PERIODICO di MINERALOGIA Volume 91, No. 5, 2022
ISSN: 0369-8963

Recognition of Emotion with Deep Learning using EEG signals - The Next Big Wave for Stress Management in this Covid-19 outbreak

RAJESWARI RAJESH IMMNAUEL[1], S K B SANGEETHA[2]

[1]*Departmnet of Computer Science and Engineering, SRM Institute of Science and Technology, Vadapalani, Chennai, India*

[2]*Departmnet of Computer Science and Engineering, SRM Institute of Science and Technology, Vadapalani, Chennai, India*

Corresponding author: Rajeswari Rajesh Immanuel

ABSTRACT

Emotion is a reaction of mind and body together for any action which originates from the brain. Electroencephalogram (EEG) signals are the primary resource of human emotion which is produced in the brain for any reaction. These EEG signals are the electrical activities that are captured with the EEG signal recorder and are used to analyze the state of emotion. Stress being one of the emotions plays a major role in health disorders of humankind in the recent year of pandemic and other changes in the cultural and social behaviors. Sometimes, the stress level increases and the same contributes to many health problems. As a result, several machine learning algorithms have been utilized in the past few years to examine different attributes of emotion recognition with regards to stress. Deep learning-based methods, on the other hand, have recently gained popularity due to their great performance and accuracy. To explore emotions, particularly stress, the deep learning approach was adopted because Deep Neural Networks (DNN) performs better when compared to machine learning algorithms due to its long-cascaded architecture with several processing hidden layers. Various input parameters are compared and explained why EEG signals are the primary source of input for detecting stress among humans. Due to nonlinearity, the EEG signals were less considered in the past for identifying the human emotion. This work gives a detailed analysis of the most extensively used deep learning models in emotion identification, discusses the implications of popular deep learning architectures, and proposes an emotion recognition classification. Many researchers' contributions are emphasized, with a special emphasis on deep learning approaches employing EEG inputs for recognizing the emotions. The most commonly used deep datasets, the deep learning architecture that were used on them, the accuracy obtained from them, and a comparison of other deep learning models are explained. Studies that perform emotion recognition using various deep learning techniques are also compared and concluded. In results, different emotion recognition models are compared with the same dataset and CNN classifier performs better than other models. The accuracy of the CNN is 98% and validation loss is 1.16.

Key words: *Emotion Recognition, Deep Learning, EEG Signals, Classification, CNN, Stress.*

 https://doi.org/10.37896/pd91.5/91542

INTRODUCTION

Emotion is a complicated state of feeling that influences one's ideas and behavior. From curiosity and satisfaction to love and joy, positive emotions cover a wide variety of pleasurable or desirable situational responses. Stress is one among the many emotions of human beings. An unpleasant or unhappy emotion evoked in people to express a negative reaction to an event or person is referred to as a negative emotion. Neutral emotion is the feeling between positive and negative emotions such as being calm and quiet. Stress is a kind of thing when body's reaction to pressure due to various stimuli around us which is called a negative emotion. Currently, the term stress is used widely and it seeks primary attention due to modern lifestyle, smartphone usage, social media, online video games, and work from home, online classes, pandemic situations like the Covid- 19 and many more related issues that the world is facing today [1].

Stress can be caused by a variety of conditions or life events that occur in one's life. When people encounter something new and unexpected that threaten us, or when people do not have control over our situations, stress is developed in the body. When people are stressed, the bodies release a stress hormone called Cortisol. The Cortisol hormone communicates with the brain regions that control mood, motivation, and fear. Immediately the brain sends signals to control the body. Sometimes, this stress hormone helps to handle the situation quickly and positively. When the same stress prolongs for a long time, it causes psychological and psychological problems in humans. This also leads to many heart diseases, diabetes, blood-pressure, obesity, mental health related problems and many more. Early detection of stress helps humans to avoid or prevent many diseases related to its [2].In recent years, stress has gained a predominant role in society which needs to be addressed.

Stress can be further classified as physiological and psychological which can be measured using different types of devices. Physical stress can be identified using different physical parameters. Psychological or mental emotion can be detected using Electroencephalography (EEG) which is a device that monitors and records brain wave patterns signals are our body's primary source of emotion. In the year 1924, Han Berger recorded the first EEG signals. Deep learning algorithms outperform machine learning algorithms when it comes to EEG readings. When compared to machine learning, Brain Computer Interaction (BCI) research publications claim to improve the system's accuracy. Despite the fact that machine learning classification algorithms are primarily built for classification, deep learning obtains better results when EEG inputs are used [3].

The research in emotion recognition is less focused with EEG signals or brain signals because EEG signals are not stationary in nature. Hence, developing an intelligent framework for emotion recognition is difficult. The development of an intelligent emotion recognition system has taken a lot of time and effort for researchers. Some research has been conducted on non-physiological signals like body movements, facial emotion recognition, speech signals and many more. The above-mentioned methodologies are highly subjective and also depend on one's cultural background and physical health. Therefore, it is hard to conclude that the obtained result is accurate [4].

Some research has been conducted on physiological signals like pulse rate, impedance of the skin, heart rate and functional magnetic resonance imaging (fMRI). However, all of these parameters cannot contribute reliable output when compared. The accuracy for valence, arousal and the dominance of the emotion identified with EEG signals. The parameter valence refers to the pleasantness of a stimulus and varies from negative to positive in values. Arousal is a variable that describes the intensity of emotion elicited by stimuli and ranges from mild to severe. Parameter dominance refers to the extent of control exercised by the stimuli. To get the desired effect, EEG waves are input into a neural network [5].

The emotion can be of different types and most of the models in the studies look at different emotions such as happiness, delight, excitement, contentment, relaxation, calmness, depression, boredom, tiredness, tension, anger and frustration as explained [6] in Figure 1.

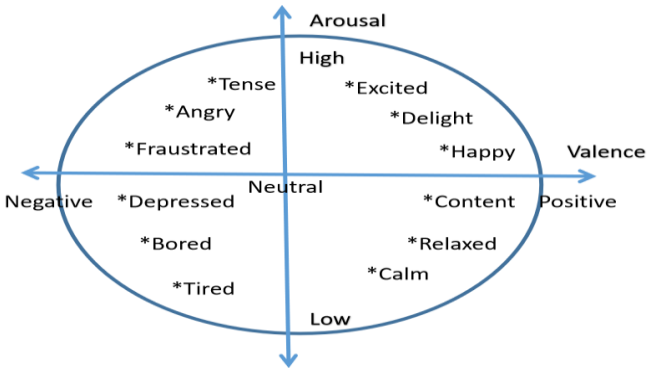

Figure 1. 2-Dimensional Emotion Recognition Model for Arousal - Valence space. Valence values vary from negative to positive. Arousal values vary from low intensity to high intensity.

Deep learning, which is a cutting-edge technology, is used to effectively forecast these emotions. Deep learning approach focuses on the problem contrary to machine learning which processes a huge volume of data in a linear way. Hence, the deep learning approach has been chosen to study emotions, especially stress. A long-cascaded architecture with multiple processing layers makes Deep Neural Networks (DNN) perform better when compared to machine learning algorithms. In DNN, the extracted data becomes more abstract from the lowest to the topmost layers. One of the benefits of deep learning is that information expression might be more meaningful as it is passed on to higher layers. These higher layers of DNN help to predict or recognize the emotion to an accurate level [7] [8]. This research is carried out to throw a glimpse on the different techniques in deep learning and related studies in stress management with regards to EEG signals. This review article will provide a more in-depth look at the prior papers. Table 1 compares the recent studies related to the study of EEG signals with emotion recognition in deep learning networks. It also explains the detailed study given in this paper more than others.

This review clearly answers the following queries

1. What is the complete roadmap of emotion recognition using various input features?
2. How are EEG signals more important than other input features?

3. Show the performance evaluation of various EEG based deep learning techniques with suitable survey records.

TABLE 1. Recent Studies Vs Proposed Study

Sr.no	Author	Content focused	Proposed study
1	Zhang Yaqing et al.,2020	Explains different deep learning techniques especially focuses on CNN only	Comparison of different deep learning algorithms explained in tabular form elaborately
2	Alexander Craik et al., 2019	Only EEG signal applications and their related deep learning methods are shown	EEG signal are the primary source of emotion origins from the brain are mentioned with proof
3	Gen Li et al.,2019	EEG signal properties are described clearly and limited explanation of the deep learning methods	Why EEG signal is better than other inputs for emotion recognition is tabulated clearly and deeply
4	Yannick Roy et al.,2019	Displays various fields where EEG emotion recognition are used and throws light on few deep learning methods	Each deep learning algorithm is explained well with related studies in tabular reference

The papers obtained through this search are then carefully evaluated to eliminate works by the same author that are similar and incremental, leaving only unique works towards EEG-based emotion detection that can be demonstrated. In section 1, emotions, basics of EEG signals and the importance of deep learning techniques are discussed. The paper also goes over emotions, as well as the fundamental features of EEG signals and deep learning networks. Section 2 includes recent works within the last five years by using specific keywords search such as EEG based emotion recognition. Section 3 will go over the full context of EEG signals in different perspectives and present various works to demonstrate how important EEG signals are in comparison to other input features. Section 4 will go through the details of several deep learning approaches which classifies the emotions using EEG signals with better comparison

analysis using performance evaluation as never before. Section 5 will talk about the implementation and evaluation results of the different models. Section 6 concludes the study.

RELATED STUDY

More than five decades, EEG signals have been used to recognize emotions. Many new feature extraction and classification approaches for emotion detection have recently been suggested. In the meantime, different machine learning classification approaches have been utilized for emotion recognition. Apart from that, many recent works within the last five years have been studied in this section. Emotion recognition is carried out with general lifestyle features like gender, age, habits conducted with KNHANES VI dataset. The Deep Belief Network (DBN) algorithm is compared with other conventional machine algorithms such as Support Vector Machine (SVM), Naive Bayes Classification (NBC) and Random Forest (RF). The specificity and accuracy in predicting the stress among people are 75.32% and 66.23% respectively. The first layer of the model is dedicated to data collecting and data cleaning, while the second layer is dedicated to data analysis. Feature extraction is performed using t-test and Chi-square test. Finally, the model predicts the stress [9].

Speech is a way to communicate the thoughts of humans. Hence, speech signals were used to detect emotions of the 56 subjects with the help of RNN and algorithm to obtain the maximum accuracy with Long Short-Term Memory (LSTM). Using pre-processed speech samples, the system first obtains mel-filterbank coefficients. Then the output is derived using binary decision criterion using LSTM and the accuracy of the model was 66.4%. (Hyewon Han et al., 2018). Work related stress among selected Smart factory workers was conducted with DAIC -WOZ dataset with 116 audio recordings. Smart factories are digital environments that are highly automated with a cyber-ecosystem. Speaker diarization and sampling strategy techniques are used for noise reduction in the audio samples [10].

Social media posts are also taken as a dataset to identify whether the person is emotionally balanced. Depression lexicon is manipulated and used the same to derive for the conclusion. Kaggle.com dataset from Twitter is used to find post-depression of cancer survivors by their Twitter post. CNN is compared to the other algorithms such as MLP, NBC, SVM, CNN n-gram. The accuracy achieved by this model is 91.29%. In this paper, they have used a text analysis tool to find the depression lexicon in the tweet. Then the extracted data is fed to filtering and to CNN layers. The output is whether the cancer survivor is stressed or not [11]. Deep learning approaches are applied in criminals' psychological problems by recording their EEG signals. Around 109 criminal EEG signals are recorded. The EEG is preprocessed using a time-domain regression approach to denoise it and extracted EEG signal features are given to the back propagation (BP) algorithm and obtained emotion recognition is 93.5%.The recorded dataset helps to evaluate the stress level in them with MATLAB Tool. The BP algorithm performance is compared with Bayes classifier, SVM and Hidden Markov Model (HMM) [12].

The study proposes using a late-fusion strategy with two phases to merge acoustical and textual information. The acoustic and text features from the speech have been fed to the SVM to find Arousal-Valence-Dominance accuracy of the model. The dataset used in this model are from IEMOCAP and MSP-

IMPROV datasets and measuring devices used here are Biosemi Active and The Emotiv EPOC. The accuracy of emotion recognition is compared between two datasets. It also compares SVM to Recurrent Neural Network (RNN) - LSTM [13].

In these recent studies as shown in Table 2, the types of emotion stimuli and how they are presented, the size of the study, EEG technology, and machine learning classifiers are all highlighted. All the effort was dedicated solely to examining classification works in this field of study. This review paper is aimed for researchers who are working on EEG signals for emotion recognition, especially stress management.

DIFFERENT PERSPECTIVES OF EEG SIGNALS - AN EXCEPTIONAL TOOL FOR HUMAN BEHAVIOR

A. EEG SIGNALS - AN OVERVIEW

Emotion recognition can be very accurate when it is measured using EEG signals as the emotions originate from the brain. The human brain functions with spontaneous electrical activity and that can be captured with an EEG recorder. A particular EEG recorder can translate this activity into brainwaves. EEG is a rhythmic electrical activity that occurs spontaneously and has a frequency that ranges from 1 to 30 times per second. In the study of EEG, there are at least four major bands. The four different bands of EEG are produced during different emotions [14]. The placement of the different electrodes is given below in Figure 2 and EEG signal image is depicted in Figure 3.

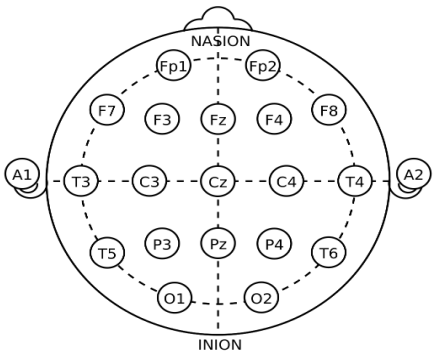

Figure 2. Position of the electrodes on the skull - Explains the position of electrodes on the skull to capture the brain signals using an EEG recorder.

ISSN: 0369-8963

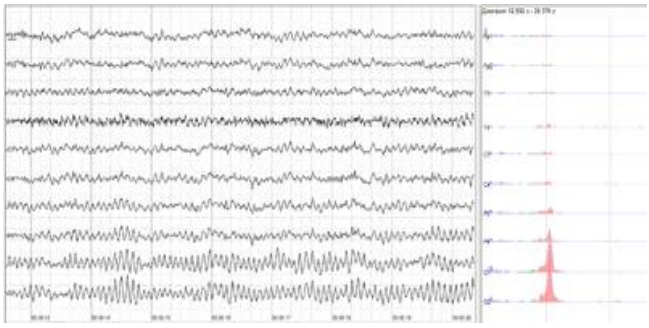

Figure 3. EEG Signal image

TABLE 2. EEG based Emotion Recognition - Recent Studies Comparison

Sr. no	Author Name	Input	Output	Method/ Algorithm	Performance Metric/accuracy	Content	Dataset
1.	Se-Hui Song et al. (2017)	9 features- Gender of the subject, Life style inputs	Low Stress High Stress	Stress Classification DBL	Accuracy – 66.23% Specificity – 75.32%	DBL performs better than SVM, NBC, RF DL4J Java based tool is used	KNHANES VI
2.	Hyewon Han et al. (2018)	25 subjects Speech Signals	Stressed Unstressed	LSTM-RNN Feedback Network	Accuracy - 66.4%	Stress detection model using LSTM	Recorded 56 subject (audio-visual Data, Bio-signals)
3.	Ishara Madhavi et al. (2020)	116 Audio recordings Spectral and Temporal Feature	Stressed Distressed	CNN GSOM	Accuracy – 82 % Effectiveness – 64%	Industry 4.0 technologies and Operator 4.0 behaviors. This model detects the stress among the workers in smarter factories	DAIC -WOZ Dataset

 https://doi.org/10.37896/pd91.5/91542

4.	Nur Hafieza İsmail et al. (2020)	Depression Lexicon from Twitter Post	Stressed Distressed	CNN	Accuracy - 91.29%	CNN is compared NBC, SVM, MLP, CNN n-gram This article discusses the stress detection of the cancer survivor's post on the social media(twitter)	Kaggle.com Dataset
5.	Qi Liu et al. (2020)	109 criminals EEG signals are recorded	Stressed Distressed	DLEEG over DSP, Back Propagation Algorithm	Accuracy - 93.5%	deep learning used with EEG to detect the stress level in criminals MATLAB	Recorded EEG dataset
6.	Bagus Tris Atmaja et al. (2021)	Speech input Acoustic and Text features	Emotion Recognition	SVM	Arousal, Valence, Dominance Accuracy is compared between two dataset	Measuring device Biosemi Active The Emotiv EPOC	IEMOCAP MSP-IMPROV

EEG has a really high time resolution and may record cognitive abilities in the moment they happen. The cognitive, perceptual, verbal, emotional, and motor processing are all extremely quick. The majority of cognitive processes take tens to hundreds of milliseconds to complete, which is far faster than the blink of an eye. Furthermore, the triggers that initiate cognitive processes occur in time sequences that range from a few seconds to hundreds of milliseconds. EEG has a high time resolution, similar to that of a high-speed camera, and can catch physiological changes behind cognitive processes far better than other brain imaging (such as MRI or PET scanners) [15].

EEG detects brain activity directly because the brain is continually working, producing electrical activity that is very subtle (less than a 9V battery), yet measurable with the correct instrument. These small signals from the scalp surface can be picked up by EEG devices. Principal forms and well-accepted theories on how EEG signals relate to cognitive, emotional, and selective attention processing have been obtained through neuroscientific research. While techniques like MRI have a high spatial resolution, they essentially detect brain activity indirectly, necessitating a much deeper knowledge of the link between what is observed and how it relates to other variables [16].

B. INPUT PARAMETERS FOR EMOTION RECOGNITION

Automated emotion recognition is usually done by studying changes in several human bodily parameters or electric impulses in the nervous system, the comparison analysis of various input parameters are discussed in Table 3. From the comparison, it is observed that the best way to detect the emotion is EEG signals as emotion originates from the brain and its activity is captured properly to identify the state of the human.

Furthermore, recognising human emotions through biological brain signals is gaining popularity. EEG signal is a cost-effective and reliable way of assessing brain activity when compared to other input parameters. Other methods are also effective but the accuracy obtained by using EEG is better than others. And also, time series properties of the EEG signals are very effective in detecting the emotions [17].

C. CLASSIFICATION OF EEG BASED DEEP LEARNING AREAS

There are many studies emerging in the EEG deep learning areas. They can be classified as many categories are motor imagery(22%), mental work estimation (16%), sleeping state scoring (9%), seizure detection (14%), event classification (10%) and rest of the areas (13%) respectively [18].Motor imaging requires the patient to see specific muscle actions on the hands and legs/or tongue. The motor imaging applications are widely used in the events where user movements are taken as input such as video games and medical processing.Mental work estimation is determined by recording the EEG signals of the subjects during various degrees of mental work complexity. Lots of studies have been conducted among drivers and pilots of mental stress. The subject’s behavior and response time during their work.

Sleeping state pattern is calculated by recording the EEG overnight. The sleeping condition of a human is rated 1, 2, 3, or 4 by specialists. The rapid eye movement is also studied, with the ultimate goal of this type of research being to eliminate the need for qualified people to assess and comprehend patient sleep stages.In seizure detection, the patients recorded the EEG signal with the period of seizure and when they are seizure-free to find the difference in mental health. Some normal subjects EEG are also recorded to understand the difference between the both. This helps to predict the next seizure occurrence and precaution measures can be given to the patients. In event related detection, subjects are asked to view a visual representation and EEG signals are recorded during the task. Frequency range of EEG signal and its band are illustrated in table 4.

Emotion recognition related work is less focused in the previous years. The prediction of the emotion is done through various ways such as speech signals, social media sentiment analysis, behavior changes, facial expression and so on involved [19]. The unsupervised learning methods are transfer learning, self-organizing map, K means and Fuzzy C means. Many novel methods have emerged in the deep learning networks for EEG based emotion recognition. These methods perform better than machine learning mechanisms as they create a deep neural network to recognize human behavior more accurately. Many algorithms developed in DNN achieve high accuracy in Brain-Computer Interface (BCI). This review is a promising part in discussing the various research works of EEG based emotion.

TABLE 3. comparison analysis of various input features for emotion recognition

Sr.No	Input Type	Measure	Advantage	Disadvantage
1.	Electroencephalography (EEG) Signals	Electrical activity of human brain is recorded	Accurate emotion recognition due to brain signals	Wearable electrodes
2.	Electrocardiography (ECG) Signal	Function of heart is recorded	Recognizes anxiety level	Huge amount of data produced
3.	Galvanic Skin Response (GSR)	Electrical parameter of the skin	Level of Arousal is measured well	valence is not clearly explained
4.	Heart Rate Variability (HRV)	Heart rate rhythm	Measures State of relaxation or stress	Complexity of sensors
5.	Photoplethysmography (PPG)	Detect a change of microvascular blood volume in tissues	Complexity is less - only one sensor	Delay in the signals
6.	Respiration Rate (RR)	Monitoring the respiratory activity	Breathing rate represents the emotions	Functional limitations
7.	Skin temperature (SKT)	Reaction of the autonomic nervous system	Non-contact measurements	Takes lot of time
8.	Electromyography (EMG)	Measures muscle cells in the skin	Signal analysis is easy	Strict requirements for skin preparation
9.	Electrooculography (EOG)	Movement of the eyes - blinking of eyes	Rare emotions can be measured such as	Less effective technique

 https://doi.org/10.37896/pd91.5/91542

			drowsiness and concentration	
10.	Gesture Detection (GD)	Facial reaction, body movements and gesture	Explains the reaction of emotion	Huge data and tracking body movement is difficult

TABLE 4

FREQUENCY OF THE EEG SIGNALS - THE DIFFERENT BANDS, THEIR FREQUENCIES AND RESPECTIVE MEAN VALUES WITH CORRESPONDING REACTIONS OR EMOTIONS

Sr.No	**Frequency of EEG Signals in(Hz)**	**Band of EEG**	**Means of EEG**
1.	1 to 4	Delta	deep slumber (no dreams)
2.	4 to 7	Theta	when individuals are distressed emotionally
3.	8 to 15	Alpha	In relaxed state but not sleeping
4.	15.5 to 16	Beta 1	Relaxed while concentrating
5.	16.5 to 20	Beta 2	pondering, processing, and receiving data from the outside world
6.	20.5 to 28	Beta 3	enthusiasm and apprehension
7.	25 to 100	Gamma	awareness, pleasure, stress reduction, and mediation
8.	Potentially Evoked	Lambda	When light stimulates the eyes, they are activated after 100 milliseconds.(P100)
9.	Potentially Evoked	P300	300 milliseconds cause's people to see or hear what is imaged in the brain.

COMPARISON ANALYSIS OF EEG BASED EMOTION RECOGNITION USING DNN

The Neural Network works actually like a human brain, hence, deep learning is a better option for studying human emotions. Neurons are information messengers that send electrical and chemical messages within brain. The same neural network is also built. Each neuron in the brain gets activated to a specific reaction in the body. Millions of neurons are available in the brain. Each of them creates electric activity in the brain. The electrical signal is captured with the help of EEG signals. The neurons in the neural network also react the same way when input is passed through them. Basic Neural Network has different layers which is illustrated in the below given Figure 4.

 https://doi.org/10.37896/pd91.5/91542

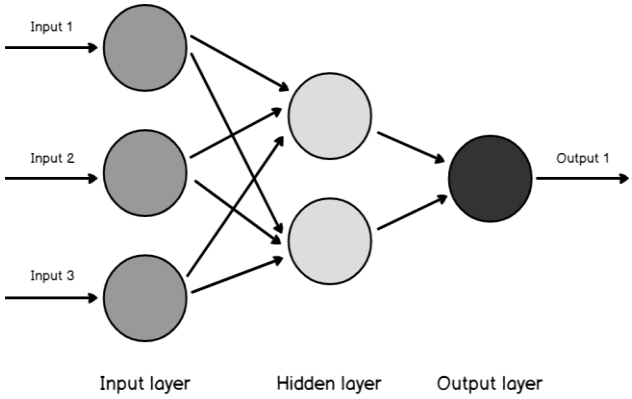

Figure 4 Basic neural network with different layers

A neuron sums up each signal multiplied by its proper weight and sends it to an activation function after receiving input from the neurons in the preceding layer of the model. The next step is passing the value to the activation function. The output is between 0 to 1 which is passed to the next layer in the neural network. The neural networks can be classified as Deep Neural networks (DNN), Recurrent Neural Network(RNN) and Convolutional Neural Network(CNN) [20].

A. DEEP NEURAL NETWORK MODEL

Artificial Neural Network (ANN) which is a relatively wide phrase that refers to any type of Deep Learning algorithm. Shallow or deep ANNs exist in ANN, when there is only one concealed layer, they are called shallow (one layer is between input and output).When there are multiple layers concealed beneath the surface, they are referred to as deep layers (mostly multiple layers are implemented). The multiple layered ANN are called Deep Neural Network (DNN) as seen in Figure 5 given below.

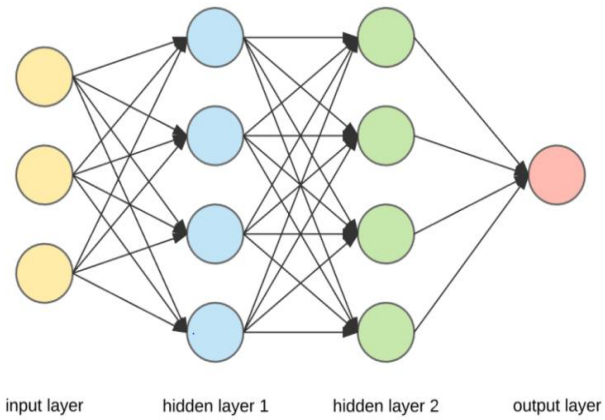

Figure 5. Architecture of Multi Layered DNN

Spicy is a tool kit for EEG filtering and preprocessing of EEG Signals. Python signal processing toolbox (MNE) is also processing EEG-specific signals. Keras Library is used for deep learning purposes. With DEEP datasets, this model achieved 82.5% accuracy [21].

A 2-D spectrogram is used to extract from raw EEG signals for each channel and the k-means algorithm is used to cluster. DEAP and SJTU SEED datasets are used for better comparison of classification accuracy. Extracted features are fed to the DNN layers for the classification. Accuracies are 77.4% and 93.8% for SJTU SEED and DEAP datasets respectively. Temporal and spatial features are used.The most consistent technique is to normalize the Common Spatial Pattern (CSP).To obtain good decoding accuracy, normalized CSP is employed to extract features.The goal of normalizing the CSP is to eliminate the impacts of noise and artifacts in the raw EEG data.CSP features were also applied in Virtual Reality (VR) control [22].

The temporal and spatial features of the EEG signals are discussed. TSception is used in this research. EEG signals are recorded with virtual reality from 18 healthy subjects. In this study, we evaluate the model's classification performance in comparison to LSTM, SVM, and EEGNet. Comparison is carried out with the different features such as temporal and spatial respectively. The best results are obtained with the SVM which has an accuracy of 86.03% and ($p<0.05$). When comparing the features separately, they perform poorly but when combined together (Temporal and Spatial feature) , desired output is obtained [23].

The non-linear pattern of EEG data makes it difficult to recognise emotions are the downside of DNNs. It can be overcome by a deep forest method. The noise in the raw EEG signals is removed using the baseline removal approach. The 2 D frame sequences are created using the spatial position. 2D frame sequences are given into a deep forest classification model that can classify EEG emotions by mining the spatial and temporal information in the signals. The DREAMER and DEAP database are involved in this method. Accuracies of 2 dimensional model- Arousal and Valence for DEAP databases are 97.69% and 97.53% respectively. Accuracies of 3 dimensional model- Arousal, Dominance and Valence for DREAMER database are 89.03%, 90.44% and 89.89% respectively [24]. Table 5 shows the comparison of various recent studies in DNN using EEG signals. MAHNOB, DEAP and SEED databases combined to do classification performance analysis. The 3 datasets are used with three different algorithms SVM, KNN and RF and the classification performances are compared with each other [25].

RECURRENT NEURAL NETWORK MODEL

In deep learning, recurrent neural networks (RNN) are networks that can process a sequence of inputs while maintaining their present state while processing the next set of inputs regardless of the sequence in which the inputs are received, traditional neural networks will process one before proceeding on to the next. Time series follows a sequential order that need to be understood to execute them. The structure of the RNN is explained with feedback in Figure 6. EEG signal is also a kind of sequential data and RNN uses sequential architectures that can be used to train better sequential data. So, RNN could perform for extracting the data and training the same features RNN can be given frequency spectrogram feature as input by converting the unprocessed EEG signals. .There are many studies that work on diagnosis of a brain related disorder like decoding of the mind which also involves emotion detection in RNNs may process sequences of inputs by using their internal state (memory). They work not only with the information that was fed, but also with relevant information from the past, implying that whatever

was previously fed and taught to the network RNNs are implemented in many areas such as sport action modeling, composition of music, image generation, forecasting financial asset, next word prediction and many more [26].

There is a feedback loop that occurs in every layer of RNN. LSTM is one of the types of the RNN. The input or information is stored in the memory of the LSTM for a long time. Special and standard units are included in LSTM networks, which are a form of RNN. A memory cell is featured in LSTM units, which can retain data for lengthy periods of time. The LSTM layers are trained with a temporal sequence of input. The input vector is spitted with a time sequence and then it is looped into the network to be trained. The working of the RNN -LSTM is clearly illustrated in Figure 7.

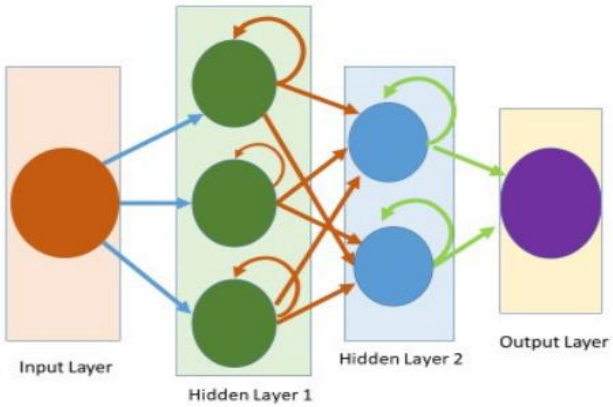

Figure 6. Structure of Recurrent Neural Network [27]

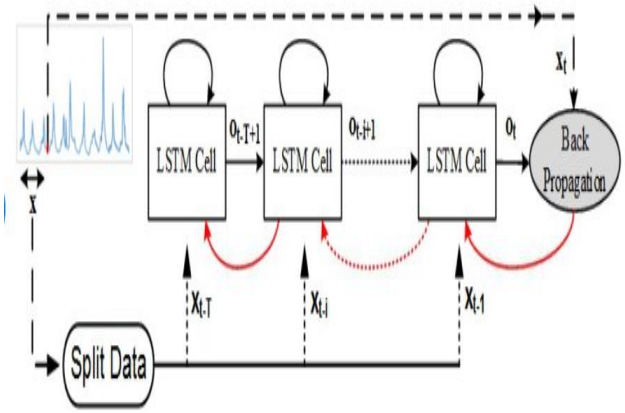

Figure 7. Structure of RNN - LSTM [27]

ISSN: 0369-8963

TABLE 5. DNN - EEG Signals based Comparison Analysis

Sr.no	Author	Input	Output	Method/ Algorithm	Performance Metric/accuracy	Content	Dataset
1.	Muhammad Adeel Asghar et al., 2019	EEG Signals 2D features Spatial Temporal	Emotion Detection	DNN	Accuracy SEED -93.8% DEAP - 77.4%	K-mean cluster algorithm is used	DEAP SEED dataset
2.	Abeer Al-Nafjan et al., 2017	Power Spectral Density (PSD)	Emotion Detection	DNN	Accuracy - 82.5%	Python.Scikit-learn toolbox Scipy for filtering	
3.	Emanuele Maiorana et al., 2020	EEG Signal Feature Temporal Spatial	Emotion Detection	DNN	Accuracy- 86.03%	SVM is compared with EEGNet, LSTM	18 subjects EEG signals recorded from MUSE EEG headband
4.	Juan Cheng et al.,2021	EEG Signal Spatial	Emotion recognition Valence Arousal Dominance	Deep Forest	Accuracy DEAP- 97.69 97.53 DREAMER- 89.03 90.44 89.89(Juan Cheng et al.,2021)	This model overcomes the shortcoming of too many hyper parameters and lots of training data	Deap and Dreamer dataset

 https://doi.org/10.37896/pd91.5/91542

ISSN: 0369-8963

						using deep forest	
5.	Muhammad Adeel Asghar et al., 2020	EEG signal	Emotion Detection	DNN	Accuracy SVM - 97.5% RF - 94.7% K-NN - 92.3%	Empirical Mode Decomposition(EMD) DIfferent Dataset and different algorithm compared	SEED DEAP MAHNOB Dataset

 https://doi.org/10.37896/pd91.5/91542

LSTM is developed to deal with the vanishing gradient problem which was there in the RNN. Hence, the performance of the LSTM is better than traditional RNN. The vanishing gradient problem is solved in this LSTM. The DEAP dataset is used with the LSTM algorithm. The 3 dimensional model (Valence- Arousal- Liking) is achieved by passing the preprocessed inputs to the LSTM network.The accuracies achieved in Valence- Arousal- Liking are 85.45, 85.65% and 87.99% respectively [27].

Studies have been conducted in RNN - LSTM based on EEG as shown in Table 6. The performance of LSTM is evaluated with comparison with other algorithms such as CNN,MLP and LSTM-CNN. The dataset has been manually collected from 10 subjects with 32 channel EEG at 350Hz. No model gave results for a single subject. Temporal and spatial features are used. The combination of both LSTM-CNN achieved high accuracy to 64.36 %. When comparing separately, the CNN performance is slightly lesser than LSTM. The accuracy for CNN is 62.09% and for LSTM is 63.61%. But the CNN got quickly trained in 45 minutes whereas the LSTM took 4 to 5 hours of time [28].

The combination of LSTM and Stack AutoEncoder (SAE) is used to improve the accuracy of 2 D model as Valence and arousal. This model is a linear EEG mixing system and Context Correlation is used to achieve higher accuracy for the model. To decompose source signals and solve the linear problem SAE algorithm is involved. The LSTM is fed with the DEAP database and optimizer algorithm also used to reduce the gradient descent. Parameter tuning is done to get a significant raise in the output. There are three comparisons involved in this study. The LSTM is compared to SVM and the accuracy of LSTM ($p<0.01$) is greater than SVM. Then, the performance of SAE is done with Independent Component Analysis (ICA) and SAE proved to be ($p<0.01$) higher than ICA. The final comparison is between the features. The classification performance for the valence and arousal are 81.10% and 74.38% respectively [29].

CNN MODELS

Convolutional Neural Network (CNN) is one of the most widely used neural networks [3, 8]. It is used in segmentation, image processing, auto correlated data, natural language processing, classification and many more. It is specially used for object recognition and emotion recognition. CNN works with large amounts of data and uses multi-channeled data.Each convolution is a sort of affine function, and they are all linear transformations. The working pattern of the CNN is explained clearly in Figure 8.

TABLE 6. RNN - EEG Signals based comparison analysis

Sr.no	Author	Input	Output	Method/ Algorithm	Performance Metric/ accuracy	Content	Dataset
1.	Jeevan Reddy Koya et al., 2019	EEG Signals	Emotion Recognition	LSTM Over CNN LSTM-CNN	Accuracy - 63.61%	No model gave result on single subject	10 subjects manually data collected

2.	Alhagry et al.,2017	Processed EEG Signals	Emotion Recognition	LSTM	Accuracy Valence 85.45, Arousal-85.65% Liking - 87.99%	Three dimensional Model	Deap dataset
3.	Xiaofen Xing et al., 2019	EEG signals	Emotion Recognition	LSTM over SVM SAE over ICA	Accuracy valence - 81.10% Arousal - 74.38%	Optimizer algorithm is used reduced gradient descent	Deap dataset

TABLE 7. CNN - EEG Signals based Comparison Analysis

Sr.no	**Author Name**	**Input**	**Output**	**Method/ Algorithm**	**Performance Metric/ accuracy**	**Content**	**Dataset**
1.	Yurui Ming et al., 2020)	EEG Signals	Emotion Recognition	CNN	Accuracy compared with various parameters	Reinforcement Learning	37 subjects manually data collected
2.	Seong-Eun Moon et al.,2018	EEG Signals	Emotion Recognition	CNN over SVM	Accuracy 99..72%	CNN-2, CNN -5 and CNN-10 layers are compared	DEAP Dataset
3.	Zhiyuan Wen et al., 2017	EEG signals	Emotion Recognition	LSTM over SVM SAE over ICA	Accuracy valence - 77.98% Arousal - 72.98%	Pearson correlation Coefficient is used to rearrange EEG signals	DEAP Dataset
4.	Tengsei Song et al., 2020	EEG signals	Emotion Recognition	CNN over SVM, DBN	Accuracy - 90.4% DREAMER Valence - 86.23%	System dynamically learns relationship between channels in	SEED DREAMER Datasets

 https://doi.org/10.37896/pd91.5/91542

					Arousal - 84.54% Dominance - 85.02%	adjacency matrix	
5.	J.X.Chen et al., 2019	EEG signals Spatial Temporal feature	Emotion Recognition	CNN over BT, SVM,LDA, BLDA	Accuracy Arousal - 85.65% Valence - 85.45% Likeness - 87.99%	Three different CNN is compared with other algorithms	DEAP Dataset
6.	Yilong Yang et al., 2018	EEG signals Spatial Temporal feature	Emotion Recognition	CNN	Accuracy Arousal - 91.03% Valence - 90.80%	Inter-channel correlation is mined using CNN	DEAP Dataset

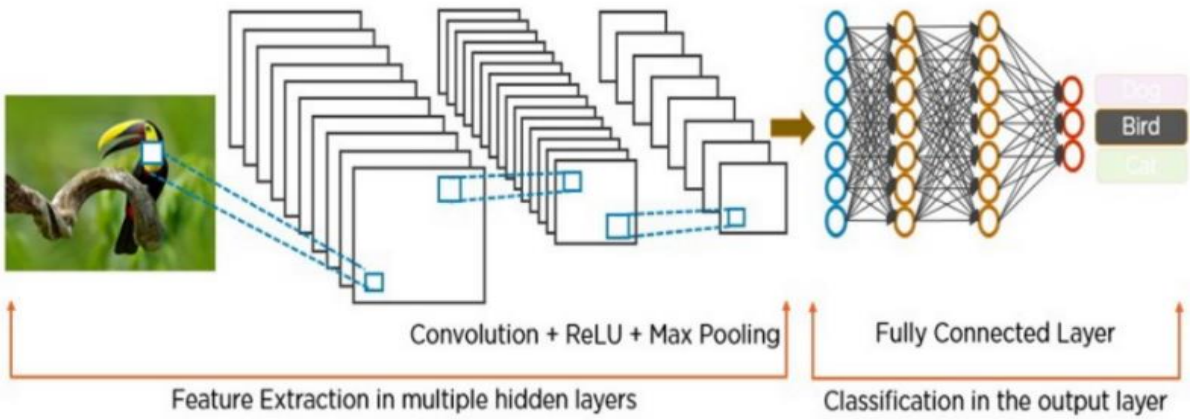

Figure 8. Structure of CNN

Manual feature engineering requires a lot of effort and has transformation problems. Hence there is a need for a better model with end to end solutions. CNN is used for emotion recognition with DEAP dataset. Two dimensional valence and arousal model is used .The accuracy of the valence and arousal are 77.98% and 72.98% respectively [30]. Model efficiency is evaluated using various permutations of convolutional layers. Various CNN networks, including CNN-2, CNN-5, and CNN-10, are employed. CNN-2 stands for a single CNN layer plus a maximum pooling layer. CNN-5 can be denoted as 3 CNN layers and two max pooling layers [3, 8]. Finally, CNN-10 has five CNN layers and five max pooling layers. Table 7 depicts the comparison analysis of recent studies on CNN based EEG signals.

Theano is used to implement the CNN layers. The loss is reduced using Adam algorithm for cross-entropy function. The Tesla K80 GPU is used to train the network. The performance metrics

ISSN: 0369-8963
for CNN is better than SVM and the accuracy is 99.72% for CNN [31].To optimize the hyper parameter, Gradient - Priority Particle Swarm Optimization(GPSO) is replaced with the Particle Swarm Optimization(PSO) method [32]. Priority concept is used by GPSO to achieve selective optimization. This is done by gradient penalties. 15 healthy students who were right-handed of age 20 to 24 were shown the famous film clips in the category of sadness, happiness and fear. EEG signals are recorded from them with the help of ESI NeuroScan System with 40 electrodes. These signals were preprocessed through the EEGLAB toolbox. The noise is removed with the help of ICA from the Signals. his paper, compare the performance of CNN algorithm over SVM and DBN. The higher accuracy obtained is 91.01% [33].

Interesting studies have been done with the car driving with EEG signals. The research talks about drowsiness and driving responses. Reinforcement learning is achieved in deep learning. 37 subjects drive in a simulated highway for 90 minutes with an EEG cap with 32 electrodes and Scan SynAmp2 Express Sys is used. Response Time is measured with temporal and spatial features. CNN model is used to predict drowsiness and response time after drowsiness. Rooted Mean Square Errors(RMSE) is measured response time and predicted response time [34].The short-time Fourier Transform(STFT) [25, 26] is used to plot the Electrode Frequency Distribution Map(EFDM) from the EEG signals. The features are automatically extracted from the EEG images. The grayscale image is treated with EFDM to get the `2 D convolution operation. This model tries to aim at shortcoming of fewer samples. The CNN model is trained with SEED dataset and the accuracy for the same is 4.5 % above other models. The pre-trained model is given a few samples of the DEAP dataset and the accuracy obtained is 82.84% [35].

B. DISCUSSIONS

This study has discussed the different techniques in deep learning and the corresponding studies that have been carried out in those techniques. The bais (weights) and activation function are the base for the ANN system. ANN artificially reconstructs the human brain as a neural network.When it makes a mistake, it returns to its original state and "modifies" its thinking, just like a human might. In an ANN, the "layers" are the rows of data points hosted by neurons sharing the same neural network. To learn, ANN employs the use of weights. The weights of ANN are changed after each loop over the neuron. The weights are subsequently adjusted by ANN according to the accuracy assessed by a "cost function."CNN, on the other hand, is devoid of both neurons and weights. CNN applies numerous layers to images and uses filters to interpret visual data. The three levels are the math layer, ReLU layer, and fully connected layer. These layers are in charge of decoding. The ability to interpret temporal information or data that arrives in sequences, such as a sentence, is the fundamental distinction between CNN and RNN when they are compared. Furthermore, convolutional and recurrent neural networks are utilized for quite different purposes, and their topologies differ to suit those differences. CNNs employ filters within convolutional layers to modify data. CNNs employ filters within convolutional layers to modify data. Table 8 compares various neural networks.

 https://doi.org/10.37896/pd91.5/91542

ISSN: 0369-8963

TABLE 8. Comparison Analysis of ANN, RNN AND CNN

Types of Data	**Artificial Neural Network (ANN)**	**Recurrent Neural Network (RNN)**	**Convolutional neural networks (CNN)**
Description	Many perceptrons or neurons make up each layer of an Artificial Neural Network. For this reason, we refer to ANN as a Feed-Forward Neural Network, meaning that the inputs are only processed forwards. [6]	A recurrent neural network (RNN) is a type of artificial neural network in which nodes are connected in a directed graph that follows a temporal sequence [4].	Convolutional neural networks (CNNs) are a type of deep neural network used to analyze visual information [3, 8].
Data types can be used	Table, Text inputs	Sequential input	Image input
Type of Input	Feed-forward type of Network (only one direction input)	Has internal memory to remember the precious input and process accordingly	Feed-forward type of Network (only one direction input)
Sharing of Parameter	Does not share	Does share the parameter	Does share the parameter
Occurrence of Vanishing Gradient and Exploding Gradient	Occurs	Occurs	Occurs
Recurrent Input	Does not occur	Occurs	Does not occur
Input with a fixed length	Takes fixed input and gives fixed length output	Takes arbitrary inputs and gives arbitrary output	Takes fixed input and gives fixed length output
Spatial Input	Processes temporal data	Processes temporal data	Processes spatial data
Performance of the Network	Less is performance when compared with RNN and CNN	RNN has less feature compatibility than CNN	CNN's performance is high when compared to ANN and RNN

 https://doi.org/10.37896/pd91.5/91542

ISSN: 0369-8963

Area of Implementation	Computer Vision and facial recognition	Speech analysis, Text analysis, Language Translation and Natural Language Processing	Face Detection, Medical Image Processing, Image Classification, Image Recognition, Computer Vision
Advantage	Fault tolerance is high	Every information is stored	High performance in image and recognition inputs
Disadvantage	Hardware dependence	Vanishing and Exploding Gradient	Large training data is required

EVALUATION RESULTS

A. DATASET USED

In this research methodology, the EEG signal is collected by using an EEG Signal recorder and there are 14 channels for collecting EEG across the entire brain. The 12 participants participated in the experiment and subjects were shown different movie clips. The EEG signals are recorded when subjects were watching the different movie clips for 2 minutes. Various emotion related videos are shown to the subjects such as happy and unhappy scenes. The electrodes are placed around the head according to the international 10 – 20 system [15]. The selected EEG [8] channels are AF4, F4, F8, FC6, T8, P8, O2, O1, P7, T7, FC5, F7, F3 and AF3 [13] are used for processing. These channels are different locations of the electrode placed on the skull. The electrodes are potentials of the brain which are produceds with various emotions. These emotions are stimulated by watching videos. This is a real time dataset conducted in an EEG lab. The table 9 describes the attributes of the dataset used in this study.

B. DATA PREPROCESSING

Basically EEG signals are very noisy due different factors such as eye blink, body movement and heartbeat. Band pass filters can be used to remove or filter the noise from the EEG signal. In EEG signals, emotion related frequencies are available below 40 Hz. Hence, a low band pass filter is used to filter the band less than 40 Hz. The preprocessed dataset is shown in Figure 9 and the preprocessed EEG signal is plotted using the Matplot function of python in Figure 10. The preprocessed data is sent to the feature extraction module where the required features for emotion detection are used.

TABLE 9. DETAILS OF DATASET

Source	File Type	No of Files	Total no of features	No of features extracted	Size of the file
Real Time Dataset	CSV	24	420	100	5 MB (each file)

 https://doi.org/10.37896/pd91.5/91542

ISSN: 0369-8963

	# mean_0_a	mean_1_a	mean_2_a	mean_3_a	mean_4_a
0	4.620	30.3	-356.0	15.60	26.3
1	28.800	33.1	32.0	25.80	22.8
2	8.900	29.4	-416.0	16.70	23.7
3	14.900	31.6	-143.0	19.80	24.3
4	28.300	31.3	45.2	27.30	24.5
...	...	...	...	...	...
2127	32.400	32.2	32.2	30.80	23.4
2128	16.300	31.3	-284.0	14.30	23.9
2129	-0.547	28.3	-259.0	15.80	26.7
2130	16.800	19.9	-288.0	8.34	26.0
2131	27.000	32.0	31.8	25.00	28.9

2132 rows × 2549 columns

Figure 9. Preprocessed Dataset

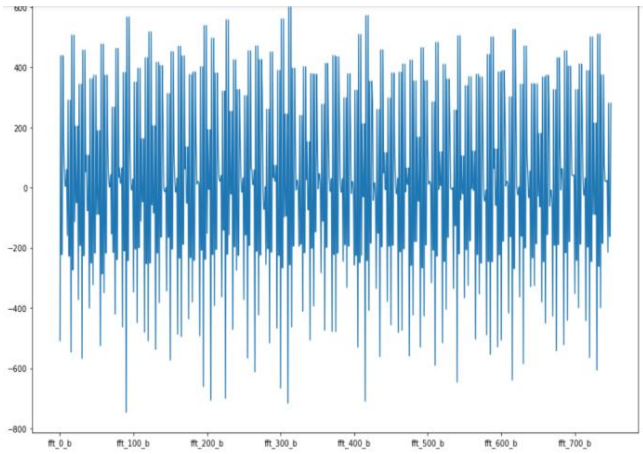

Figure 10. Preprocessed EEG Signals

C. FEATURE EXTRACTION

The feature extraction module is the next one which works on the necessary features for identification of emotions. The proposed model works with the features extracted from the dataset. Extracting features from EEG signals is a difficult due to its nonlinear properties. Since, research with EEG signals [8] for emotion recognition is less conducted. The time-frequency analysis is employed to obtain features of time and frequency domain. Power spectrum features, wavelet energy entropy features, statistical features and Horthy features are extracted from EEG signals to obtain the different emotions [21].

Short-Time Fourier transform (STFT) [25, 26]

STFT is used to extract the frequency features from EEG signals with time concern. These features can be used to control and perform various brain related tasks using BCI. Hence, STFT is used to extract different features from different bands of EEG signals.

STFT is defined as follows:

 https://doi.org/10.37896/pd91.5/91542

ISSN: 0369-8963

$$\sum_{k=-\infty}^{\infty} X(k)\omega(k-nR)e^{-j\omega m} \quad (1)$$

where x(k) is the input signal of the time n,ω(n) is a window function whose length is m , j is the Bias value[4].

Statistical features

To recognize the emotion, the time-series of EEG signals which has a set of signal statistics data are used. From EEG signals, five statistical features are extracted for each channel producing 70-dimensional (14(channels) x 5(statistics) = 70). This statistical feature vector is used to characterize the EEG signals.

Mean from Statistical Feature

$$y=\frac{1}{M}\sum_{M}^{m=1} Y(m) \quad (2)$$

Standard Deviation from Statistical Feature

$$\sigma_y = \sqrt{\frac{1}{M}\sum_{M}^{m=1} Y(m-\mu_y)^2} \quad (3)$$

Mean of absolute values of first difference [45]

$$\theta_y = \frac{1}{M-1}\sum_{M}^{m=1} |Y(m-\mu_y)^2| \quad (4)$$

Mean of absolute values of second difference [45]

$$\gamma_y = \frac{1}{M-2}\sum_{M-2}^{m=1} |Y(m+2)-Y(m)| \quad (5)$$

Mean of absolute values of second difference of normalized

$$\gamma_y - \frac{1}{M-2}\sum_{M-2}^{m=1} |Y(m+2)-Y(m)| \quad (6)$$

These parameters can be extracted from the statistics features and it can be used to recognize the emotions of human beings more accurately.

 https://doi.org/10.37896/pd91.5/91542

Hjorth parameter

TABLE 10. The Hjorth Feature

Parameter of Hjorth Feature	**Notation**
Activity of Parameter	var(x(i))
Mobility of Parameter	$\sqrt{\frac{\mathrm{var}(x'(i))}{\mathrm{var}(x(i))}}$
Complexity of Parameter	$\frac{\mathrm{mobility}(x'(i))}{\mathrm{mobility}(x(i))}$

The Hjorth parameter is one way of expressing statistical features of a signal in the time domain, and it has three types of parameters shown in Table 3. Additionally, they can be used to produce lower computing complexity [5].

Wavelet entropy feature

Entropy is one of the powerful algorithms developed for nonlinear analysis of EEG signals [37] with the advancement of nonlinear dynamics [4]. The discrete wavelet transform can automatically adjust the window size based on the frequency. The discrete wavelet transform is defined mathematically as follows:

$$Wx(i,j)=\sum_j \quad x\varphi_{i,j}(t) \qquad (7)$$

Wx(j, k) are the convoluted data with scale function i, j(t): where x(t) is the EEG data, i is the number of wavelet decomposition layers, and Wx(i, j) are the convoluted data with scale function i , j(t):

$$\varphi_{i,j}(t) = \frac{1}{2^j}\left(\frac{t}{2^j} - j\right) \qquad (8)$$

Classifier

We have used different classifier models in deep learning for classifying whether the person is positive (Happy), Neutral (relaxed) and negative(Sad) in python 3.6.9 version. The classifier is implemented using Keras library. First step involved in the process is capturing an EEG signal from the subject using the EEG recorder. The received EEG signal is pre-processed using filters to remove noise. The next step involved is the extraction function. The proposed model has an input layer, a hidden layer, and an output layer, which provides the results explained in Figure 11. We trained the model using a dataset that is divided into training and test samples as 7:3. The model consists of different layers and a summary of the model, which is the output is shown in Figure 12 below.

Classifier Parameter Analysis

The calculation of the parameters and summary of the neural network for this model is discussed in Figure 11. The summarization (parameter or neurons) of the model is given from the output to the input in stepwise:

1. Parameters in the sixth layer (dense_16 (Dense)) = ((Layer $_{(Current)}$ (c) * Layer $_{(Previous)}$ (p) + 1 * c))

= (3 * 100) + 1 * 3

= (300) + 3 = 303 (Parameters in output layer)

2. Parameters in the fifth layer (dense_15 (Dense)) = ((Layer$_{(Current)}$ (c) * Layer$_{(Previous)}$ (p) + 1* c))

= (100 * 20288) + 1 * 100

= 2028800 + 100 = 2028900 (Parameters in dense layer)

3. Third layer (Max_pooling) and fourth layer(Flatten) has no parameters

4. Parameters in the second layer (conv1d_3) = (output_channel *(input_channel * window_size + 1))

= (16 * (16 * 3 +1))
= (16 *(48 + 1))
= (16 * (49))
= 784 (Parameters in convolutional second layer)

5. Parameter in the first layer (conv1d_2) = (output_channel *(input_channel * window_size + 1))

= ((16 * (10 + 1))

= (16 * 11)

= 176 (Parameters in convolutional first layer)

6. Total number of trainable parameters in the different layers = (conv1d_2 + conv1d_3 + dense_15 + dense_16)

= (176 + 784 + 2028900 +303)

 https://doi.org/10.37896/pd91.5/91542

= 2030163 (Total number of trainable parameter)

```
Model: "sequential_5"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv1d_2 (Conv1D)           (None, 2539, 16)          176

 conv1d_3 (Conv1D)           (None, 2537, 16)          784

 max_pooling1d_1 (MaxPooling  (None, 1268, 16)         0
 1D)

 flatten_1 (Flatten)         (None, 20288)             0

 dense_15 (Dense)            (None, 100)               2028900

 dense_16 (Dense)            (None, 3)                 303

=================================================================
Total params: 2,030,163
Trainable params: 2,030,163
Non-trainable params: 0
```

Figure 11. Model Summary

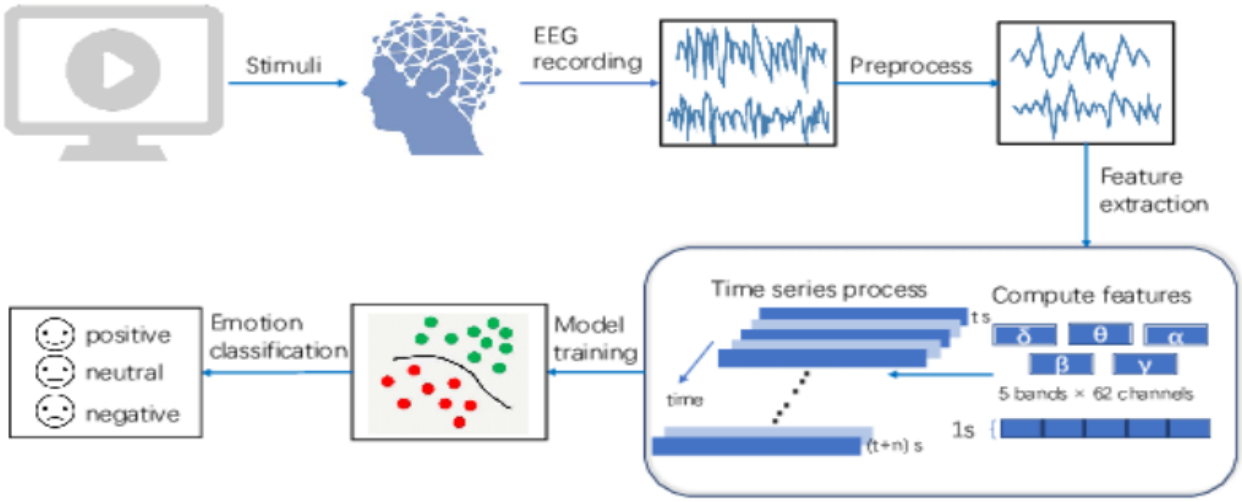

Figure 12. The EEG emotion recognition architecture **[**26**]**

Hyperparameter Optimization

The variables that determine the neural network structure and working are called as hyper parameters of the classifier. The fine tuning of the parameters makes the classifier perform better. The different hyper parameters used in this classifier are explained in Table 11 .Optimizers are methods or approaches to reduce loss by adjusting parameters of your neural network such as weights and learning rate [23]. The Adam optimizer is used in this paper to optimize the neural network output. The learning rate is the difference between the loss gradient of the network and its weight. An epoch is a complete cycle of forward and backward passage of all training variables. The number of variables passed in one epoch is called the batch size. Thus, by tuning these variables or parameters, the classifier is performed effectively and efficiently.

 https://doi.org/10.37896/pd91.5/91542

ISSN: 0369-8963

Performance Analysis of different deep learning algorithms

The execution of CNN classifier [36] is shown in Figure 13 and accuracy of this model is 97.82 and validation loss is 1.16 which is very less compared to the other models.

TABLE 11. Hyper parameters used in the classifier

Sr.No	Hyperparameter Name	Units
1.	Neurons Input	(2548, 16)
2.	Neurons output	(2548, 3)
3.	Optimizer	Adam
4.	Batch_Size	35
5.	Epoch	40
6.	Learning Rate	0.01

```
Epoch 00014: val_accuracy did not improve from 0.97812
Epoch 15/40
47/47 [==============================] - 552s 12s/step - loss: 0.0000e+00 - accuracy: 1.0000 - val_loss: 1.1602 - val_accuracy: 0.9781

Epoch 00015: val_accuracy did not improve from 0.97812
Epoch 16/40
47/47 [==============================] - 553s 12s/step - loss: 0.0000e+00 - accuracy: 1.0000 - val_loss: 1.1602 - val_accuracy: 0.9781

Epoch 00016: val_accuracy did not improve from 0.97812
Epoch 17/40
47/47 [==============================] - 552s 12s/step - loss: 0.0000e+00 - accuracy: 1.0000 - val_loss: 1.1602 - val_accuracy: 0.9781

Epoch 00017: val_accuracy did not improve from 0.97812
Epoch 18/40
47/47 [==============================] - 551s 12s/step - loss: 0.0000e+00 - accuracy: 1.0000 - val_loss: 1.1602 - val_accuracy: 0.9781

Epoch 00018: val_accuracy did not improve from 0.97812
Epoch 19/40
47/47 [==============================] - 553s 12s/step - loss: 0.0000e+00 - accuracy: 1.0000 - val_loss: 1.1602 - val_accuracy: 0.9781

Epoch 00019: val_accuracy did not improve from 0.97812
Epoch 20/40
47/47 [==============================] - 555s 12s/step - loss: 0.0000e+00 - accuracy: 1.0000 - val_loss: 1.1602 - val_accuracy: 0.9781

Epoch 00020: val_accuracy did not improve from 0.97812
Epoch 00020: early stopping
```

Figure 13. Validation accuracy and loss calculation of CNN Classifier

 https://doi.org/10.37896/pd91.5/91542

The accuracy of the model is described in Figure 14. Fig. A plot of the model's accuracy is compared with the train data and the test data. The X-axis consists of the number of epochs and the Y-axis consists of the percentage of accuracy to 1.0. As the number of epochs increases, so does the accuracy. For the train data, the achieved accuracy is 1.0, while for the test data, around 10 epochs, it reaches a stagnant accuracy of 0.98. In this situation, the early stopping function should neither overfit nor underfit the model to properly train the system. Figure 15 shows that the model loss is compared between the train data and the test data for the CNN model. The loss of training data decreases slowly, while it is low for test data from the beginning because training is already done with train data. Finally, the loss of the training data is reduced to zero, and the loss of the test data is 1.16, respectively.

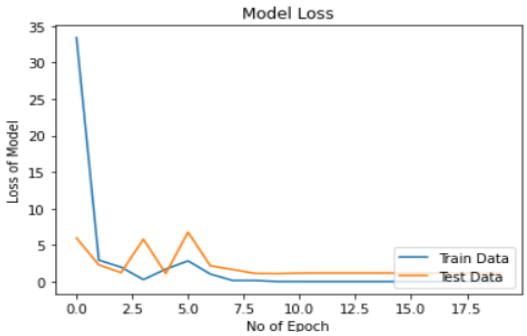

Figure 14. Analysis of model accuracy of CNN classifier

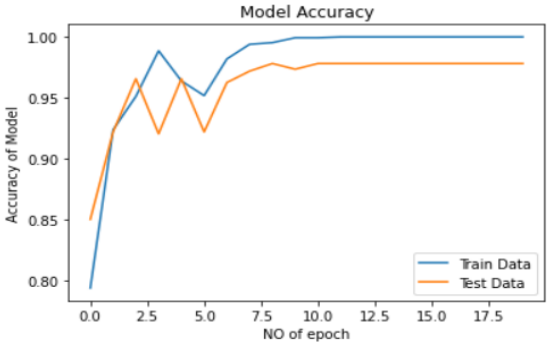

Figure 15. Analysis of model loss of CNN classifier

The Confusion matrix is plotted for the proposed model in the Figure 16 has actual output against predicted output. The labels for emotions positive, negative and neutral are represented as 0, 1,2 respectively. The other performance metrics are also plotted in the Figure 17 as precision, recall and F1-score respectively.

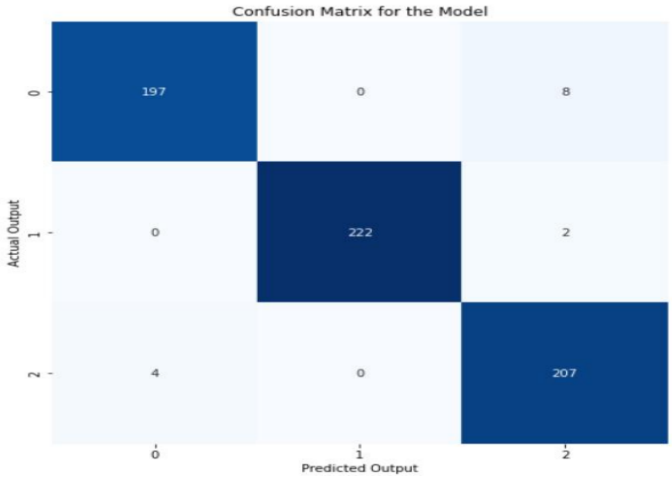

Figure 16. Confusion matrix for the accuracy of the CNN Classifier

```
Classification Report For Different Performance Metrics:
----------------------
              precision    recall  f1-score   support

           0       0.98      0.96      0.97       205
           1       1.00      0.99      1.00       224
           2       0.95      0.98      0.97       211

    accuracy                           0.98       640
   macro avg       0.98      0.98      0.98       640
weighted avg       0.98      0.98      0.98       640
```

Figure 17. Performance metrics for the CNN classifier

TABLE 12. Performance comparison of different models

Sr.no	Method/ Classifier	Performance Metric/accuracy	Validation Loss
1.	SVM	86.03	1.80
2.	Deep Forest	90.25	1.63
3.	DNN (Xiaofen Xing et al., 2019)	88.53	1.46
4.	LSTM (Jeevan Reddy Koya et al., 2019)	89.00	1.75
5.	RNN (Xiaofen Xing et al., 2019)	92.30	2.00
6.	c	97.82	1.16

For the evaluation, different existing deep learning algorithms are implemented with the same dataset. The performance and loss are also calculated. The accuracy and validation loss are also shown in Table 9.The final accuracy for the CNN classifier [36] is 0.98 which is higher than other previous models. CNN is performing better in the accuracy of the image classification, image recognition and medical image processing than other two techniques. The accuracy of challenging classification tasks that require understanding abstract concepts in images has been the subject of substantial research in Convolution models. Another reason why CNN performs well than most neural networks. The best part is that no feature extraction is required. The system learns to extract features and the key concept of CNN [3, 8] is that it generates invariant features by convolution of images and filters which are then passed to the next layer. The elements in the next layer are concatenated with different filters to obtain more invariant and abstract elements, and the process is repeated until the final element. Deep convolutional networks are

very adaptive and work well with image data, which is an important aspect. Convolutional layers take advantage of the fact that an interesting pattern can appear in any part of an image, and regions are continuous blocks of pixels. CNN's potential to discover meaningful features from raw data is one of the reasons why researchers are excited about it. CNN [3, 8] can now extract information features from images, avoiding the requirement of manual image processing. Thus, CNN is a more powerful and accurate approach to solving classification problems in general than ANN and RNN. All these aspects of CNN allow researchers to choose this algorithm for further development and improvement of image processing and emotion recognition.

CONCLUSION

Deep learning is a blossoming technology that can be used in many areas of research. Due to the increase in data that is generated through social media, smartphone usage and many more. This increase in data paves the way for the use of deep learning algorithms rather than machine learning. Recognizing emotions has become a necessity due to the recent situation. Therefore, emotion recognition is the focus of this article, especially with EEG signals. In principle, EEG signals have non-linearity. As a result, less work is done in this area. Thus, this paper sheds light on a potential research area where research can be conducted. Comparison of different techniques in deep learning. CNN is a better solution due to its ability to see images as data. When it comes to emotion recognition, CNN outperforms ANN and RNN. In particular, CNN performs well with emotion recognition using EEG signals. The accuracy of CNN exceeds that of RNN and ANN. Experimental results also prove that CNN performs better than other methods. The accuracy of the CNN method of the above data set is 98% and the loss is also relatively less than others. Experimental results are based on a database of EEG signals. In future work, a multi-modal HMI system can be considered for emotion recognition, such as other parameters, as shown in Table 3.

DECLARATION OF COMPETING INTEREST

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

REFERENCES

[1] USMAN, S.M., KHALID, S. AND BASHIR, Z., 2021. Epileptic seizure prediction using scalp electroencephalogram signals. Biocybern. Biomed. Eng, 41(1), pp.211-220.

[2] KANG, X., HANDAYANI, D.O.D. AND YAACOB, H., 2021, February. Comparison between Butterworth Bandpass and Stationary Wavelet Transform Filter for Electroencephalography Signal. In IOP Conference Series: Mater. Sci. Eng (Vol. 1077, No. 1, p. 012024). IOP Publishing.

[3] OPAŁKA, S., STASIAK, B., WOSIAK, A., DURA, A. AND WOJCIECHOWSKI, A., 2021, June. EEG-Based Emotion Recognition–Evaluation Methodology Revisited. In International Conference on Computational Science (pp. 525-539). Springer, Cham.

[4] AFSHAR, S., BOOSTANI, R. AND SANEI, S., 2021. A combinatorial deep learning structure for precise depth of anesthesia estimation from EEG signals. IEEE Journal of Biomedical and Health Informatics, 25(9), pp.3408-3415.

[5] LI, F., TAI, Y., YU, H., ZHOU, H. AND ZHANG, L., 2021. Research Status of Motor Imagery EEG Signal Based on Deep Learning. 3D Imaging Technologies—Multi-dimensional Signal Processing and Deep Learning, pp.11-17.

[6] RASHID, M., SULAIMAN, N., PP ABDUL MAJEED, A., MUSA, R.M., BARI, B.S. AND KHATUN, S., 2020. Current status, challenges, and possible solutions of EEG-based brain-computer interface: a comprehensive review. Front. Neurorob, p.25.

[7] LI, G., LEE, C.H., JUNG, J.J., YOUN, Y.C. AND CAMACHO, D., 2020. Deep learning for EEG data analytics: A survey. Concurr Comput, 32(18), p.e5199.

[8] HAN, H., BYUN, K. AND KANG, H.G., 2018, October. A deep learning-based stress detection algorithm with speech signal. In proceedings of the 2018 workshop on audio-visual scene understanding for immersive multimedia (pp. 11-15).

[9] SONG, S.H. AND KIM, D.K., 2017. Development of a stress classification model using deep belief networks for stress monitoring. Healthc Inform Res, 23(4), pp.285-292.

[10] MADHAVI, I., CHAMISHKA, S., NAWARATNE, R., NANAYAKKARA, V., ALAHAKOON, D. AND DE SILVA, D., 2020, September. A deep learning approach for work-related stress detection from audio streams in cyber physical environments. In 2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA) (Vol. 1, pp. 929-936). IEEE.

[11] ISMAIL, N.H., LIU, N., DU, M., HE, Z. & HU, X., 2020. A deep learning approach for identifying cancer survivors living with post-traumatic stress disorder on Twitter. BMC Medical Informatics and Decision Making, 20(4), pp.1-11.

[12] LIU, Q. AND LIU, H., 2021. Criminal psychological emotion recognition is based on deep learning and EEG signals. Neural. Comput. Appl., 33(1), pp.433-447.

[13] ATMAJA, B.T. AND AKAGI, M., 2021. Two-stage dimensional emotion recognition by fusing predictions of acoustic and text networks using SVM. Speech Commun., 126, pp.9-21.

[14] GADHOUMI, K., LINA, J.M., MORMANN, F. AND GOTMAN, J., 2016. Seizure prediction for therapeutic devices: A review. J. Neurosci. Methods, 260, pp.270-282.

[15] ACHARYA, U.R., 2018. Oh SL, Hagiwara Y, tan JH, Adeli H. Deep convolutional neural networN for the automated detection and diagnosis of seizure using EEG signals. comput Biol med, 100, pp.270-278.

[16] MATHUR, N., GUPTA, A., JASWAL, S. AND VERMA, R., 2021. Deep learning helps EEG signals predict different stages of visual processing in the human brain. Biomed. Signal Process, 70, p.102996.

[17] GAO, Z., DANG, W., WANG, X., HONG, X., HOU, L., MA, K. AND PERC, M., 2021. Complex networks and deep learning for EEG signal analysis. Cogn Neurodyn, 15(3), pp.369-388.

[18] JENKE, R., PEER, A. AND BUSS, M., 2014. Feature extraction and selection for emotion recognition from EEG. IEEE Trans. Affect. Comput, 5(3), pp.327-339.

[19] WEN, Z., XU, R. AND DU, J., 2017, December. A novel convolutional neural networks for emotion recognition based on EEG signal. In 2017 International Conference on Security, Pattern Analysis, and Cybernetics (SPAC) (pp. 672-677). IEEE.

[20] ROSA, G.A., 2022. Understanding and predicting interruptions index of medium voltage customers using fully connected networks.

[21] AL-NAFJAN, A., HOSNY, M., AL-WABIL, A. AND AL-OHALI, Y., 2017. Classification of human emotions from electroencephalogram (EEG) signal using deep neural network. Int J Adv Comput Sci Appl, 8(9).

[22] ASGHAR, M.A., KHAN, M.J., AMIN, Y., RIZWAN, M., RAHMAN, M., BADNAVA, S. AND MIRJAVADI, S.S., 2019. EEG-based multi-modal emotion recognition using bag of deep features: An optimal feature selection approach. Sensors, 19(23), p.5218.

[23] ALHAGRY, S., FAHMY, A.A. AND EL-KHORIBI, R.A., 2017. Emotion recognition based on EEG using LSTM recurrent neural network. Int J Adv Comput Sci Appl, 8(10).

[24] LOTTE, F., BOUGRAIN, L., CICHOCKI, A., CLERC, M., CONGEDO, M., RAKOTOMAMONJY, A. AND YGER, F., 2018. A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update. J. Neural Eng, 15(3), p.031005.

[25] ASGHAR, M.A., KHAN, M.J., RIZWAN, M., MEHMOOD, R.M. AND KIM, S.H., 2020. An innovative multi-model neural network approach for feature selection in emotion recognition using deep feature clustering. Sensors, 20(13), p.3765.

[26] WANG, Y., QIU, S., LI, J., MA, X., LIANG, Z., LI, H. AND HE, H., 2019, July. EEG-based emotion recognition with similarity learning network. In 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 1209-1212). IEEE.

[27] NAGABUSHANAM, P., THOMAS GEORGE, S. AND RADHA, S., 2020. EEG signal classification using LSTM and improved neural network algorithms. Soft Computing, 24(13), pp.9981-10003.

[28] GLADWIN, A., BASKAR, R. AND GANAPATHY, C., 2012. Effect of micro and nano-particles in M-Sand cement mortar. Int. J. Struct. Civil Eng, 1, pp.71-75.

[29] GLADWIN, A., BASKAR, R. AND GANAPATHY, C., 2012. Study of marble power as partial replacement for cement mortar using M-Sand, IJCSEIERD, 1, pp.21-29.

[30] WEN, Z., XU, R. AND DU, J., 2017, December. A novel convolutional neural networks for emotion recognition based on EEG signal. In 2017 International Conference on Security, Pattern Analysis, and Cybernetics (SPAC) (pp. 672-677). IEEE.

[31] TORRES, E.P., TORRES, E.A., HERNÁNDEZ-ÁLVAREZ, M. AND YOO, S.G., 2020. EEG-based BCI emotion recognition: A survey. Sensors, 20(18), p.5083.

[32] RAJESWARI, I.B. AND PATIL, D., Detection of Intrusion and Recovery for Smartphones using Cloud Services.

[33] GARCIA-MORENO, F.M., BERMUDEZ-EDO, M., RODRÍGUEZ-FÓRTIZ, M.J. AND GARRIDO, J.L., 2020, July. A CNN-LSTM deep Learning classifier for motor imagery EEG detection using a low-invasive and low-Cost BCI headband. In 2020 16th International Conference on Intelligent Environments (IE) (pp. 84-91). IEEE.

[34] MING, Y., WU, D., WANG, Y.K., SHI, Y. AND LIN, C.T., 2020. EEG-based drowsiness estimation for driving safety using deep Q-learning. IEEE Trans. Emerging Top. Comput. Intell, 5(4), pp.583-594.

[35] WANG, F., WU, S., ZHANG, W., XU, Z., ZHANG, Y., WU, C. AND COLEMAN, S., 2020. Emotion recognition with convolutional neural network and EEG-based EFDMs. Neuropsychologia, 146, p.107506.

[36] MING, Y., WU, D., WANG, Y.K., SHI, Y. AND LIN, C.T., 2020. EEG-based drowsiness estimation for driving safety using deep Q-learning. IEEE Transactions on Emerging Topics in Computational Intelligence, 5(4), pp.583-594.

[37] ALEX, A.G., GEBREHIWET, T., KEMAL, Z., 2021. M-Sand cement mortar with partial replacement of alpha phase nano alumina. Journal of Building Pathology and Rehabilitation, 6(1), pp.1-6.