed by individuals through their facial expressions, gestures, vocal intonation, or other behavioural cues. There are physiological and non-physiological signals in humans to detect their emotions.

1.1. PHYSIOLOGICAL SIGNAL

Physiological signals refer to measurable and quantifiable signals that are derived from the functioning of the human body. These signals provide valuable information about the physiological state and processes occurring within an individual. Electrocardiogram (ECG) signals, which measure the electrical activity of the heart, can be utilized in emotion detection through the analysis of specific features and patterns present in the signal. Emotions elicit physiological responses within the body, causing variations in heart rate, heart rate variability (HRV), and other ECG parameters. Electroencephalogram (EEG) signals are used in emotion recognition by analysing the patterns of brain activity associated with different emotional states. EEG signals are acquired by positioning electrodes on the scalp, enabling the recording of electrical activity originating from the brain. These signals offer valuable insights into the neural mechanisms associated with emotions, thereby facilitating the detection and classification of various emotional states. Electromyography (EMG) signals can be utilized in emotion detection involves the examination of the electrical activity generated by facial muscles, which corresponds to different emotional expressions. By analyzing these signals, it becomes possible to identify and interpret the specific muscle activations associated with various emotions. EMG signals provide insights into the muscle contractions and movements that occur during emotional responses. Galvanic Skin Response (GSR), also referred to as Electrodermal Activity (EDA) or Skin Conductance, is a physiological signal that quantifies the electrical conductance of the skin. It primarily reflects variations in sweat gland activity. GSR is commonly used as a measure of sympathetic nervous system arousal, and it can provide insights into emotional states, stress levels, and other physiological responses.

1.2. NON-PHYSIOLOGICAL SIGNAL

Non-physiological signals from humans are the data signals that are not directly derived from physiological processes within the human body. These signals are often associated with human behaviour, communication, or interactions, and they can provide valuable insights into cognitive, social, or psychological aspects of individuals. Facial expression serves as a natural and instinctive reflection of an individual's mental state, encompassing a wealth of emotional information. It stands as a crucial form of interpersonal communication, carrying immense importance. Consequently, facial expressions hold a significant role in the process of emotion recognition due to their strong connection to the underlying emotional state of an individual. The analysis of facial expressions can provide valuable cues for identifying and categorizing different emotions. Using hand gestures for emotion detection involves analysing and interpreting the movements and gestures made by an individual's hands to infer their emotional state. While hand gestures are not direct indicators of emotions, they can provide valuable contextual cues that, when combined with other modalities, contribute to emotion recognition. Speech analysis is a widely used modality for emotion detection due to the various vocal cues present in an individual's speech that can convey their emotional state. By examining speech features, such as pitch, intensity, speech rate, and spectral characteristics, it is possible to infer emotions expressed through speech.

1.3. EMOTIONS IN PARALYZED PEOPLE

Paralysis refers to the loss or impairment of voluntary movement in a part or parts of the body. It can result in a loss of sensation, muscle weakness, or a complete inability to move the affected body parts. Emotion recognition in paralyzed individuals can be challenging due to the physical limitations they may have. Paralysis in some people may affect the whole body or partially, based on the severity. People with paralysis in vocal cord may not be able to control their vocal cord movements and cannot communicated with their voice. These problems result in limited body movement, facial expression and speech impairment. Thus normal method of using non-physiological signals like facial expression, hand gestures and speech cannot be used efficiently for paralyzed people.

2. RELATED WORKS

2.1. ECG SIGNALS

Wonju Seo et al. in [20] proposed deep learning model exhibits high accuracy in detecting work-related stress by leveraging multimodal signals. To ensure optimal performance, the datasets underwent pre-processing, and specifically, 10-second segments of Electrocardiogram (ECG) and Respiration (RESP) signals, along with a sequence of facial features, were inputted into the deep neural network. This integrated approach is used to capture and analyze multiple modalities of information for enhanced stress detection. This paper only focused on facial expression, which cannot be obtained from people with nervous disorder and people with paralysis. Also, facial expressions can be faked and thus relying only on facial expression is not efficient. Minh Pham et al. in [15] proposed a negative emotion management system has been developed, capable of identifying negative emotions through Electrocardiogram (ECG) signals and facilitating emotion regulation through a robot assistant. This system holds promising potential for mitigating health risks associated with negative emotions. The ECG signals are obtained from publicly available datasets such as RECOLA and DECAF. Recurrence Quantitative Analysis is employed to extract relevant features from the ECG signals, enabling accurate classification of emotions. The paper only focused on ECG data for recognizing emotions, which can be misleading for people with heart problems like abnormal heart rhythms and blocks in the heart valve. Amita Dessa et al. in [14] had used physiological signals such as ECG, GSR from AMIGOS database. They hey have eliminated the baseline drift using the baseline wandering pathfinding algorithm, extracted the peaks in the GSR signal, and inter beat measurements of the ECG signals. GSR signal is difficult to obtain from people with Hidradenitis disorder and Miliaria rubra disorder. JIANQIANG XU et al. [6] multi-channel convolutional autoencoder neural network was employed to extract features from both electrocardiogram (ECG) data and emotional text. This neural network architecture facilitated the extraction of relevant and discriminative features from both modalities, enabling the integration of physiological and textual information for emotion-related tasks or analyses. The extracted features were combined to detect the emotions. The paper only focused on ECG data for recognizing emotions, which can be misleading for people with heart problems like abnormal heart rhythms and blocks in the heart valve. Hany Ferdinando et al. in [1] had shown that ECG signals and HRV analysis can be used for emotion recognition. Numerous studies have explored the utilization of specific Electrocardiogram (ECG) features for the purpose of emotion recognition. Prominent examples include the examination of features like the amplitude of the R-wave peak and the width of the QRS complex. These features have been investigated extensively as potential indicators of emotional states, contributing to the development of robust and accurate emotion recognition models based on ECG data. Xia et al. (2019) used ECG-derived features such as the R-wave amplitude and slope to classify emotions. HRV signals lack specificity and varies for every individual. HRV can be influenced by stress, caffeine and medications. Measuring HRV requires specialized equipment and is time consuming and invasive. Etemad Sarkar et al. in [23] introduce an ECG-based emotion recognition system using self-supervised learning with two networks. The first network detects signal transformations from unlabeled data, and its weights are transferred to the emotion recognition network, outperforming fully-supervised models and achieving state-of-the-art results on SWELL and AMIGOS datasets with significantly less data. Mahmood Khan et al. in [24] emphasizes the importance of emotional states' impact on human well-being and health. Using portable ECG devices, it investigates changes in ECG waveforms corresponding to emotions like happiness, sadness, and anger. The findings align with existing research on affect recognition, indicating the potential of ECG data for understanding emotional changes. Trivizakis Giannakakis ei al. in [25] addressed stress recognition, a complex issue with subjective experiences but shared characteristics. Deep Learning (DL) techniques applied to electrocardiography (ECG) data unveil intricate patterns, achieving high classification accuracy (up to 99.1%) and consistency, surpassing traditional methods in 6-fold cross-validation. Etemad Sarkar et al. in [26] employs a self-supervised multi-task learning framework for ECG-based emotion recognition. It involves two stages: learning ECG representations via signal transformation recognition on unlabeled data, and transferring these weights to an emotion

Table 1: Summary of previous methodologies on Emotion Recognition

Literature/Year	**Modalities**	**Findings**	**Methods/Features**	**Results**	**Challenges**
[20] (2020)	ECG, RESP, Facial	Detects work-related stress	Deep learning model	Optimal performance	Facial expression limitations
[15] (2021)	ECG	Identifies negative emotions	Emotion management system	Potential health risk mitigation	Limited focus on ECG
[14] (2021)	ECG, GSR	Utilizes physiological signals	Baseline drift elimination	Eliminates baseline drift	Challenges with disorders
[6] (2020)	ECG, Emotional Text	Extracts features for emotion	Convolutional autoencoder network	Relevant features extraction	Focus on ECG data
[1] (2020)	ECG, HRV	Uses ECG features for emotion	Exploration of R-wave amplitude	Robust emotion recognition models	Lack of specificity in HRV
[18] (2022)	EEG	Multimodal Recognition	EEG, Facial Fusion	Accurate classification	Limited to deaf individuals
[7] (2021)	EEG	EEG Transfer Learning	Deep transfer models	Higher accuracy in preference	Noise modification, misleading
[4] (2019)	EEG	ECG to EEG Transformation	Fourier transform, features	Analysis of frequency components	Noise modification, misleading
[5] (2020)	EEG	Interpretable Emotion Recognition	Autoencoder, Voting Strategies	Constructing novel activation curves	Applicable for second half
[10] (2020)	EEG	Two-stage, Three-stage Classification	LSTM, CNN Fusion	High accuracy, Kappa coefficients	Limitations in capturing features
[17] (2022)	Speech	CoSTGA Model, IEMOCAP dataset	Acoustic, lexical features	Enhanced emotion classification	Vocal cord problems
[16] (2021)	Speech	Efficient model, Emotion embedding	Autoencoder, SMILE toolkit	Relevant features for emotion analysis	Vocal cord problems, irregular rhythms
[2] (2020)	Speech	Synthesized facial expressions, Inferential mechanisms	Blended facial expressions, Interactive scenario	Socio-affective factors impact	Varying interpretations among individuals
[16] (2021)	Speech	Efficient model, Emotion embedding	Deep emotion features extraction	Relevant features for emotion	Vocal cord problems

recognition network, resulting in significant performance improvements and state-of-the-art results across four datasets.

2.2. EEG SIGNALS

Dahua Li et al. in [18] proposed a study focusing on subjects with hearing impairment, a multimodal emotion recognition framework was devised to classify four distinct emotions. This framework integrated Electroencephalogram (EEG) topographic maps and facial expressions by employing a multichannel fusion method. The fusion technique facilitated the combination of information from both modalities to enhance the emotion recognition process. To accomplish emotion classification, a deep learning model was created by incorporating ResNet and a Convolutional Block Attention Module. This model provided a powerful approach to accurately classify emotions within the study population. This paper only focused on deaf people and used facial expression for emotion recognition, which can be faked. It may or may not be possible to obtain facial expressions from people who are partially or completely paralyzed. MASHAEL S. ALDAYEL et al. in [7] investigated the use of EEG-based transfer learning and proposed deep transfer learning models. The aim was to transfer knowledge gained from emotion recognition tasks to preference recognition tasks, thereby improving the accuracy of classification predictions. This approach showed promising results in achieving higher classification accuracy in preference recognition tasks based on EEG data. The EEG signal is obtained from the pre-processed DEAP dataset, and the power spectral and valence features are extracted. The deep neural network model and other conventional classification algorithms were developed to recognize the emotions. The background noise will modify the EEG signal while collecting it. EEG signals can be misleading for the people with brain dysfunction and sleep disorder. Sejal Chopra et al. in [4], ECG data was collected using the Neurosky Mindwave headset and the data was then transformed into the frequency domain using Fourier transform, enabling the analysis of different frequency components. Various feature extraction techniques were subsequently applied to extract relevant information from the ECG signals. The background noise will modify the EEG signal while collecting it. EEG signals can be misleading for the people with brain dysfunction and sleep disorder. CHUNMEI QING et al. in [5], had used Machine learning and EEG signals to create a new approach named interpretable emotion recognition. Autoencoder was used to process the different features which are extracted from EEG signals. The soft Voting Strategies was used as a classifier. The novel activation curves of emotions are constructed based on correlation coefficients and entropy coefficients; after that we can get weight coefficients which were used to get the final emotional table. In this method, only the second half of the EEG dataset was able to predict the emotions. This method is not applicable for the first half of the EEG dataset. SOBHAN SHEYKHIVAND et al. in [10], introduced a novel two-stage and three-stage classification model for emotions, utilizing signals extracted from Electroencephalogram (EEG) data. In the two-stage classification, emotions were categorized as either negative or positive, while the three-stage classification involved the additional category of neutral emotions. To enhance stability and reduce oscillation, a fusion of Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) networks was employed. The proposed algorithm achieved high accuracy and Kappa coefficients, with 97.42% and 96.78% accuracy and Kappa coefficients of 0.94 and 0.93, respectively, when evaluated using 12 active channels. It should be noted that while CNN was used for feature extraction from EEG signals, it might not capture all relevant features necessary for accurate emotion recognition. Authors used a relatively small dataset, which may limit the generalizability of the results. Paweł Tarnowski et al. in [3] had shown that both GSR and EEG signals can be used for emotion recognition. Few studies have combined these two signals for emotion recognition Lee et al. (2019) found that a combination of GSR and EEG signals led to higher classification accuracy than either signal alone. Other studies have used GSR or EEG signals alone to classify emotions based on physiological responses to music or visual stimuli. GSR is non-specific and was influenced by external temperature, humidity and electric interference. GSR cannot capture rapid changes in emotional arousal's can habituate and cannot be used for repeated stimuli. Wibawa Oktavia et al. in [27] explores emotion recognition via EEG signals, specifically differentiating between happy and sad emotions. Utilizing time domain features and a Naïve Bayes classifier, it achieved an 87.5% accuracy, with the combination of alpha and beta frequency bands showing better results. Chopda Gupta et al. in [28] addresses the challenge of recognizing emotions from EEG signals, using flexible analytic wavelet transform (FAWT) to decompose signals and extract features via information potential (IP). The proposed method outperforms existing techniques and offers channel-specific subject classification across different databases. X Du et al. in [29] introduces ATDD-LSTM, a deep learning model for EEG-based emotion recognition, addressing limitations of hand-crafted EEG features. ATDD-LSTM uses Long Short-Term Memory (LSTM) to capture nonlinear relations among EEG signals from different electrodes and achieves superior performance across various emotion recognition evaluations on three public EEG databases. J Cheng et al. in [30] presents a multi-channel EEG-based emotion recognition method using deep forest, addressing DNNs' hyperparameter and data requirements. The approach achieves high accuracy on DEAP and DREAMER databases, outperforming state-of-the-art methods with accuracies reaching up to 97.69% and 97.53% for valence and arousal on DEAP, and 89.03%, 90.41%, and 89.89% for valence, arousal, and dominance on DREAMER.

2.3. FACIAL EXPRESSION AND SPEECH

SAMUEL KAKUBA et al. in [17] proposed a deep learning-based model called the concurrent spatial-temporal and grammatical (CoSTGA) model. This model was developed using the interactive emotional dyadic motion capture (IEMOCAP) dataset. Acoustic features and lexical features were extracted from the dataset to capture relevant information for emotion analysis. The CoSTGA model incorporated spatial-temporal information and linguistic patterns to enhance the understanding and classification of emotions in the dataset. The CoSTGA model is trained using these features to make the emotion classification. This paper focused on using speech for emotion recognition, which cannot be obtained from people with problems in vocal cord. Sara Casaccia et al. in [13] proposed a technique that utilizes the facial movements of different people using EMG signals. The direct measurement of an EMG signal involves burdensome procedures. So, the authors have approached an alternative measurement called laser Doppler vibrometry to obtain facial movements for emotion classification. EMG signals are difficult to obtain from paralyzed and overweight people. CHENGHAO ZHANG et al. in [16] had proposed an efficient model to extract deep emotion features for speech emotion recognition. The proposed model includes an emotion embedding path, which enables the identification of features that are highly correlated with human emotion. These features are obtained using an autoencoder, while the acoustic features are concatenated using the open SMILE toolkit. By incorporating this approach, the model aims to highlight and leverage the most relevant features for emotion analysis and recognition. This method cannot be used for people with vocal cord problems and people with irregular vocal cord rhythms. C. Mumenthaler et al. in [2] utilized synthesized facial expressions to investigate how socio-affective inferential mechanisms affect the recognition of social emotions. Participants were presented with blended facial expressions on a target avatar while a contextual avatar displayed an emotion or remained neutral. Manipulations of head and gaze movements created an interactive scenario. The research aimed to explore how individuals interpret ambiguous facial expressions, considering factors such as past experiences, cultural background, personality, and other individual characteristics. The findings shed light on the intricate nature of emotion recognition and the impact of socio-affective factors in this process. Thus, perception of ambiguous facial expressions may vary among individuals, and it is not possible to make a generalization that applies to everyone. Xiao Liu et al. in [9] had shown that facial expression recognition based on geometric features can achieve high accuracy, but traditional machine learning methods often suffer from the problems of overfitting and low generalization ability. To address these issues, recent studies have used evolutionary algorithms, like genetic algorithms, to optimize hyperparameters for facial expression recognition. Support Vector Machines (SVMs) are favored due to their ability to handle high-dimensional data and capture nonlinear relationships. By applying evolutionary algorithms, researchers aim to improve accuracy and performance of facial expression recognition systems. The authors did not address the issue of model efficiency or discuss ways to optimize the algorithm for real-time implementation, which is an important consideration for many potential applications of facial expression recognition technology. K. N. V Satyanarayana1 et al. in [19] noted that KNN has the advantage of being simple to implement and computationally efficient, but it may not perform well on high-dimensional datasets. Recurrent Neural Networks (RNNs) are a type of neural network designed specifically for sequential data processing, making them useful for emotion recognition tasks involving time-series data, such as speech or physiological signals. The authors cited several previous studies that have used KNN and RNN models for emotion recognition with varying degrees of success, and noted that there is still a need to improve the accuracy of these models. Dong Cai et al. in [31] presents a multimodal emotion recognition approach combining speech and facial expression features, enhancing human-computer interaction. Using CNN, LSTM, and multiple small-scale kernel convolution blocks, the model achieves a 10.05% accuracy improvement over single-modal speech recognition and an 11.27% improvement over single-modal facial expression recognition on the IEMOCAP dataset. Yang Song et al. in [32] presents a method for converting sign language integrated with facial expressions into emotional speech, aiding communication with speech disorders. Objective tests show high recognition rates for both static sign language and facial expressions, while subjective evaluation confirms the emotional expressiveness of the synthesized speech. Palaniswamy Keshari et al. in [33] recognize the importance of automatic emotion recognition using computer vision in various real-world applications. Humans inherently rely on both facial expressions and body language to understand emotional behavior. Previous studies primarily focused on single modality, but our objective is to efficiently combine emotions detected from both facial expressions and upper body poses in images, aiming to enhance the accuracy of emotion recognition in a bimodal approach. Yadav Pathak et al. in [34] recognize the importance of real-time facial expression recognition and emotion prediction for seamless human-machine interactions. We've developed a Python model using supervised learning, which learns to identify expressions and forecast emotions in real-time through webcam input, bringing us closer to creating emotionally intelligent AI systems.

2.4. CNN AND GENETIC ALGORITHM

TING-WEI SUN et al. in [11] had proposed deep learning model which combines with a gender information in speech to recognize the emotion using speech. The proposed deep learning model is Residual Convolutional Neural Network (R-CNN) and gender can be classified using SVM model. Gender can be obtained by the shape of vocal cord trace features. This paper focused on speech dataset, which could be inappropriate for people with vocal problems. Obtaining gender details solely from vocal cord features could be misleading, since people could mimic voices of the opposite gender. HONGLI ZHANG et al. in [12] utilized deep autoencoders and Convolutional Neural Networks (CNN) to improve emotion recognition. By combining these techniques, the performance of the classification system was enhanced. However, the proposed method of classifying video clips was noted to lack subjectivity, suggesting the need for further improvement and consideration of subjective factors in emotion recognition. Sabrina Begaj et al. in [8] trained CNN models from scratch or fine-tuned on pre-trained models to classify emotions based on facial expressions. Data augmentation techniques are employed to increase training dataset size in emotion recognition using CNNs. Integration of additional information, like audio or contextual

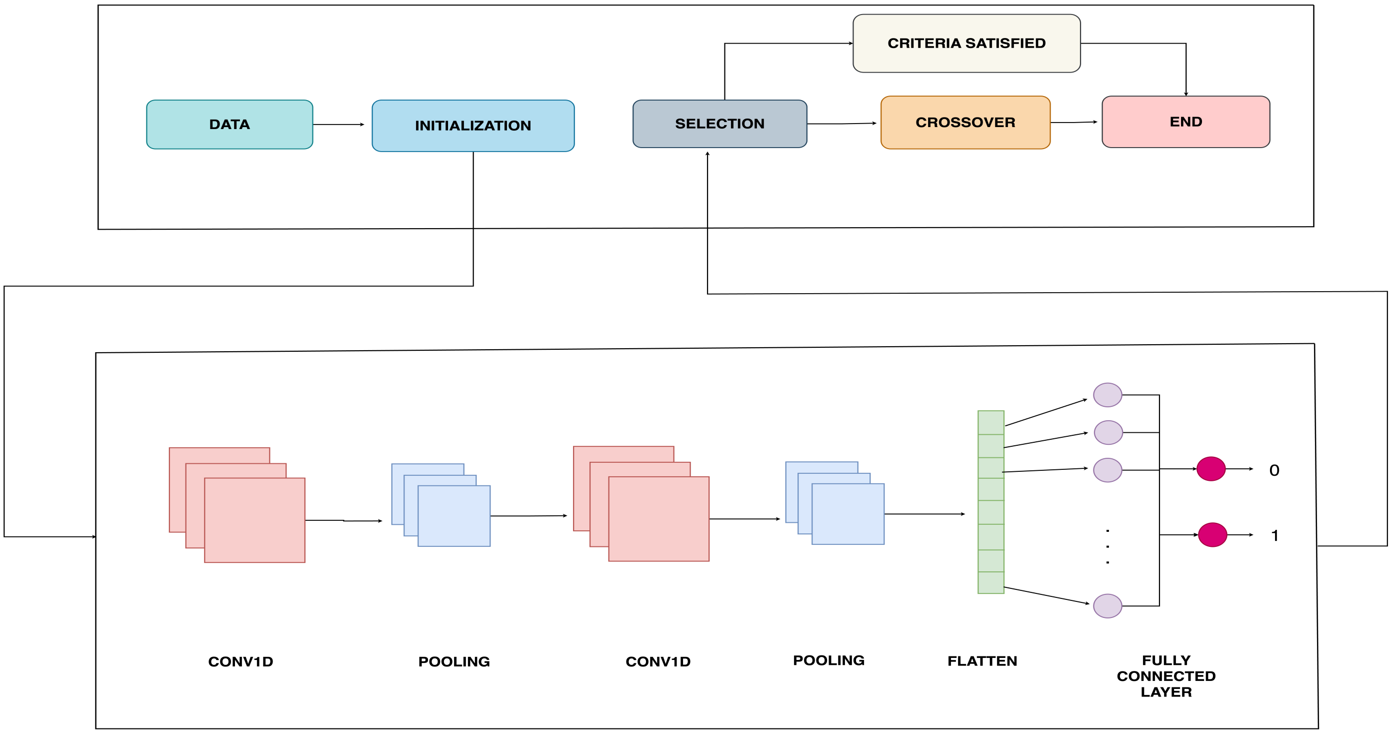

Figure 1: CNN AND GENETIC ALGORITHM

data, enhances CNN model accuracy and robustness. Ongoing research focuses on developing efficient CNN architectures and incorporating other deep learning techniques. However, the model's generalizability to different datasets and real-world scenarios with diverse facial expressions and emotions may be limited. There is a need for comprehensive evaluation metrics in this field. In [21], the genetic algorithm to optimize the nonlinear pulse repetition interval (PRI) variation strategy for staggered synthetic aperture radar (SAR) imaging. The results demonstrated the effectiveness of the knowledge-guided genetic algorithm approach, resulting in high-quality SAR images and reduced computational time compared to other optimization methods. This research highlights the potential of genetic algorithms in enhancing SAR imaging techniques through efficient parameter optimization. Ramachandran et al. in [22] proposed a genetic algorithm-based approach to optimize the sensing matrix for compressive sensing in image classification. The proposed method outperformed other state-of-the-art techniques, achieving higher accuracy with fewer measurements. This enhanced efficiency and practicality make the approach suitable for real-world applications. CHETOUANI GUETARI et al. in [35] acknowledges the persistent technological challenges, including occlusion and feature similarity between emotions. To address these, we employ deep learning techniques like bilinear pooling (B-CNN) and Fusion Feature Net (F-CNN) to enhance precision and performance, presenting a promising step forward in this field. Mall Rajak et al. in [36] recognize text's limitations in capturing nuanced elements like humor and cadence. Departing from traditional discrete emotion classification, our approach, placing emotions in quadrants on the valence-arousal axis, attained a promising 76.2% accuracy on the RAVDESS dataset, outperforming previous models relying on discrete categories. Lu Peng et al. in [37] recognize the complexity of the task and propose an efficient neural network framework. Our approach, utilizing multi-scale convolutional layers (MSCNN), statistical pooling units (SPU), and attention mechanisms, outperforms previous state-of-the-art methods on the IEMOCAP dataset, demonstrating substantial gains in weighted accuracy (WA) and unweighted accuracy (UA) under the ASR setting for four emotion categories. Zafeiriou Kollias et al. in [38] present a novel CNN-RNN approach for dimensional emotion recognition from visual data, focusing on the OMG-Emotion dataset. Leveraging pre-training on relevant emotion databases and multi-level feature extraction, our method outperforms state-of-the-art techniques, ranking prominently in the OMG-Emotion Challenge, and significantly enhancing arousal estimation by integrating low-level and high-level features. Sewisy Mohammed et al. in [39] address the limitations of the Black Hole (BH) algorithm by introducing a novel hybrid approach, GA-BH, which combines BH for local search and Genetic Algorithm (GA) for global search. Extensive evaluations confirm that our hybrid algorithm surpasses both BH and GA, enhancing optimization performance across various problem domains. Machavaram Gupta et al. in [40] focus on the optimization of induction motors, crucial components in Electric Vehicles (EVs), using Genetic Algorithm (GA). Our approach enhances motor efficiency by 1.02% and reduces weight by 1.11 kg, delivering improved performance over conventionally designed motors while adhering to design constraints.

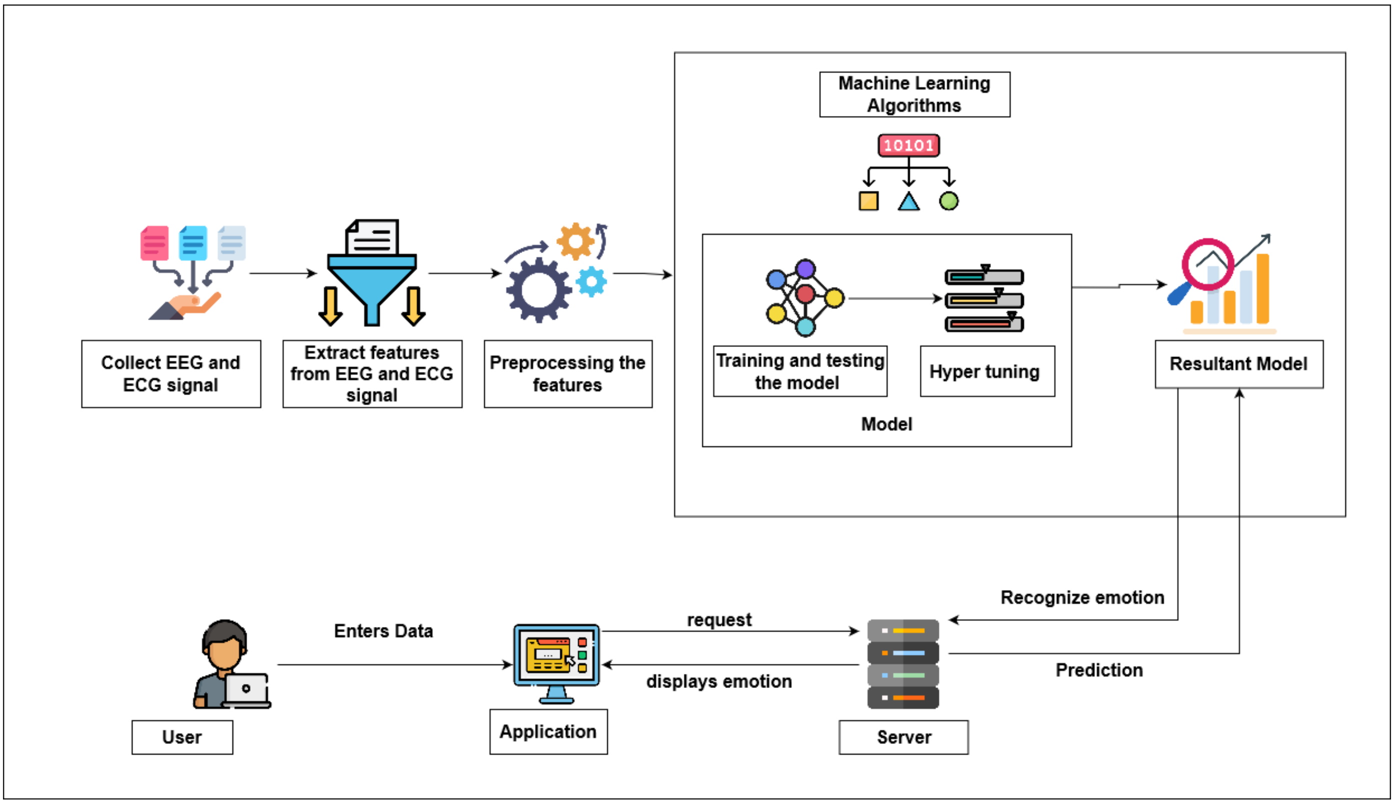

Figure 2: Proposed System Workflow

3. PROPOSED SYSTEM

The proposed system workflow shown in Fig. 10 has five parts: dataset collection, feature extraction, data augmentation, model training and hyper parameter tuning.

3.1. NOVELTY OF PROPOSED SYSTEM

The proposed model suggests using EEG (electroencephalography) and ECG (electrocardiography) as alternative methods for emotion recognition in paralyzed individuals. This approach addresses the limitations associated with obtaining facial expressions, hand gestures, and voice recordings, which are challenging in paralyzed individuals. EEG measures the electrical activity of the brain and provides valuable insights into brainwave patterns and can indicate different emotional states. ECG, or electrocardiogram, measures the heart's electrical activity. The heart rate variability (HRV) derived from ECG signals reflects emotional states. HRV analysis examines the variations in heartbeat intervals to understand physiological changes associated with emotions. ECG serves as a valuable tool for studying and recognizing individuals' emotional responses. Since EEG and ECG are non-invasive, it is safer and more comfortable for paralyzed individuals for measuring the signals. Therefore, the proposed system of utilization of EEG and ECG signals represents an effective and reliable approach. By analyzing the data derived from EEG and ECG, it becomes possible to precisely evaluate emotional responses, thereby offering valuable insights into the emotional states of paralyzed individuals.

3.2. FEATURE SELECTION

Feature selection involves determining which samples or instances are relevant for the specific task. Three different feature selection techniques are used in determining the relevant features which includes Select K Best, Extra Trees Classifier and Correlation Matrix.

3.2.1. SELECT K BEST METHOD

Select K Best technique involves three main steps which includes scoring, ranking, and feature selection. It applies a scoring function to each feature in the dataset to calculate the score. After calculating the scores for all features, it ranks the features based on their scores in descending order. It finally selects the top k features with the highest scores.

3.2.2. EXTRA TREE CLASSIFIER

Extra Trees is a variant of Random Forest that builds multiple trees and randomly selects subsets of features to split nodes. Unlike Random Forest, it samples observations without replacement and uses random splits instead of the best splits for node partitioning. These differences introduce additional randomness to the training process, making Extra Trees distinct from Random Forest.

3.2.3. CORRELATION MATRIX

By analysing the correlation coefficients between features, highly correlated or redundant features can be identified in Correlation Matric method. By setting threshold, the redundant features can be excluded from the dataset which improves the efficiency of the dataset.

3.3. DATA AUGMENTATION

Data augmentation is a machine learning technique that expands the training dataset by generating modified or synthetic versions of the existing data. Its purpose is to enhance the generalization and resilience of machine learning models by exposing them to a broader range of variations and scenarios. One common method of data augmentation is Random Noise Addition, which involves introducing random noise to the data and thereby increasing the number of records in the dataset. This technique helps improve the performance and effectiveness of machine learning models by providing more diverse and representative training examples.

3.4. MODEL TRAINING

The augmented dataset is trained using five different models and the best model is selected based on the accuracy which is further hyper parameter tuned for improving the performance.

3.4.1. SUPPORT VECTOR MACHINE (SVM)

Support Vector Machine (SVM) is a powerful supervised learning algorithm used for classification and regression tasks. It excels in handling complex datasets, including both linearly separable and non-linearly separable data. By mapping the data to a higher-dimensional feature space, SVM finds an optimal hyperplane to separate different classes or predict continuous values. Its versatility and effectiveness make SVM a valuable tool in machine learning.

Algorithm 1 Emotion Prediction using Support Vector Machine (SVM)

```
Input: EEG and ECG data samples X, Emotion labels Y, Number of epochs E, Regularization parameter C, Kernel function K
Parameters: Regularization parameter C, Kernel, Gamma
Output: Trained SVM model M
Initialize SVM model M
for e = 1 to E do
    for mini-batch in X do
        Forward propagation:
        Apply kernel function K on input data
        Train SVM model with regularization parameter C
        Compute loss using ground truth Y
        Backpropagation:
        No explicit backpropagation for SVM; parameters are updated during training
        Update model parameters
    end for
end for
return Trained SVM model M
```

3.4.2. GAUSSIAN NAIVE BAYES (GNB)

Gaussian Naive Bayes is a machine learning classification technique based on probability and Gaussian distribution. It assumes that each feature has an independent predictive power for the output variable. By combining predictions from all features, it calculates the probabilities of the dependent variable belonging to different groups. The final classification is assigned to the group with the highest probability. Gaussian Naive Bayes is a simple and effective algorithm for classification tasks, especially when features follow a Gaussian distribution.

Algorithm 2 Emotion Prediction using Gaussian Naive Bayes (GNB)

```
Input: EEG and ECG data samples X, Emotion labels Y
Parameters: None
Output: Trained GaussianNB model M
Initialize GaussianNB model M
Fit the model with input data X and labels Y
return Trained GaussianNB model M
```

3.4.3. RANDOM FOREST CLASSIFIER

Random Forest is an ensemble learning algorithm that combines multiple decision trees to create a robust predictive model. It randomly selects subsets of training data and features to build each tree. By aggregating the predictions from multiple trees, Random Forest improves accuracy and handles overfitting. It is widely used for classification and regression tasks, especially with large and complex datasets. Random Forest is a popular and effective algorithm in machine learning.

Algorithm 3 Emotion Prediction using Random Forest

```
Input: EEG and ECG data samples X, Emotion labels Y, Number of trees N, Maximum depth D, Minimum samples split S
Parameters: Number of Trees, Maximum Depth, Minimum Samples Split
Output: Trained Random Forest model M
Initialize Random Forest model M with N trees, maximum depth D, and minimum samples split S
Fit the model with input data X and labels Y
return Trained Random Forest model M
```

3.4.4. MULTI-LAYER PERCEPTRON (MLP)

Multilayer Perceptron (MLP) is a popular artificial neural network used for classification, regression, and pattern recognition tasks in machine learning. It consists of interconnected layers of neurons. Each neuron performs a weighted sum of its inputs, applies an activation function, and produces an output. The output of each neuron serves as input to the neurons in the next layer. With multiple layers and non-linear activation functions, MLP can learn complex relationships in data and make accurate predictions. It is a versatile and widely-used algorithm in machine learning.

```
augment_data
```

	LnHF	HTI	HF	SampEn	SD2	TINN	LFn	MedianNN	MadNN	MCVNN	...
0	0.148437	0.016814	-0.192414	-0.753730	0.322955	0.515258	-0.404926	0.626130	0.231884	0.128559	...
1	-0.747262	-0.545404	-0.467872	-0.806302	0.107772	-0.204271	1.028633	0.430015	-1.113931	-1.125123	...
2	-1.149718	2.494122	-0.525760	1.803011	-0.394467	-0.463804	-0.455935	-2.388904	0.077533	0.422851	...
3	0.048217	-0.771669	-0.222482	0.270719	-0.785849	-1.118437	-0.312018	-0.321003	-1.346474	-1.298967	...
4	0.437140	1.422369	-0.049044	-0.735160	0.808205	0.919499	-0.276979	0.338567	1.687993	1.558128	...
...	...	...	...	...	...	...	...	...	...	...	...
3859	1.017970	0.212810	0.686720	0.292980	1.546280	1.140530	0.769290	1.113450	1.400170	1.363320	...
3860	0.393910	2.090410	0.378710	0.337970	1.736320	1.212830	1.196360	0.246860	1.400640	1.511470	...
3861	0.343670	2.550870	0.487000	1.283700	0.516540	1.492470	0.866760	2.631260	0.517200	0.326390	...
3862	0.307900	0.241780	0.379310	0.094560	0.519610	-0.058690	1.030780	-0.410390	1.350310	1.584080	...
3863	0.631040	1.198180	0.586830	1.920730	0.211260	0.826360	0.893450	1.015960	0.447380	0.492640	...

3864 rows × 26 columns

Figure 3: Dataset after Feature Selection and Data Augmentation

Algorithm 4 Emotion Prediction using Multi-Layer Perceptron (MLP)

Input: EEG and ECG data samples X, Emotion labels Y, Number of epochs E, Learning rate η, Number of hidden layers H, Number of neurons per hidden layer N, Activation function σ
Parameters: Hidden Layer, Activation Function
Output: Trained MLP model M
Initialize MLP model M with H hidden layers, N neurons per hidden layer, and activation function σ
for $e = 1$ to E **do**
  for mini-batch in X **do**
    Forward propagation:
    for each hidden layer h in H **do**
      Perform linear transformation
      Apply activation function σ
    end for
    Perform linear transformation for output layer
    Apply activation function σ
    Compute loss using ground truth Y
    Backpropagation:
    Compute gradients with respect to the loss
    for each hidden layer h in H **do**
      Update weights using gradients and learning rate η
    end for
    Update output layer weights using gradients and learning rate η
  end for
end for
return Trained MLP model M

3.4.5. CONVOLUTIONAL NEURAL NETWORK (CNN)

Convolutional Neural Networks (CNNs) are a type of artificial neural network commonly used for processing grid-like data, especially images. However, CNNs can also be adapted for processing numerical data with a structured representation, such as time series or one-dimensional sequences.

3.5. HYPER PARAMETER TUNING

Hyperparameter tuning, also known as hyperparameter optimization, is the process of finding the best hyperparameter values for a machine learning algorithm. Hyperparameters are configuration settings that affect the performance of the model but are not learned from the data. Genetic algorithms (GAs) are a type of optimization technique inspired by natural selection and genetic evolution. GAs iteratively search for the optimal hyperparameter values by generating a population of solutions, evaluating their fitness, and using genetic operators like selection, crossover, and mutation to evolve the population over generations. GAs can be effective in finding approximate solutions to complex optimization problems in machine learning.

4. RESULT AND ANALYSIS

4.1. DATASET DESCRIPTION

The DREAMER dataset is a multi-modal database that contains electroencephalogram (EEG) and electrocardiogram (ECG) signals recorded during the elicitation of four different emotions (happy, anger, fear, and sad) using audio-visual stimuli. The dataset includes signals from 23 participants, along with their self-assessment of affective states in terms of valence, arousal, and dominance after each stimulus. The signals were

Algorithm 5 Emotion Prediction using Conv1D-CNN

Input: EEG and ECG data samples X, Emotion labels Y, Number of epochs E, Learning rate η
Parameters: No. of filter, Size of Kernal, Stride of Convolution layer, Size of Max Pooling, Stride of Max Pooling
Output: Trained CNN model M
Initialize Conv1D-CNN model M
for $e = 1$ to E **do**
    for mini-batch in X with size BS **do**
        Forward propagation:
        for each Conv1D layer l in L **do**
            Apply 1D convolution with K filters of size F
            Apply activation function
        end for
        for each pooling layer l in P **do**
            Perform max-pooling or average-pooling
        end for
        Flatten the output
        for each fully connected layer l in FC **do**
            Perform linear transformation
            Apply activation function
        end for
        Compute loss using ground truth Y
        Backpropagation:
        Compute gradients with respect to the loss
        Update model parameters
    end for
end for
return Trained Conv1D-CNN model M

Algorithm 6 Genetic Algorithm for CNN Parameter Optimization

Input: EEG and ECG data samples X, Emotion labels Y, Population size P, Number of generations G, Crossover probability C_p, Mutation probability M_p
Parameters: No. of filter, Size of Kernal, Stride of Convolution layer, Size of Max Pooling, Stride of Max Pooling
Output: Optimized CNN hyperparameters
Initialize population P with individuals representing random hyperparameter combinations within the specified ranges
for $g = 1$ to G **do**
    Evaluate fitness of each individual in P using CNN with corresponding hyperparameters
    Select individuals for crossover based on their fitness
    while new population size $< P$ **do**
        Perform crossover with probability C_p to create new individuals
        Perform mutation with probability M_p on new individuals
        Add new individuals to the population
    end while
end for
Select the best individual from the final population based on fitness
return Optimized CNN hyperparameters

recorded using portable, wearable, wireless, low-cost, and commercially available equipment, making it suitable for everyday applications of affective computing methods. This dataset has 184 records and 69 columns.

4.2. FEATURE SELECTION

All the three methods: Select K Best, Extra Tree Classifier and Correlation Matrix are used in determining the top features to be selected to train the model. In Select K Best method, the top 58 features gave an accuracy of 41%. In Extra Tree Classifier method, the top 26 features gave 43% accuracy. The Correlation Matrix method gave the highest accuracy of 32% for the threshold value of 1.0. Thus the Extra Tree Classifier method is used in determining the required features to train the model. Thus the initial 68 features in the dataset is reduced to 26.

4.3. DATA AUGMENTATION

The existing DREAMER dataset consists of 184 records, which is not sufficient to build an efficient model. Larger training data sets yield more reliable results with smaller margins of error and lower standard deviations. They provide greater diversity and allow the model to capture a wider range of patterns, improving accuracy. Increased sample size reduces the impact of outliers and random variations, enhancing model stability and robustness. Standard deviation measures how spread out the data values are from the mean. Random Noise Addition method is used for data augmentation, which adds a random value to the existing dataset based on the standard deviation of each column. This increased the dataset from 184 records to 3864 records. The final dataset which is used to train the model consists of 3864 rows and 26 columns is shown in Fig. 3.

4.4. MODEL TRAINING

The DREAMER dataset with 3864 rows and selected 26 features is used to train five models namely SVM, GNB, Random Forest Classifier, MLP and CNN. The training loss of MLP is shown in Fig. 4 and model loss of CNN in Fig. 5. The models are compared based on their accuracy in Fig. 5. Among the five models, the CNN outperformed all other models and has an accuracy of 72%. Thus CNN is selected for hyper parameter tuning.

4.5. HYPER PARAMETER TUNING

The selected CNN model is hyper parameter tuned using Genetic Algorithm. The different parameters of CNN model which is hyper parameter tuned using Genetic Algorithm are number of layers, kernel size, convolution stride, pooling size, and pooling stride. The initial population of individuals is created to kickstart the optimization process. This population is typically generated randomly which contains all the hyperparameters to start the algorithm. A population of random individuals which contains all the hyper parameters is generated to start the algorithm. Each individual in the population is evaluated using a fitness which determines the selection of individuals for the

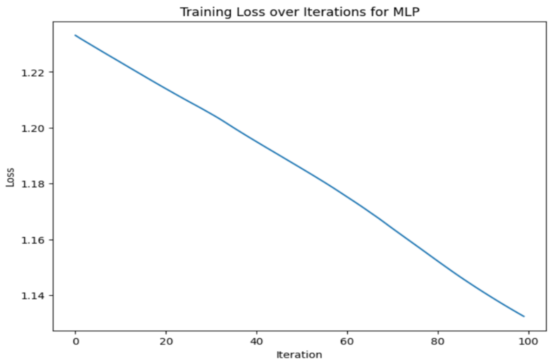

Figure 4: Training Loss Over Iterations of MLP

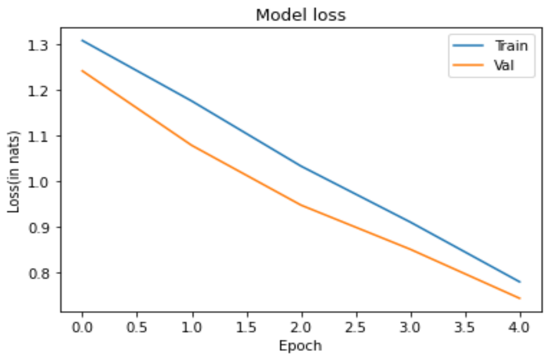

Figure 5: Model Loss of CNN

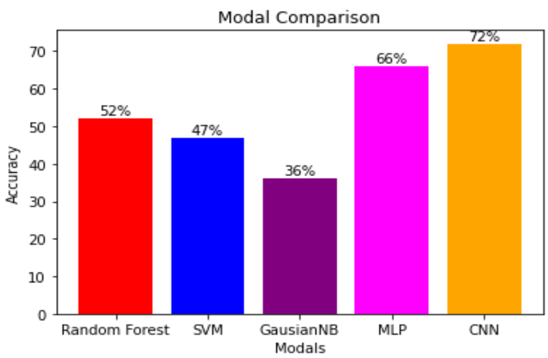

Figure 6: Comparison of Five Models Trained with Same Dataset

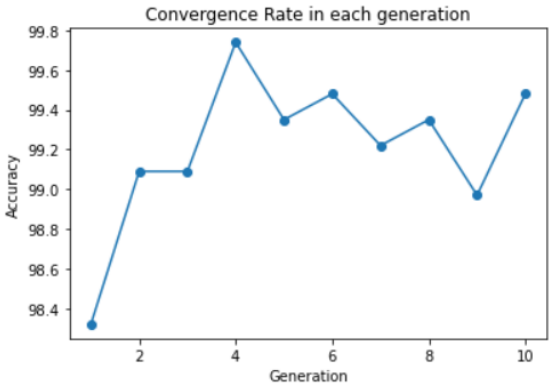

Figure 7: Convergence Rate of Genetic Algorithm

next steps based on their suitability. Individuals will be selected based on their fitness value for reproduction. This step aims to preserve the fittest individuals and promote the propagation of their genetic material. Selected individuals are used to create offspring through genetic operators such as crossover and mutation. Crossover involves exchanging genetic information between two parents to create new individuals, while mutation does a random change to the individuals to maintain diversity in the population. The offspring population replaces the previous population, ensuring that the population size remains constant. This step allows for the exploration of new solutions while maintaining some of the best solutions from the previous generation. The algorithm continues the evaluation, selection, reproduction, and replacement steps for a certain number of generations or until the desired fitness level is reached.

The average convergence rate (ACR) is a metric used to assess the convergence speed of an evolutionary algorithm. Specifically, it measures how quickly the approximation error of the algorithm decreases over consecutive generations as shown in the Fig. 7.

The optimised CNN model is developed using the parameter obtained from the fittest individual, which gives an accuracy of 98.71% and the confusion matrix of the hyper parameter tuned CNN model is shown in Fig. 9. The performance metrics of all five trained models and hyper parameter tuned CNN model is given in Table 2.

The hyper parameter tuned model is deployed using flask application in Amazon EC2 virtual server for public access. A frontend web application is developed using reactjs and deployed in Amazon S3. The users can upload their EEG and ECG signal as CSV file in the frontend application, which will interact with the deployed model using RESTful APIs and predicts the emotion as shown in Fig **??**. The predicted emotions can be downloaded as a PDF report for further usage.

5. CONCLUSION

A multimodal model system is developed to recognise the different emotions of the people. The dataset with 184 records containing the different emotions includes happy, angry, sad and fear was used for this proposed work. Data augmentation

Table 2: Performance metrics of different trained models in percentage

METRICS	SVM	GNB	RANDOM FOREST	MLP	CNN	TUNED CNN
Accuracy	47	36	52	66	72	98
Precision	49	34	53	53	66	99
Recall	35	50	64	64	82	98
F1-Score	41	40	58	58	73	99

Table 3: Hyperparameters, Range Considered, and Optimal Values for Different Models

Model	**Hyperparameter**	**Range Considered**	**Optimal Value**
CNN	No. of filter	[32, 64]	64
	Size of Kernal	[3, 5]	3
	Stride of Convolution layer	[1,3]	3
	Size of Max Pooling	[2,4]	2
	Stride of Max Pooling	[1,4]	3
RF	Number of Trees	[50, 100, 200]	100
	Maximum Depth	[5, 10, 20]	10
	Minimum Samples Split	[2, 5, 10]	5
SVM	C (Regularization Parameter)	[0.1, 1, 10]	1
	Kernel	[Linear, RBF, Polynomial]	RBF
	Gamma	[0.01, 0.1, 1]	0.1
MLP	Hidden Layers	[1, 2, 3]	2
	Activation Function	[ReLU, Tanh, Sigmoid]	ReLU

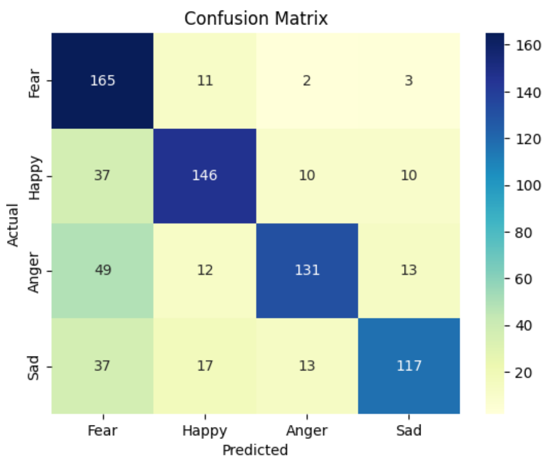

Figure 8: Confusion Matrix of CNN Before Tuning

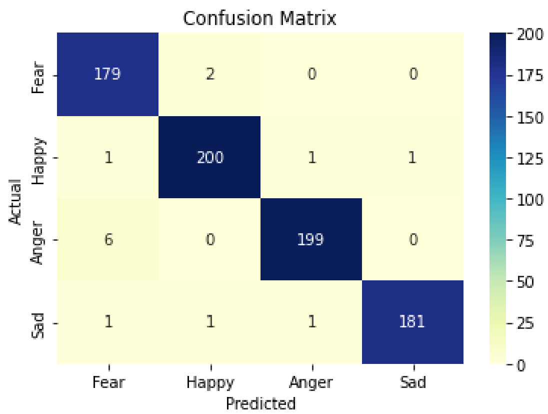

Figure 9: Confusion Matrix of CNN After Tuning

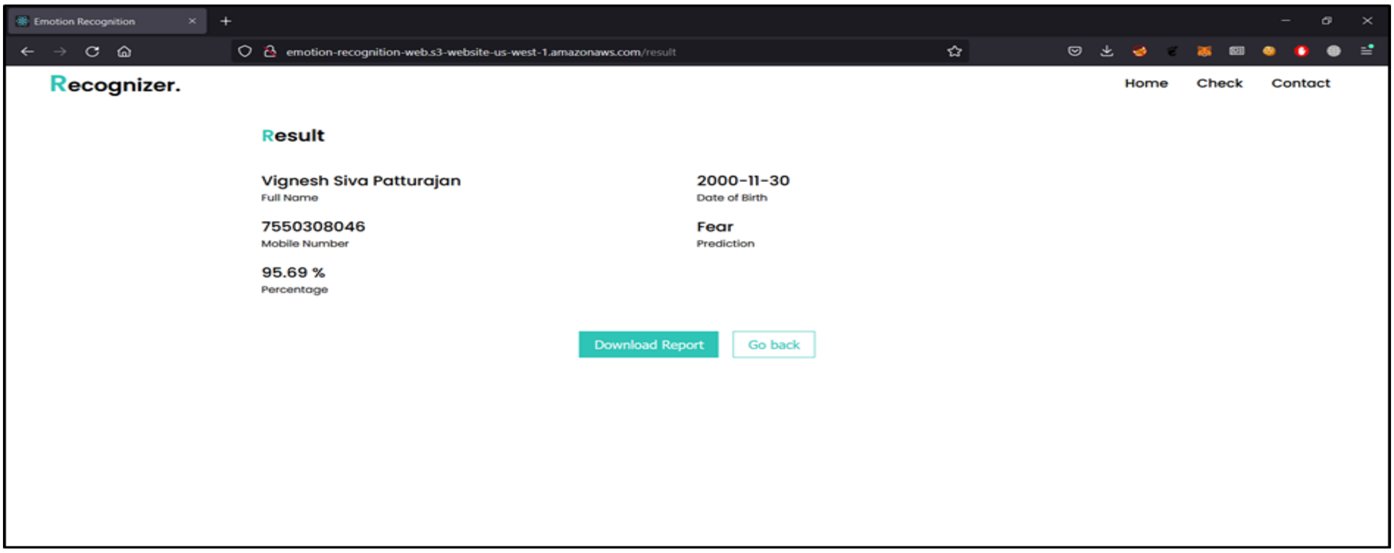

Figure 10: Proposed System Workflow

is done to increase the size of the dataset for improving the model's performance. This augmented dataset is used to train five different models, in which CNN performed better when compared to other models. The CNN model is further hyper parameter tuned using Genetic Algorithm. The resulted tuned CNN model is deployed in Amazon EC2 using flask application. A web application is developed for the users to input their signals and is hosted in Amazon S3 for public access. The web application interacts with the model using REST API to predicts the emotions and generates a PDF report for the users to download.

References

[1] Ferdinando, Hany, Tapio Seppänen, and Esko Alasaarela. "Comparing features from ECG pattern and HRV analysis for emotion recognition system", IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB). IEEE, 2016.

[2] Mumenthaler, Christian, David Sander, and Antony SR Manstead. "Emotion recognition in simulated social interactions", IEEE Transactions on Affective computing vol. 11, no. 2, pp. 308-312, 2018.

[3] Tarnowski, Paweł, Marcin Kołodziej, Andrzej Majkowski, and Remigiusz Jan Rak. "Combined analysis of GSR and EEG signals for emotion recognition", International Interdisciplinary PhD Workshop (IIPhDW). IEEE, pp. 137-141, 2018.

[4] Chopra, Sejal, Atharva Aglawe, Shyam Singh, and Amogh Ambekar. "ParaInfia-An Infotainment System for semi-paralyzed using Electroencephalogram", International Conference on Nascent Technologies in Engineering (ICNTE). IEEE, pp. 1-7, 2019.

[5] Qing, Chunmei, Rui Qiao, Xiangmin Xu, and Yongqiang Cheng. "Interpretable emotion recognition using EEG signals", IEEE Access, vol. 7, pp. 94160-94170, 2019.

[6] Xu, Jianqiang, Zhujiao Hu, Junzhong Zou, and Anqi Bi. "Intelligent emotion detection method based on deep learning in medical and health data", IEEE Access vol. 8, pp. 3802-3811, 2019.

[7] Aldayel, Mashael S., Mourad Ykhlef, and Abeer N. Al-Nafjan. "Electroencephalogram-based preference prediction using deep transfer learning", IEEE Access, vol. 8, pp. 176818-176829, 2020.

[8] Begaj, Sabrina, Ali Osman Topal, and Maaruf Ali. "Emotion Recognition Based on Facial Expressions Using Convolutional Neural Network (CNN)", International Conference on Computing, Networking, Telecommunications & Engineering Sciences Applications (CoNTESA). IEEE, pp. 58-63, 2020.

[9] Liu, Xiao, Xiangyi Cheng, and Kiju Lee. "Ga-svm-based facial emotion recognition using facial geometric features", IEEE Sensors Journal, vol. 21, no.10, pp. 11532-11542, 2020.

[10] Sheykhivand, Sobhan, Zohreh Mousavi, Tohid Yousefi Rezaii, and Ali Farzamnia. "Recognizing emotions evoked by music using CNN-LSTM networks on EEG signals", IEEE access, vol. 8 pp. 139332-139345, 2020.

[11] Sun, Ting-Wei. "End-to-end speech emotion recognition with gender information", IEEE Access, vol. 8, pp. 152423-152438, 2020.

[12] Zhang, Hongli. "Expression-EEG based collaborative multimodal emotion recognition using deep autoencoder", IEEE Access, vol. 8, pp. 164130-164143, 2020.

[13] Casaccia, Sara, Erik J. Sirevaag, Mark G. Frank, Joseph A. O'Sullivan, Lorenzo Scalise, and John W. Rohrbaugh. "Facial muscle activity: High-sensitivity noncontact measurement using laser Doppler vibrometry", IEEE Transactions on Instrumentation and Measurement, vol. 70, pp. 1-10, 2021.

[14] Dessai, Amita, and Hassanali Virani. "Emotion Detection Using Physiological Signals", International Conference on Electrical, Computer and Energy Technologies (ICECET). IEEE, 2021.

[15] Pham, Minh, Ha Manh Do, Zhidong Su, Alex Bishop, and Weihua Sheng. "Negative emotion management using a smart shirt and a robot assistant", IEEE Robotics and Automation Letters vol.6, no. 2, pp. 4040-4047, 2021.

[16] Zhang, Chenghao, and Lei Xue. "Autoencoder with emotion embedding for speech emotion recognition", IEEE access, vol. 9, pp. 51231-51241, 2021.

[17] Kakuba, Samuel, Alwin Poulose, and Dong Seog Han. "Deep Learning-Based Speech Emotion Recognition Using Multi-Level Fusion of Concurrent Features", IEEE Access, vol. 10, pp. 125538-125551, 2022.

[18] Li, Dahua, Jiayin Liu, Yi Yang, Fazheng Hou, Haotian Song, Yu Song, Qiang Gao, and Zemin Mao. "Emotion Recognition of Subjects with Hearing Impairment Based on Fusion of Facial Expression and EEG Topographic Map", IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2022.

[19] Satyanarayana, KN V., V. Tejasri, YS Naga Srujitha, K. Nitya Sai Mounisha, Sai Tejasri Yerramsetti, and Gayathri Devi Darapu. "Human Emotion Classification using KNN Classifier and Recurrent Neural Networks with Seed Dataset", 6th International Conference on Computing Methodologies and Communication (ICCMC). IEEE, pp. 1717-1722, 2022.

[20] Seo, Wonju, 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, 2022.

[21] Qi, Xin, Yun Zhang, Yicheng Jiang, Elisa Giusti, and Marco Martorella. "Optimized Nonlinear PRI Variation Strategy Using Knowledge-Guided

Genetic Algorithm for Staggered SAR Imaging", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 16, pp. 2624-2643, 2023.

[22] Ramachandran, Prakash. "Compressive Wavelet Domain Deep CNN for Image Classification Using Genetic Algorithm Based Sensing Mask Learning", IEEE Access, vol. 11, pp. 35567-35578, 2023.

[23] P. Sarkar and A. Etemad, "Self-Supervised Learning for ECG-Based Emotion Recognition," ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 2020, pp. 3217-3221, doi: 10.1109/ICASSP40776.2020.9053985.

[24] N. Y. Khan, N. Al Mahmood, H. Kamal, T. Jebin and M. H. Imam, "Emotional State Analysis Using Handheld Portable ECG Device," 2019 IEEE International Conference on Biomedical Engineering, Computer and Information Technology for Health (BECITHCON), Dhaka, Bangladesh, 2019, pp. 33-36, doi: 10.1109/BECITHCON48839.2019.9063178.

[25] G. Giannakakis, E. Trivizakis, M. Tsiknakis and K. Marias, "A novel multi-kernel 1D convolutional neural network for stress recognition from ECG," 2019 8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW), Cambridge, UK, 2019, pp. 1-4, doi: 10.1109/ACIIW.2019.8925020.

[26] P. Sarkar and A. Etemad, "Self-Supervised ECG Representation Learning for Emotion Recognition," in IEEE Transactions on Affective Computing, vol. 13, no. 3, pp. 1541-1554, 1 July-Sept. 2022, doi: 10.1109/TAFFC.2020.3014842.

[27] N. Y. Oktavia, A. D. Wibawa, E. S. Pane and M. H. Purnomo, "Human Emotion Classification Based on EEG Signals Using Naïve Bayes Method," 2019 International Seminar on Application for Technology of Information and Communication (iSemantic), Semarang, Indonesia, 2019, pp. 319-324, doi: 10.1109/ISEMANTIC.2019.8884224.

[28] V. Gupta, M. D. Chopda and R. B. Pachori, "Cross-Subject Emotion Recognition Using Flexible Analytic Wavelet Transform From EEG Signals," in IEEE Sensors Journal, vol. 19, no. 6, pp. 2266-2274, 15 March15, 2019, doi: 10.1109/JSEN.2018.2883497.

[29] X. Du et al., "An Efficient LSTM Network for Emotion Recognition From Multichannel EEG Signals," in IEEE Transactions on Affective Computing, vol. 13, no. 3, pp. 1528-1540, 1 July-Sept. 2022, doi: 10.1109/TAFFC.2020.3013711.

[30] J. Cheng et al., "Emotion Recognition From Multi-Channel EEG via Deep Forest," in IEEE Journal of Biomedical and Health Informatics, vol. 25, no. 2, pp. 453-464, Feb. 2021, doi: 10.1109/JBHI.2020.2995767.

[31] L. Cai, J. Dong and M. Wei, "Multi-Modal Emotion Recognition From Speech and Facial Expression Based on Deep Learning," 2020 Chinese Automation Congress (CAC), Shanghai, China, 2020, pp. 5726-5729, doi: 10.1109/CAC51589.2020.9327178.

[32] N. Song, H. Yang and P. Wu, "A Gesture-to-Emotional Speech Conversion by Combining Gesture Recognition and Facial Expression Recognition," 2018 First Asian Conference on Affective Computing and Intelligent Interaction (ACII Asia), Beijing, China, 2018, pp. 1-6, doi: 10.1109/ACIIAsia.2018.8470350.

[33] T. Keshari and S. Palaniswamy, "Emotion Recognition Using Feature-level Fusion of Facial Expressions and Body Gestures," 2019 International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2019, pp. 1184-1189, doi: 10.1109/ICCES45898.2019.9002175.

[34] K. M. Pathak, S. Yadav, P. Jain, P. Tanwar and B. Kumar, "A Facial Expression Recognition System To Predict Emotions," 2020 International Conference on Intelligent Engineering and Management (ICIEM), London, UK, 2020, pp. 414-419, doi: 10.1109/ICIEM48762.2020.9160229.

[35] R. GUETARI, A. CHETOUANI, H. TABIA and N. KHLIFA, "Real time emotion recognition in video stream, using B-CNN and F-CNN," 2020 5th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), Sousse, Tunisia, 2020, pp. 1-6, doi: 10.1109/ATSIP49331.2020.9231902.

[36] R. Rajak and R. Mall, "Emotion recognition from audio, dimensional and discrete categorization using CNNs," TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON), Kochi, India, 2019, pp. 301-305, doi: 10.1109/TENCON.2019.8929459.

[37] Z. Peng, Y. Lu, S. Pan and Y. Liu, "Efficient Speech Emotion Recognition Using Multi-Scale CNN and Attention," ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Toronto, ON, Canada, 2021, pp. 3020-3024, doi: 10.1109/ICASSP39728.2021.9414286.

[38] Kollias and S. Zafeiriou, "Exploiting Multi-CNN Features in CNN-RNN Based Dimensional Emotion Recognition on the OMG in-the-Wild Dataset," in IEEE Transactions on Affective Computing, vol. 12, no. 3, pp. 595-606, 1 July-Sept. 2021, doi: 10.1109/TAFFC.2020.3014171.

[39] O. S. Mohammed, A. A. A. -M. Sewisy and A. I. Taloba, "Solving Optimization Problems using Hybrid Metaheuristics: Genetic Algorithm and Black Hole Algorithm," 2020 2nd International Conference on Computer and Information Sciences (ICCIS), Sakaka, Saudi Arabia, 2020, pp. 1-5, doi: 10.1109/ICCIS49240.2020.9257717.

[40] A. Gupta, R. Machavaram, T. Kshatriya and S. Ranjan, "Multi-Objective Design Optimization of a Three Phase Squirrel Cage Induction Motor for Electric Propulsion System using Genetic Algorithm," 2020 IEEE First International Conference on Smart Technologies for Power, Energy and Control (STPEC), Nagpur, India, 2020, pp. 1-6, doi: 10.1109/STPEC49749.2020.9297776.