Emotion Recognition using EEG Signal and Deep Learning Approach

by

Sayedi Hassan Bin Islam

19341036

Md. Quamar Mehdi

19141036

Bhuiyan Yash Rohan

19341031

Syed Atif Imtiaz Mahmood

14201015

A thesis submitted to the Department of Computer Science and Engineering

in partial fulfillment of the requirements for the degree of

B.Sc. in Computer Science

Department of Computer Science and Engineering

Brac University

August 2019

Declaration

It is hereby declared that

1. The thesis submitted is my/our own original work while completing degree at Brac University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

Sayedi Hassan Bin Islam	Md. Quamar Mehdi
19341036	19141036
Bhuiyan Yash Rohan	Syed Atif Imtiaz Mahmood
19341031	14201015

Approval

The thesis/project titled "Emotion Recognition Using EEG Signal And Deep Learning Approach" submitted by

1. Sayedi Hassan Bin Islam (19341036)
2. Md. Quamar Mehdi (19141036)
3. Bhuiyan Yash Rohan (19341031)
4. Syed Atif Imtiaz Mahmood (14201015)

Of Summer, 2019 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on 2019.
Signatures of Supervisory Committee Members:

Supervisor :

Mohammad Zavid Parvez, PhD

Assistant Professor
Department of Computer Science and Engineering
Brac University

Session Coordinator :

Jia Uddin, PhD

Associate Professor
Department of Computer Science and Engineering
Brac University

Head of Department: (Chair)

Mahbubul Alam Majumdar, PhD

Professor and Chairperson
Department of Computer Science and Engineering
Brac University

Abstract

Emotion is a mental state, which originates in the brain and is closely related to the nervous system. Emotion can be defined as a feeling expressed through, or detectable by voice intonation, facial expression body language, as response from one's mood relationship with others and most importantly the circumstance they are in. Although, Brain Computer Interface (BCI) are being developed to find a better human-machine interaction system using brain activity and it is frequently implemented by Electroencephalogram (EEG) signals. EEG is a well established approach to measure the brain activities which can be analyzed and processed to distinguish different emotions. In this thesis, we present an approach to classify human emotions using EEG signal by Convolutional Neural Network(CNN). In our model, we use the Dataset for Emotion Analysis using Physiological signals (DEAP) dataset, a benchmark for emotion classification research, to transform the EEG signal from time domain to frequency domain and extract the features to classify the emotions. Emotion can be classified based on the two dimensions of valence and arousal. Previous researches have used fewer channels and participants. Our approach which was carried out on 32 participants, has achieved an accuracy of 94.75% for the valence and 95.75% on the arousal detection, which is quite competitive with other methods of emotion recognition.

Keywords: EEG, BCI, DEAP, CNN, FFT, DCT, DWT.

Dedicated to our loved ones for all their inspiration and support.

Acknowledgement

First and foremost, we take this opportunity to show our deepest gratitude to our supervisor, Dr. Mohammad Zavid Parvez, whose contribution in encouraging suggestions helped us to coordinate our thesis especially in compiling modern ideas and advancing to the next level. We feel privileged and honoured to have had him as our supervisor, whose relentless motivation and guidance pushed us to learn more and to absorb more knowledge throughout the whole process. Furthermore, we would like to thank the Department of Computer Science and Engineering for providing an amazing working environment that stimulates learning and also for providing resources that aided our research. We would also like to thank all the academic and admin staff for guiding us through the academic procedures within the University. Lastly we thank our Creator, for blessing us, and immense thanks to all the friends and family for their roles and moral support which enabled us to complete our Thesis on due time.

Table of Contents

List of Figures

List of Tables

Nomenclature

The next list describes several symbols & abbreviation that will be later used within the body of the document

AR	Auto Regressive
BCI	Brain Computer Interface
CNN	Convolutional Neural Network
DCT	Discrete Cosine Transformation
DEAP	Dataset for Emotion Analysis using Physiological signals
DNN	Deep Neural Network
DWT	Discrete Wavelet Transformation
EEG	Electroencephalography
EMD	Empirical Mode Decomposition
FFT	Fast Fourier Transform
FN	False Negative
FP	False Positive
HA	High Arousal
HV	High Valence
LA	Low Arousal

LV	Low Valence
ROC	Receiver Operating Characteristic
SAM	Self Assessment Manikin
SVM	Support Vector Machine
TN	True Negative
TP	True Positive

Chapter 1

Introduction

In this fast advancing modern world, technologies are advancing at an immense rate, and all of this to make the world a better place. The application of emotion recognition in technology and in our daily life can help us detect emotions of autistic or mentally challenged people who cannot express themselves. Advertisers and content creators can integrate emotion recognition in their business to sell their products more effectively as people rely heavily on emotions while making purchase decisions. Emotion recognition will allow to determine fatigue in the case of driving and alert in advance. When a person is scared while withdrawing money, the ATM will not dispense money. In case of video content creators, emotion recognition will help them measure the perceived effectiveness of their short and long form video content. In the case of E-learning, studying the emotions will aid in adjusting to the learning techniques and presenting it according to the style of learner. Emotion recognition for enhancing the above mentioned areas and also future possibilities, is what motivated us to pursue our research. Emotion recognition based on text, speech, facial expression and posture are possible, but by these methods, it is difficult to guarantee the accuracy and authenticity of emotion. According to studies, physiological signals tend to reflect people's real emotions much more accurately than facial expressions, postures or voice. Physiological signals such as Electrooculography (EOG), Electrocardiogram (ECG) and Electromyography (EMG) are indirect reactions caused by emotions, although they are far better than facial expressions, postures or voice. However, these signals lack reasonable evaluation criteria and low emotional recognition accuracy compared to EEG [1] since the generation of emotion is closely related to the cerebral cortex of the brain. Because of its strong objectivity and high accuracy of classification, we have chosen EEG signals to apply our methods for classifying and recognizing emotions.

1.1 Thesis Overview

Emotions are responses to external or internal events, they are discrete although case-wise consistency can be observed. There are two ways of classifying emotions [2]–[5]. The first one is the dimensional approach where emotions are mapped into two dimensions such as valence and arousal. Dimensional representation has been mostly used for emotion recognition. We have opted to use the model proposed by Russel, which uses only valence and arousal. There are other methods of emotion recognition based on facial expression, speech, posture etc [6]–[8]. However detailed emotional inputs from hand gesture, voice can often lead to difficulties while processing, ultimately causing inconsistency in Emotion recognition [9]–[12]. Physiological signals such as EEG, EOG, ECG and EMG are more accurate in reflecting people's emotions. Among these EEG signals have the highest emotional recognition accuracy [6], [13]. Generation of emotions cause signals to generated in the brain, more commonly known as brainwaves. These signals are recorded the using an EEG machine (Please see Figure 1.1).

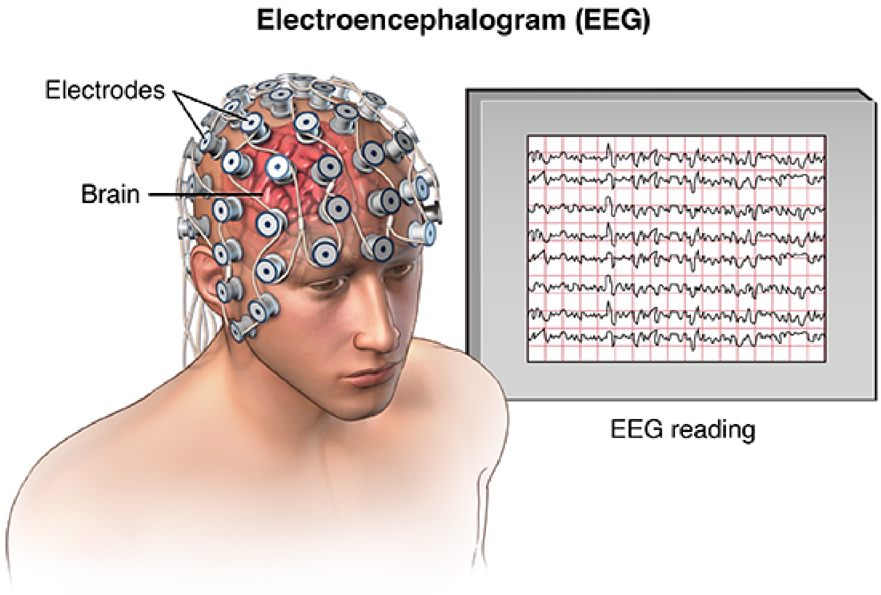


Figure 1.1: Acquiring EEG Signals [14].

The human brain has complex interconnections and billions of neurons. Brain activity is monitored by using the EEG machine [15], [16]. Readings are taken using electrodes placed on the human skull. This makes the EEG signals non linear, stationary and random in nature, which makes it difficult to recognize emotion [17]. In this paper we have decided to use Discrete Cosine Transformation (DCT), Discrete Wavelet Transformation (DWT) and Fast Fourier Transform (FFT) [18], [19] to extract features from the EEG signals. The proposed method is evaluated on the DEAP dataset [20]–[23]. The extracted features are then fed into a CNN to classify the emotion into valence and arousal dimensions, as proposed by James Russell. They can be recognized through his Circumplex Model of Emotion, as illustrated in Figure 1.2.

We used DEAP dataset [24]–[26]. We extracted features using DCT, DWT and FFT and

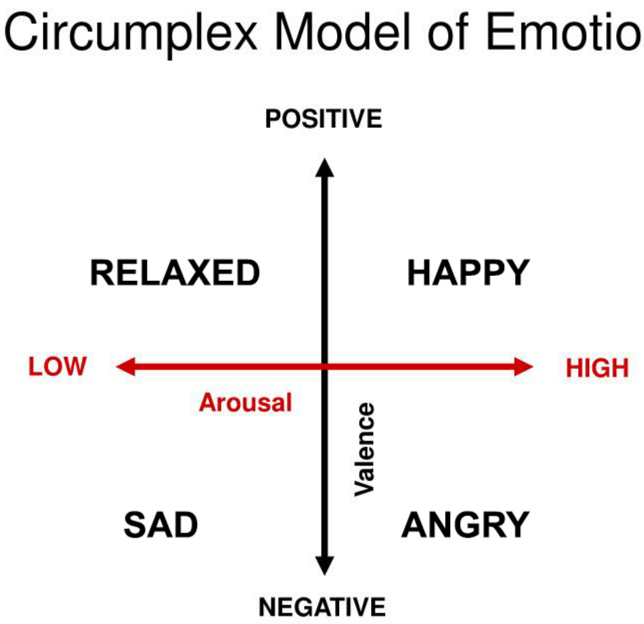


Figure 1.2: Circumplex Model of Emotions

fed them into CNN to classify emotions [27]–[30].

1.2 Thesis Contribution

Emotion Recognition is a revolutionary idea in today's modern world. By using EEG Signals instead of other approaches seen in existing methods, we have contributed to this research by acquiring a more desirable and accurate outcome using valence and arousal, which competes well with other methods. Moreover, our deep learning approach give a more meaningful classification.

1.3 Thesis Orientation

The consequent chapters of the paper have been covered in the following order. Chapter 2 discusses similar work in the field and existing methodologies and their limitations. Chapter 3 gives an intensive analysis of the topics preceding data and information associated with our work. Chapter 4 presents the proposed model in detail. Chapter 5 give the test results and the interrelated discussions. At last, Chapter 6 closes and outlines the the thesis along with future plans.

Chapter 2

Literature Review

BCI of emotion detection can be very beneficial for us and future generations in various fields, giving rise to interest in this field of research. Using DEAP dataset gives us data of EEG signals through forty channels from thirty two different participants watching forty different videos. Emotions can be detected in other ways which most often can be misleading, however emotion detection using EEG signals prove to be more accurate than other methods since it is based on effects on brainwaves due to emotion.

Previous work in detecting human emotion used different approaches, for instance, Feng et al. in [31] conducted a research on emotion detection from speech. Their study uses a corpus containing emotional speech. It includes 721 short utterances, that express four emotions anger, happiness, sadness and neutral state. These utterances were taken manually from movies and plays SVM is applied to them to recognize emotions within the four categories.

Emotions are essential to the way a person behaves or communicates with each other as well as computers [32]. Individuals are fit for understanding an emotional condition of other beings and carry on in like manner so as to improve the state of the circumstance. They can do this since feelings are perceived through words, by means of voice pitch, facial expression and in particular non-verbal communication such as body language, which a machine fails to accomplish. Busso et al. in [33] used database recorded from an actress, by using marks on her face and capturing detailed facial motions, along with speech recordings. They used this database to classify four emotions sadness, anger, happiness and neutral. Moreover, in the field full of affective computing, various methodologies have been connected to understanding emotions dependent on signals received. However, even though most of our emotions come internally, some research are based on body movements like gesture [34], facial expression, and voice intonation [13]. Examining a case where a person is smiling, does not necessarily mean that person is happy.

Human beings convey emotional information, through speech both intentionally and unintentionally. Morrison et al in [35] aimed to improve perception of vocal emotion in two steps. First they compared two emotional data sources, natural, spontenous emotional speech and acted or portrayed emotional speech. They they applied stack generalisation, and unweighted vote to classify emotions. Additionally in the field of computational linguistics, approaches have been taken to detect emotions based on text. Shivhare et al in [36] conducted a research to detect emotions from text document written by a person. Their model works by using predefined keywords, and matching it with emotion words from the text. Their justification for taking this approach is based on the fact that textual expressions use emotional words directly.

Researchers have been trying to distinguish a pattern that might exist for all human beings. Advanced studies have been carried out to understand the neural correlation of emotions. Although, the outcome shows that there is no fixed criteria to recognize specific type of emotions, there are pattern in emotional activity that point towards very different neural signature of specific type of emotions. Zhang et al. in [37] attempted to investigate feature extraction by mainly focusing on Empirical Mode Decomposition (EMD) and Autoregressive (AR) model. To classify these emotional states, he constructed an EEG based emotion recognition model. Tripathi el. at. [38] explored the DNN and CNN, using Dropout (Advanced Machine-Learning Techniques) for classifying emotions. This study also used DEAP dataset to classify emotion based on high/low valence and high/low arousal. Brain activity has been used for recognition and control for quite a few years. Extraction of brain information from neuro-physiological signals about the psychological states has increased interest in human-computer interaction. Making human interactions with computers more natural directly relates to the human emotional states. Among the different methods of recognizing emotions, EEG signal based algorithms are considered to be more efficient due to its high accuracy and stabilization [9], [10].

Chapter 3

Background Analysis

In this chapter, we take an intensive analysis of the background data associated with our work. The analysis of the biological and electrochemical connection of the brain is the root of our research. In Section 3.1, we study the biological features, funtions and structure of the brain as an organ. We learn deeper about its lobes and their functions. Section 3.2 gives us a detailed idea about brainwaves and their types, like Alpha, Beta, Gamma etc. Section 3.3 explains the different algorithms that can be used in relation to the EEG Analysis of the brain, such as ANN, DNN, CNN, SVM etc.

3.1 Basic structure of the Human Brain

The human brain is one of the most complex organs of our body, which is the key part of our nervous system. Brain consists of billions of neurons that communicate in trillions of connections to control most of the activities of our body. It is about 2 percent of the total body weight. It is situated inside the head near all the other neural organs

[39]. The human cerebrum works as an operational centre of the human body for all its roles. It potentially [40] exhibits the importance of conscious and unconscious personality. By creating a consistent flow of sensory signals, the brain makes the body mindful of all inward and outside environments. It assembles the messages in a way that has meaning for us, and can store that information in our memory. A couple of things done by the brain are : Memory, feelings, knowledge, imagination, breathing, gland secretion and internal temperature. The brain gets data of the outside world through the five senses of the human body : sight, taste, smell, touch and hearing. Moreover, it has the capability to control our comprehension of a specific circumstance. The main functions of the brain are motor control, sensory, cognition lateralisation, emotion, language, and regulation. A brain is composed of 3 major segments : Brain Stem, Cerebrum and Cerebellum.

3.1.1 Brain Stem

The Brain Stem is situated below the cerebrum, it interfaces the spinal cord with the brain. By linking the spinal cord to the cerebrum and cerebellum, it acts like a relay centre. It comprises of three parts; medulla oblongata, mid-brain, pons. It is made out of a blend of white and grey matter. It works like a switch to regulate the brain wake and sleep cycles and also controls muscle tone of the body. Moreover, it oversees functions like blood pressure, oxygen levels, puking, wheezing, coughing, swelling reflexes, homeostasis.

3.1.2 Cerebellum

The Cerebellum is found posterior to the Brain Stem and inferior to the Cerebrum. The posterior and anterior lobes are linked in the middle by the vermis. It is hemispherical in form and is wrinkly. It is tasked with the control of motor functions like posture, balancing and aligning muscle actions. It even leads the synchrony and tactfulness of motor activities like walking, writing and speech.

3.1.3 Cerebrum

The Cerebrum makes up most of the brain. It contains both the right and left hemispheres.fibres called corpus callosum connect the two hemispheres. The left hemisphere controls the right half of the body and vice versa. The cerebral cortex is an outer layer of grey matter which covers the core of white matter. These fibers send data from one side to the other. Nevertheless, some functions are independently assigned to each side of the brain. For instance, the left side controls writing, calculations, cognitions, speech while the right side regulates spatial capacity, imagination, creativity, musical aptitude etc. The dominance of left hemisphere in language and hand use can be found in almost 92 percent of people, which explains why few people use left hand to write.

In the Figure 3.1 below, we can see that each half of the cerebellum could further be ordered into four particular areas called lobes: frontal, parietal, temporal, and occipital lobe. Each lobe is obligated to perform a specific kind of function.

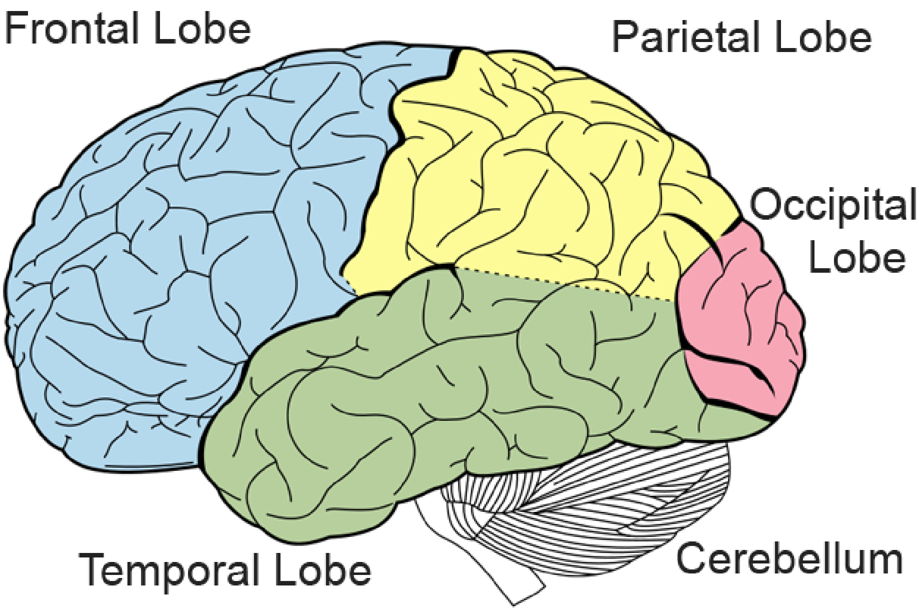


Figure 3.1: Structure Of The Brain [41].

3.1.4 Frontal Lobe

The frontal lobe as the name suggests, is located in the front part of the brain. It is involved in motor abilities, self awareness, writing and speech, intelligence, judgment, planning, problem solving, and emotions.

3.1.5 Parietal Lobe

The parietal lobe is situated just behind the frontal lobe. It is named so being close to the parietal bone. It has somatosensory cortex and deals with processing sensory information to determine touch, sense of pain and temperature.

3.1.6 Occipital Lobe

The occipital lobe is situated at the back of the brain. It has a visual cortex that, receives visual information from the eye. this information is processed in this lobe to make decisions of color, perception and vision.

3.1.7 Temporal Lobe

The temporal lobe is located on the side of the brain, and is involved with learning, memory, emotion and language. it contains auditory cortex which processes auditory information.

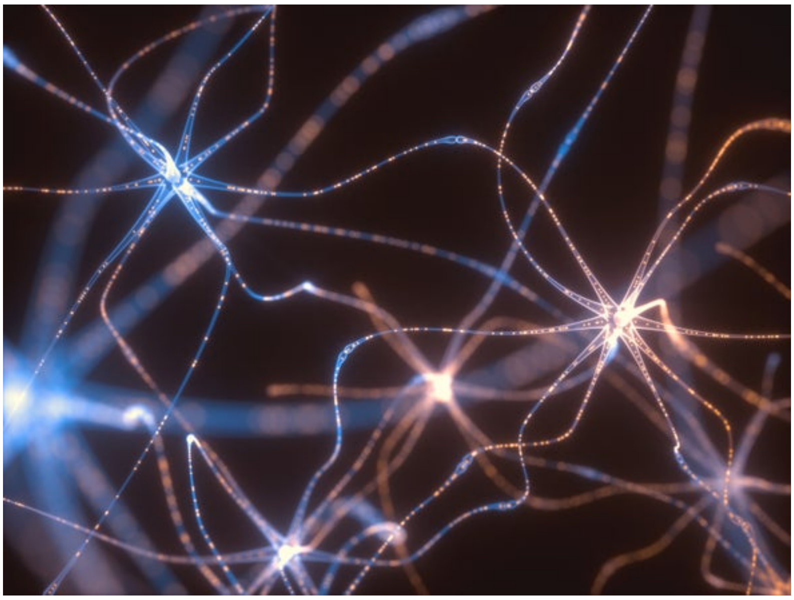


Figure 3.2: Traveling of Brain Waves In Neurons [42].

In the Figure 3.2 above, travelling brainwaves between neurons can be seen. There are billions of interconnected neurons in the brain. these transmit information through electro chemical pulses. These pathways of these pulses are made of nerve cells (neurons), brain cells and glial cells.

Neurons consist of cell body, dendrites and axons. Dendrites transfer message from neurons. Axons are found at the end of the cell, neurotransmission occurs between axon and a dendrite. The glial cells on the other hand facilitate neurons by ensuring synapses are transmitted properly, are are also responsible for repairing nerves. Therefore, messages of motor actions or other responses to stimuli travel through the nerves to the destination muscle, gland or cell. In case of emotion, for example the tearing of eyes due to experiencing a certain emotion, the signal comes all the way from the brain to dictate a certain response.

3.2 Brainwaves

Electrical pulses in the central nervous system, that have repetitive cycles or patterns are what we call "Brainwaves" [43]. These neural oscillations responsible for neural transmission are measured in Hertz (cycles per second). [8] Pulses generate from either a neuron or due to interaction between two neurons. Information is passed onto the target muscle or destination through the nervous system in order to perform motor action, send sensory information or just virtual information. Therefore brainwaves differ based on what information or instruction they are communicating [44]. High frequency brainwaves can be observed when a human is ecstatic, on the other hand low frequency brainwaves can be observed when humans feel lazy or bored. The different brainwaves of varying frequency range can be observed through an EEG. They can be classified according to their frequency range into five types [45], please see Figure 3.3 below.

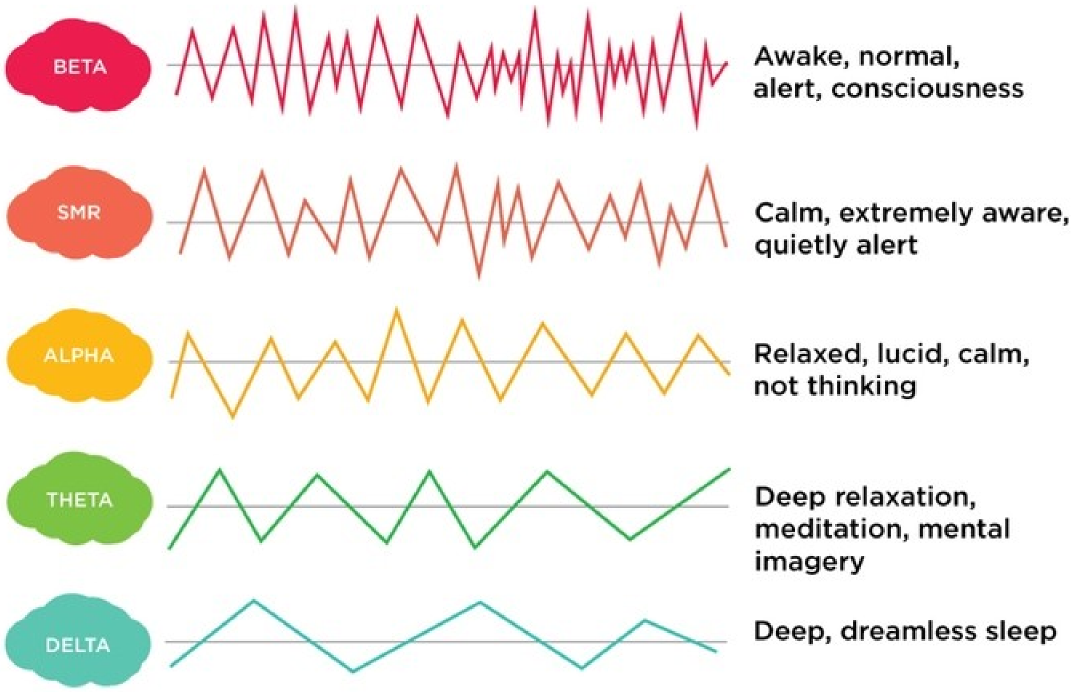


Figure 3.3: Types Of Brain Waves [46].

3.2.1 Alpha

Alpha waves range between frequencies 8 Hz and 13 Hz. This low frequency range is due to the fact that they are generated when the brain is calm and in a relaxed state. It is also generated in thoughtful or meditative state of mind.

3.2.2 Beta

Beta waves range between frequencies 13 Hz to 38 Hz. Since it is mostly generated in the frontal lobe, this means it is produced during cognitive tasks, prioblem solving, planning, self awareness. When humans are actively participating or doing tasks that require attention, brainwaves in this frequency range can be observed.

3.2.3 Delta

Delta waves range between 0.5 Hz and 3 Hz. It has a frequency range lower than that of alpha waves. Hence it is only generated while experiencing deep sleep, and there is no awareness of surroundings whatsoever. During this phase human body recovers and regenerates.

3.2.4 Gamma

Gamma waves have frequencies higher than beta waves, they range from 38 Hz to 42 Hz. Being above neuron firing frequency , it is often related to broader consciousness. Initially it was considered as noise and discarded, till later on it was found out that it is active in states of 'high virtue'.

3.2.5 Theta

Theta waves range between 3 Hz to 8 Hz. It lies between alpha and delta waves and can be observed during meditation or while dreaming. while in this range humans tap into their subconscious mind.

3.3 EEG Analysis with Various Algorithms

A German psychiatrist named Hans Berger [13] invented EEG [47] and proposed the method of recording brainwaves by placing electrodes on the human skull in 1929. Although his method was met with criticism at first it was later recognised later [9]. Eversince more and more researchers engaged in this field to study interaction between computer and humans by observing emotional reactions [48], [49].

EEG device is called noninvasive due to the fact that it does not effect or influence the brain in any way. EEG requires electrodes to be placed on the human skull, which picks up brain waves resulting from electro chemical pulses within neurons in the brain. Electro chemical pulses are received by the EEG machine and passed on to an amplifier, in order to make them view-able on a paper or a display. Simple mechanism and being noninvasive, EEG has emerged to become the best method of detecting brainwaves [10].

Information from the electrodes of the EEG cap [11] can be transferred in different methods such as Bio- Semi [12], B-Alert and Bio-Radio 150 [11]. while recording brainwaves the EEG machine picks up various noises due to eye blinking, muscle movement and also instrument noise [50]. To avoid inconsistencies the noise is removed by using band pass filters, Independent Component Analysis (ICA) can be used as accounted in [51] and [52]. noise can be reduced furher by making sure that there is proper conductivity at the contact points of the scalp [11].

The most common step of applying machine learning algorithms is feature extraction. This is done so to reduce huge dataset and to develop different combinations of variables in order to achieve good accuracy.

3.3.1 Artificial Neural Networks (ANN)

ANN [53] mimics an animal brain. It has interconnected nodes much like neurons, and there is intercommunication between these nodes. In case of image recognition, it can identify certain objects, if examples are provided. By learning from the examples and then finding similar objects in the samples.

3.3.2 Support Vector Machine (SVM)

SVM is a classification algorithm. It separates points of two classes using a line that is farthest from both the points. This line is called a hyper-plane [53] and draws the decision boundary between the two points. the hyper-plane ensure highest possible distance between the points to ensure high classification accuracy.

3.3.3 Convolutional Neural Network (CNN)

A CNN is an neural network most commonly used for image classifiaction, and can also be used for other types of data analysis and classification. This special type of neural network can detect patterns in data. This does so with its hidden layers called the convolutional layers. More specifically the filters in the convolutional layers are used to detect patterns that make the CNN so useful [1], [38], [54]–[56]. See Figure 3.4 below.

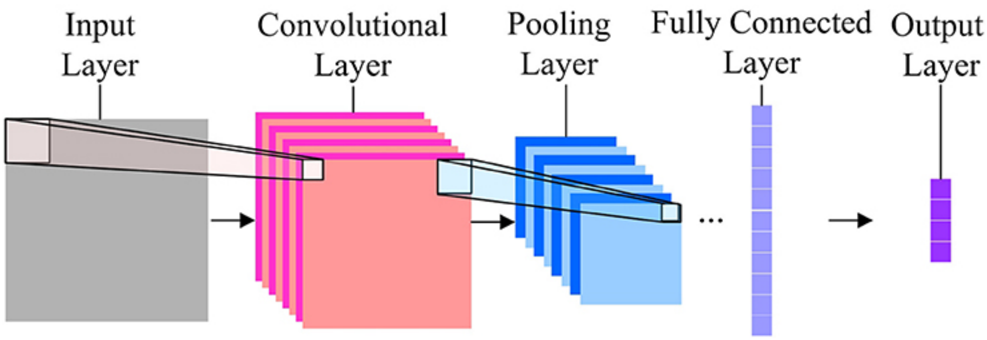


Figure 3.4: Feature Extraction using Convolutional Neural Network [57].

3.3.4 Deep Neural Network (DNN)

A deep neural network is a type of ANN. It has input layers and output layers [58], and several layers in between. These layers have nodes which hold a value and each each of these nodes are connected with nodes of the next layer forming what we call the neural network.The connections also have weighted values. Input data when fed in to this network, features are identified at output at the end.

Chapter 4

Proposed Model

The DEAP Dataset [40], [59] was used in our experiment for detection of emotions. This dataset is based on EEG signals and other peripheral physiological signals. In our research, only EEG signals were used for emotion recognition. At first, preprocessing was used to reduce deviations, noises and inconsistent signals. Preprocessed signals were further refined and then features were extracted based on time-frequency distribution using DWT, DCT and FFT. Finally, the extracted features are fed into a deep learning algorithm called CNN [60]–[62]). The proposed workflow is presented in Figure 4.1.

4.1 Experimental Setup

The experimental setup for the DEAP dataset [63]–[65] was performed to obtain EEG signals and peripheral physiological signals using video records of 40 stimuli tests. 32 participants were shown musical video records. Among them, 50% were male and 50% were female. They were aged between 19 and 37 with an average age of 26.9 years.

About 32 AgCl electrodes were placed on each participant's skull to take their EEG signals at a sample frequency of 512 Hz. Each participant watched a video and their EEG physiological signals were recorded from the following 32 positions (according to the international 10-20 positioning system, as Figure 4.2): and the electrodes are Fp1, AF3, F3,F7, FC5,FC1, C3, T7, CP5, CP1, P3, P7, PO3, O1,Oz, Pz, Fp2, AF4, Fz, F4, F8, FC6, FC2, Cz, C4, T8, CP6, CP2, P4, P8, PO4, and O2.

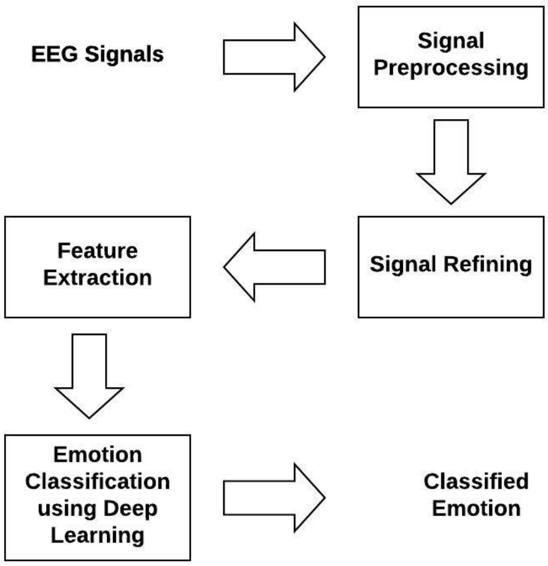


Figure 4.1: Workflow of the Proposed Model.

The videos were all the same for each participant. However, the order of viewing was random for each participant. Before viewing, participants were asked to relax watching a fixation cross on the screen and two minute long EEG signal was collected for each of them. Musical videos were shown to induce emotion in the participants. The participants rated the videos in terms of arousal, valence, like/dislike, dominance and familiarity. In our study, we only considered the dimensions of valence and arousal. See Table 4.1 above.

Table 4.1: The Arrangement of EEG Signal Channel [55].

Channel Number	**Electrode Name**	**Channel Number**	**Electrode Name**
1	Fp1	17	O2
2	AF3	18	PO4
3	F7	19	P4
4	F3	20	P8
5	FC1	21	CP6
6	FC5	22	CP2
7	T7	23	C4
8	C3	24	T8
9	CP1	25	FC6
10	CP5	26	FC2
11	P7	27	F4
12	P3	28	F8
13	Pz	29	AF4
14	PO3	30	Fp2
15	O1	31	Fz
16	Oz	32	Cz

4.2 Data Description

The DEAP is a multimodal dataset based on electroencephalogram and peripheral physiological signals created by researchers from Queen Mary University of London [59]. [66], [67] The EEG and peripheral physiological signals of 32 participants were recorded. Among them 16 male and 16 female subjects each watched 40 music videos which were selected in terms of different levels of arousal, valance, like/dislike, dominance and familiarity. The subjects rated each video of 1 minute long using the Self Assessment Manikin (SAM) in ranges of 1-9 for arousal, valance, like/dislike and from 1-5 for familiarity [8], [45].

In this paper, we used preprocessed EEG signals to detect emotions based on arousal and valence as two independent emotional model [68]. Each participant recorded their emotions into two arrays as shown in Table 4.2. The array contains data of 40 one minute long videos of the 32 participants and labels of valance, arousal, dominance

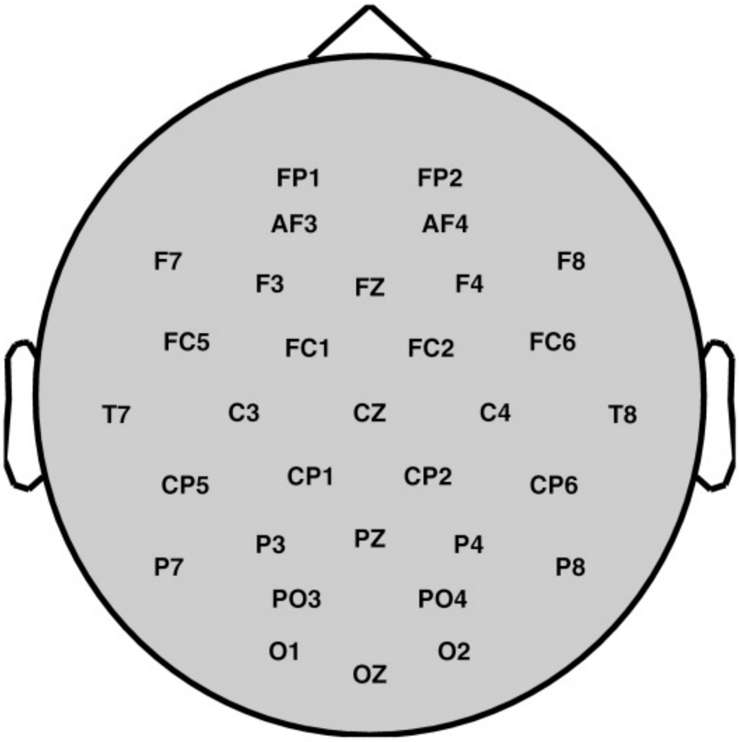


Figure 4.2: Placement of electrodes on the Human skull [21].

and liking. All the 40 videos have 40 different channels containing 8064 samples of EEG signals.

Table 4.2: DEAP Dataset Representation for each subject.

Array Name	**Array Shape**	**Array Contents**
Data	40 x 40 x 8064	video/trial x channel x data
Labels	40 x 4	video/trial x label (valence, arousal, dominance, liking)

4.3 Signal Processing

EEG signals was initially processed by down-sampling to 128Hz from 512Hz [59], [69]. Signals went through a band-pass filter with a minimum of 4Hz and a maximum of 45Hz range to remove the ocular artifact and unwanted noises.

4.4 Signal Refining

Even though each video were run for 63 seconds, the data was averaged to a common reference to achieve higher precision. A 3 seconds of pre-trial baseline from the first 63 seconds was removed to get a 60 seconds refined trial data. This way 7680 samples were achieved from the initial 8064 samples for each of the 40 channels.

4.5 Feature Extraction

In our research, there are 32 participants each of them watched 40 videos and EEG signals were recorded or 40 different channels which produced 8064 samples [70]. As mentioned in Section 4.4, we removed 3 seconds of pre-trial baseline and transformed the data to 32(participants) x 40(videos) x 40(channels) x 7680(samples). There are 40 unique videos and participants might have different reactions to the same video. For this reason, we have 40 videos and 40 channels 40x40. Each of the 40 channel has 7680 samples and for each particular sample, average of 32 participants readings were calculated to maintain a consistency of the reactions for the 40 videos. Therefore, the dataset is restructured to 40(videos) x 40(channels) x 7680(averaged samples). Then the restructured dataset will be used for feature extraction using different time-frequency distributions.

The optimal batch size for valance and arousal classification is 3-12 seconds [Investigation of window size in classification of EEG-emotion signal with wavelet entropy]. Hence, these 7680 samples are of 60 seconds as of 128Hz which further been divided into 12 batches of 5 seconds each containing 640 readings. The acquired readings are converted from time domain to frequency domain using three different feature extraction techniques such as DCT, DWT and FFT, which are described below.

4.5.1 Discrete Cosine Transformation(DCT)

Time series signals can be converted to base frequencies using DCT [71]–[73]. The DCT coefficients are arranged in descending order from high frequency which represent detail signals to low frequency which represents coarse signals. DCT [74] can be used to compress signals based on their frequencies. We took the first 60% of the DCT signals which indicates high frequency as it's more distinguishable.

$$X_k = \sum_{n=0}^{N-1} x_n \, cos[\frac{\pi}{N}(n+\frac{1}{2})k] \qquad \text{.........} [1]$$

where x_n is a set of N data values, the output X_k is a set of N Discrete Cosine Transform coefficients and N is the list of real number.

4.5.2 Discrete Wavelet Transform (DWT)

EEG signals are decomposed into a set of basis functions called wavelets by using DWT [75]–[78]. DWT can be specified by low pass filters and high pass filters. The cutoff frequency is set to be one fourth of the sampling frequency for these filters and the bandwidth of the output set to be half of the original signal. The down-sampled signals from HP and LP, which do not have any loss of information, refer to the first level approximation and details of the original signal.

$$\phi(x) = \sum_{k=-\infty}^{\infty} a_k \phi(S_x - k) \qquad \text{.........} [2]$$

where S is a scaling factor, a_k the finite set of coefficients that define the scaling function.

4.5.3 Fast Fourier Transform (FFT)

Fourier analysis can be used to convert a signal from the original domain to frequency domain. FFT is used to compute DFT (Discrete Fourier Transform) by decomposing a sequence into different frequencies[79]–[81].

$$X_k = \sum_{n=0}^{N-1} x_n e^{-i2\pi kn/N} \qquad \text{.........} [3]$$

where $e^{-i2\pi/N}$ is a primitive Nth root of 1 and there are N outputs X_k ,and each output requires a sum of N terms.

These three features extraction techniques are applied sequentially on each 640 readings and energy is computed on the higher frequency of DCT, detail coefficients of DWT and coefficients values of FFT. Thus, we reduce the dimension of 40(channels) x 7680(readings) to 40(channels) x 36(readings).

$$Energy = \sum_{i=1}^{n} |W_i X_i(t)|^2 \qquad \text{.........} [4]$$

where X(t) is the higher frequency of DCT coefficients and n is the length of the signals.

Since our data can have multiple dimensions, to limit the range of inconsistency, we used feature standardization to normalize the data. Standardization distributes the features and scales them in way such that it is centered around zero with a standard deviation of one.

$$X' = \frac{x-u}{a} \qquad \text{.........} [5]$$

where x is the original feature vector, $\bar{x}$ = average(x) is the mean of that feature vector, and σ is its standard deviation.

Classification of emotions in our approach is based on valence and arousal labels and the emotional state values range from 1-9. If the emotional state value is less than 5 it is considered as low valence(LV) or low arousal(LA) and if the emotional state value is greater than 5 is considered as high valence (HV) or high arousal(HA). The labels are processed with one hot encoding for classifying two classes of high and low valence/arousal.

4.6 Emotion Classification

Emotions are separately classified for valance and arousal in our approach [Koelstra]. For classifying emotions we have used CNN. It is a renowned model for classifying images with very high outcome. The data 40(videos) x 40(channels) x 36(features) is converted to 1600 channels with 36 features each and labels for each video is assigned to their respective channels. This 1600(channels) x 36(features) data are fed into 1D CNN to train our model. In our model, we have used three 1D CNN and max pooling layers, first 1D CNN layer has 64 filters with kernel size of 3. Some functions are not able to activate neurons of the later layers consistently. To avoid making the model defective, the first layer uses Tan Hyperbolic activation function. After the first layer, the second and third layer of CNN have 64 and 128 filters respectively, the kernel size is 3 and max pooling is done in both of the layers with a pool size of 2. Like the first layer both these layers Tan Hyperbolic activation function. We used adadelta optimizer of learning rate 0.01, a batch size of 8 and 20 epochs. This output is fed to the final neural layer, which uses Softmax as activation function to detect classes of valence or arousal.

Chapter 5

Results And Discussions

Different strategies were implemented to reach the final output. Reaching the output required rigorous research, which included the goal of extracting features in the best possible approach. In this research, we present an approach to classify human emotion using DEAP dataset, a benchmark for emotion research, to transfer EEG signals into time domain frequency analysis, and CNN to classify emotion. The following equation was used to determine the accuracy.

$$Sensitivity = \frac{TP}{TP + FN} \qquad \text{.........} [6]$$

$$Specificity = \frac{TN}{TN + FP} \qquad \text{.........} [7]$$

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP} \qquad \text{.........} [8]$$

where TP is the number of true positive, TN is the number of true negative, FP is the number of false positive, FN is the number of false negative.

It can be observed in the Table 4.2 that the data set is divided into two arrays named data and labels. The data array has a dimension of 40x40x8064 containing video, channel data. The labels array has a dimension of 40x4 containing trials and labels (valence, arousal, dominance, liking).

Firstly, DEAP dataset signal was refined by omitting 3 seconds of signals to make them 7680 from 8064, then divided into 12 batch of 5 seconds each at 128 Hz. Initially we chose different set of features extraction techniques such as statistical and PSD. However, we achieved accuracy of 65.25% and 67.75% only. As it did not not yield good accuracy, the approach had to be changed. Furthermore, we applied three feature extraction techniques: DCT, DWT and FFT, for 12 sample batches and combined them together. This brings the data to dimension of 40x36; 40 being the number of channels and 36 the features.

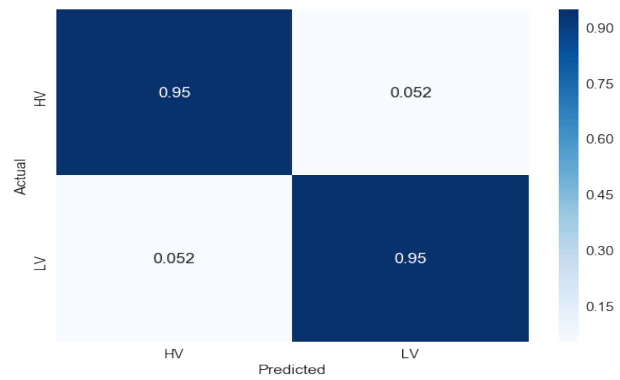
After the consideration of new features DCT, DWT and FFT, there was a significant improvement in the result. However, it was not close to the desired accuracy. Hence, further refinement was made to the signal by feature scaling, so that they can be compared to common grounds and ultimately fed to the CNN. Since data deviation was reduced before feeding to the CNN, the accuracy improved significantly. The classification brought improvement to the results, where for the valence, 94.75% accuracy, specificity and sensitivity is achieved. For the arousal, we achieved 95.75% accuracy, specificity and sensitivity.

Comparing accuracy for the valence and arousal for other methods and our approach

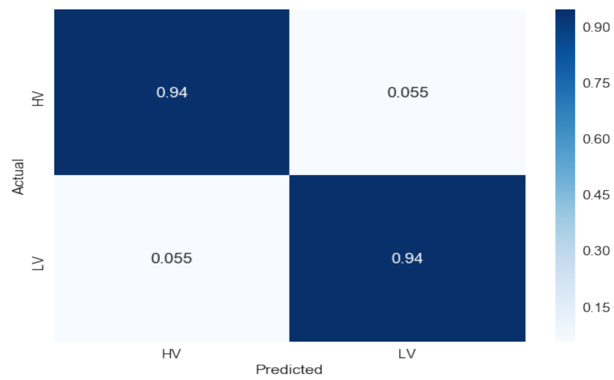
Table 5.1: Accuracy comparison with other methodologies.

Reference	**Method**	**Valence Accuracy**	**Arousal Accuracy**
[82]	MESAE	83.04%	84.18%
[83]	mRMR	73.14%	73.06%
[84]	Decision Tree Algorithm	67.89%	69.09%
[85]	SVM	58.40%	64.20%
[86]	SVM	50.80%	76.51%
[87]	DTCWPT[C]	64.30%	66.20%
[88]	Bayesian Weighted-log Posterior	70.10%	70.90%
[59]	Naive Bayes Classifier	62.70%	62.00%
[89]	SVM	92.36 %	92.36%
	Proposed Method	94.75%	95.75%

is stated in Table 5.1. However, we see that our approach with combination of three feature extraction techniques shows better result than most other available approaches. With the accuracy of 94.75% for valence and 95.75% for arousal, our approach is the most efficient for classifying emotions using the DEAP dataset.



(a) Valence.



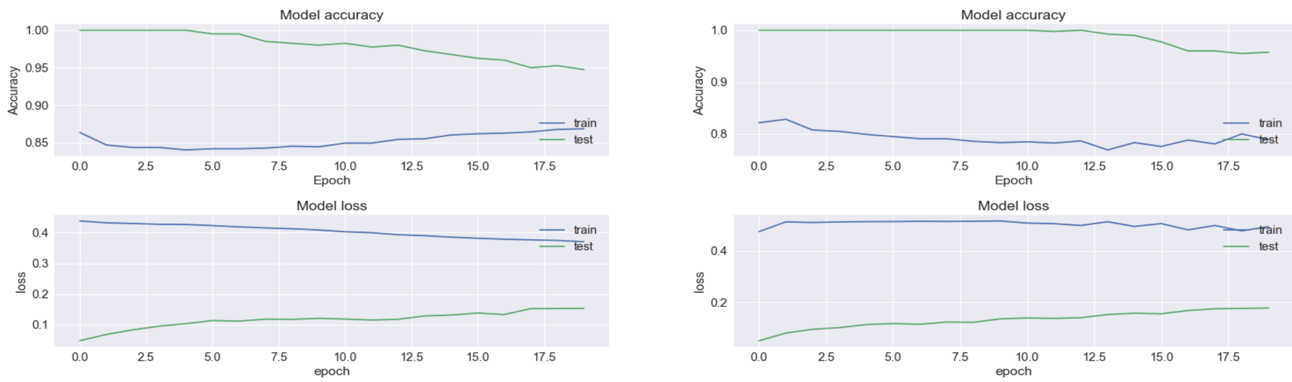
(b) Arousal.

Figure 5.1: Normalized Confusion Matrix for two classes.

In Figure 5.1, the representations of confusion matrix in normalized form for both valence and arousal is indicated. The emotion recognition was categorized in 2 classes for arousal as high arousal and low arousal and for valence 2 classes as high valence and low valence.

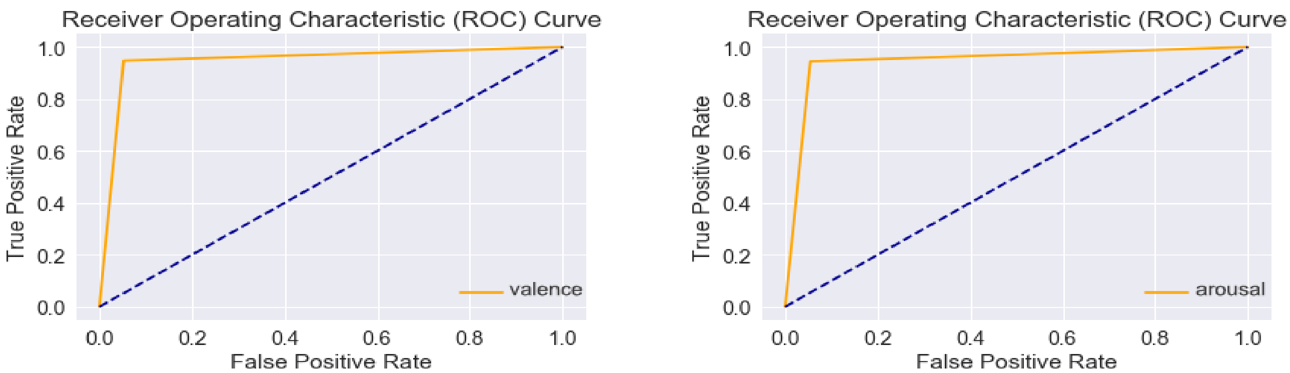
In Figure 5.3, the performance of the classifier was obtained from the true positive rate and false positive rate and plotted for both valence and arousal. The ROC curve shows exponential behaviour with 3 points only as binary cross entropy was used and output was obtained as 0 and 1.

In Figure 5.2, the accuracy and loss result for the CNN model for arousal and valence is shown for 20 epochs.



(a) Valence. (b) Arousal.

Figure 5.2: Accuracy and Loss Result for the CNN Model.



(a) Valence. (b) Arousal.

Figure 5.3: Receiver Operating Characteristic Curve (ROC) for Valence and Arousal.

Chapter 6

Conclusion

In this fast advancing modern world, technology is advancing at an immense rate and so is emotion recognition in the field of research. Developing this research and integrating it in various parts of our life may be very rewarding for us. There are numerous approaches to emotion recognition, like through facial expression, hand gesture and voices. Emotion recognition in these type of approaches have a common drawback. They can be sometimes faked at will. Hence it makes these approaches misleading and not always accurate. This is where Electroencephalogram (EEG) signals are better since it is based on the changes and detection of brainwaves that result from experiencing emotions. This research starts with DEAP dataset. The EEG signals are refined by reducing noise and deviation. Afterwards, batch-wise computation of energy of the features DCT, DWT and FFT is applied. This energy computation of the features are fed to the CNN to acquire emotion classification which is exclusive in our research. We can expect an accuracy of 94.75% for the valence and 95.75% for the arousal from our research findings.

In the future, we would like to integrate our methodology of emotion detection, to help autistic people or mentally challenged people to express their emotions, we would also like to help businesses to sell their products better by understanding the consumers emotion, as they rely on emotions while making purchase decisions.

References

[1] J. Chen, P. Zhang, Z. Mao, Y. Huang, D. Jiang, and Y. Zhang, "Accurate eeg-based emotion recognition on combined features using deep convolutional neural networks", *IEEE Access*, vol. 7, pp. 44 317–44 328, 2019.

[2] W.-L. Zheng, J.-Y. Zhu, Y. Peng, and B.-L. Lu, "Eeg-based emotion classification using deep belief networks", in *2014 IEEE International Conference on Multimedia and Expo (ICME)*, IEEE, 2014, pp. 1–6.

[3] J. Bang, J.-S. Choi, and K. Park, "Noise reduction in brainwaves by using both eeg signals and frontal viewing camera images", *Sensors*, vol. 13, no. 5, pp. 6272–6294, 2013.

[4] R. Khosrowabadi, H. C. Quek, A. Wahab, and K. K. Ang, "Eeg-based emotion recognition using self-organizing map for boundary detection", in *2010 20th International Conference on Pattern Recognition*, IEEE, 2010, pp. 4242–4245.

[5] N. Jatupaiboon, S. Pan-ngum, and P. Israsena, "Emotion classification using minimal eeg channels and frequency bands", in *The 2013 10th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, IEEE, 2013, pp. 21–24.

[6] B. Perseh and A. R. Sharafat, “An efficient p300-based bci using wavelet features and ibpso-based channel selection”, *Journal of medical signals and sensors*, vol. 2, no. 3, p. 128, 2012.

[7] Y.-P. Lin, Y.-H. Yang, and T.-P. Jung, “Fusion of electroencephalographic dynamics and musical contents for estimating emotional responses in music listening”, *Frontiers in neuroscience*, vol. 8, p. 94, 2014.

[8] T. Matlovic, P. Gaspar, R. Moro, J. Simko, and M. Bielikova, “Emotions detection using facial expressions recognition and eeg”, in *2016 11th international workshop on semantic and social media adaptation and personalization (SMAP)*, IEEE, 2016, pp. 18–23.

[9] T. Song, W. Zheng, P. Song, and Z. Cui, “Eeg emotion recognition using dynamical graph convolutional neural networks”, *IEEE Transactions on Affective Computing*, 2018.

[10] M. Singh, M. M. Singh, and N. Singhal, “Emotion recognition along valence axis using naive bayes classifier”, *International Journal of Information Technology & Knowledge Management*, vol. 7, no. 1, pp. 51–55, 2013.

[11] S. Kumar and F. Sahin, “A framework for a real time intelligent and interactive brain computer interface”, *Computers & Electrical Engineering*, vol. 43, pp. 193–214, 2015.

[12] A. K. T. Pandey *et al.*, *Empirical evaluation of machine learning algorithms based on emg, ecg and gsr data to classify emotional states*, 2013.

[13] J. G. Webster and J. W. Clark, *Medical instrumentation: application and design*. John Wiley & Sons, 1995, vol. 200.

[14] *EEG Measures of Cognition", Hvmn.com, 2019*, [Accessed: 04- Aug- 2019], 2016. [Online]. Available: https://hvmn.com/biohacker-guide/cognition/eeg-measures-of-cognition.

[15] M. Li and B.-L. Lu, "Emotion classification based on gamma-band eeg", in *2009 Annual International Conference of the IEEE Engineering in medicine and biology society*, IEEE, 2009, pp. 1223–1226.

[16] R. Ramirez and Z. Vamvakousis, "Detecting emotion from eeg signals using the emotive epoc device", in *International Conference on Brain Informatics*, Springer, 2012, pp. 175–184.

[17] K. S. Kassam, A. R. Markey, V. L. Cherkassky, G. Loewenstein, and M. A. Just, "Identifying emotions on the basis of neural activation", *PloS one*, vol. 8, no. 6, e66032, 2013.

[18] L. M. Bruce, C. H. Koger, and J. Li, "Dimensionality reduction of hyperspectral data using discrete wavelet transform feature extraction", *IEEE Transactions on geoscience and remote sensing*, vol. 40, no. 10, pp. 2331–2338, 2002.

[19] E. Feig and S. Winograd, "Fast algorithms for the discrete cosine transform", *IEEE Transactions on Signal processing*, vol. 40, no. 9, pp. 2174–2193, 1992.

[20] S. D'Mello and A. Graesser, "Automatic detection of learner's affect from gross body language", *Applied Artificial Intelligence*, vol. 23, no. 2, pp. 123–150, 2009.

[21] G. Placidi, P. Di Giamberardino, A. Petracca, M. Spezialetti, and D. Iacoviello, "Classification of emotional signals from the deap dataset", in *International Congress on Neurotechnology, Electronics and Informatics*, SCITEPRESS, vol. 2, 2016, pp. 15–21.

[22] M. L. R. Menezes, A. Samara, L. Galway, A. Sant'Anna, A. Verikas, F. Alonso-Fernandez, H. Wang, and R. Bond, "Towards emotion recognition for virtual environments: An evaluation of eeg features on benchmark dataset", *Personal and Ubiquitous Computing*, vol. 21, no. 6, pp. 1003–1013, 2017.

[23] M. Sreeshakthy and J. Preethi, "Classification of human emotion from deap eeg signal using hybrid improved neural networks with cuckoo search", *BRAIN. Broad Research in Artificial Intelligence and Neuroscience*, vol. 6, no. 3-4, pp. 60–73, 2016.

[24] A. T. Sohaib, S. Qureshi, J. Hagelbäck, O. Hilborn, and P. Jerčić, "Evaluating classifiers for emotion recognition using eeg", in *International conference on augmented cognition*, Springer, 2013, pp. 492–501.

[25] I. Wichakam and P. Vateekul, "An evaluation of feature extraction in eeg-based emotion prediction with support vector machines", in *2014 11th international joint conference on computer science and software engineering (JCSSE)*, IEEE, 2014, pp. 106–110.

[26] Z. Lan, O. Sourina, L. Wang, R. Scherer, and G. R. Müller-Putz, "Domain adaptation techniques for eeg-based emotion recognition: A comparative study on two public datasets", *IEEE Transactions on Cognitive and Developmental Systems*, vol. 11, no. 1, pp. 85–94, 2018.

[27] A. Vedaldi and K. Lenc, "Matconvnet: Convolutional neural networks for matlab", in *Proceedings of the 23rd ACM international conference on Multimedia*, ACM, 2015, pp. 689–692.

[28] H. Li, Z. Lin, X. Shen, J. Brandt, and G. Hua, “A convolutional neural network cascade for face detection”, in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 5325–5334.

[29] N. Kalchbrenner, E. Grefenstette, and P. Blunsom, “A convolutional neural network for modelling sentences”, *arXiv preprint arXiv:1404.2188*, 2014.

[30] B. Hu, Z. Lu, H. Li, and Q. Chen, “Convolutional neural network architectures for matching natural language sentences”, in *Advances in neural information processing systems*, 2014, pp. 2042–2050.

[31] F. Yu, E. Chang, Y.-Q. Xu, and H.-Y. Shum, “Emotion detection from speech to enrich multimedia content”, in *Pacific-Rim Conference on Multimedia*, Springer, 2001, pp. 550–557.

[32] F. Jabr, “The evolution of emotion: Charles darwin’s little-known psychology experiment”, *Scientific American*, vol. 24, 2010.

[33] C. Busso, Z. Deng, S. Yildirim, M. Bulut, C. M. Lee, A. Kazemzadeh, S. Lee, U. Neumann, and S. Narayanan, “Analysis of emotion recognition using facial expressions, speech and multimodal information”, in *Proceedings of the 6th international conference on Multimodal interfaces*, ACM, 2004, pp. 205–211.

[34] T. S. Rached and A. Perkusich, “Emotion recognition based on brain-computer interface systems”, *Brain-computer interface systems-Recent progress and future prospects*, pp. 253–270, 2013.

[35] D. Morrison, R. Wang, and L. C. De Silva, “Ensemble methods for spoken emotion recognition in call-centres”, *Speech communication*, vol. 49, no. 2, pp. 98–112, 2007.

[36] S. N. Shivhare and S. Khethawat, “Emotion detection from text”, *arXiv preprint arXiv:1205.4944*, 2012.

[37] Y. Zhang, S. Zhang, and X. Ji, “Eeg-based classification of emotions using empirical mode decomposition and autoregressive model”, *Multimedia Tools and Applications*, vol. 77, no. 20, pp. 26 697–26 710, 2018.

[38] S. Tripathi, S. Acharya, R. D. Sharma, S. Mittal, and S. Bhattacharya, “Using deep and convolutional neural networks for accurate emotion classification on deap dataset.”, in *Twenty-Ninth IAAI Conference*, 2017.

[39] G. F. Wilson and C. A. Russell, “Real-time assessment of mental workload using psychophysiological measures and artificial neural networks”, *Human factors*, vol. 45, no. 4, pp. 635–644, 2003.

[40] N. Jadhav, R. Manthalkar, and Y. Joshi, “Effect of meditation on emotional response: An eeg-based study”, *Biomedical Signal Processing and Control*, vol. 34, pp. 101–113, 2017.

[41] *Structure of the Brain*, [Accessed: 04- Aug- 2019], 2019. [Online]. Available: http://controlmind.info/human-brain/structure-of-the-brain.

[42] S. Makin, *"Traveling" Brain Waves May Be Critical for Cognition*, [Accessed: 04-Aug- 2019], 2018. [Online]. Available: https://www.scientificamerican.com/article/traveling-brain-waves-may-be-critical-for-cognition.

[43] R. Horlings, D. Datcu, and L. J. Rothkrantz, “Emotion recognition using brain activity”, in *Proceedings of the 9th international conference on computer systems and technologies and workshop for PhD students in computing*, ACM, 2008, p. 6.

[44] C. Mühl, B. Allison, A. Nijholt, and G. Chanel, "A survey of affective brain computer interfaces: Principles, state-of-the-art, and challenges", *Brain-Computer Interfaces*, vol. 1, no. 2, pp. 66–84, 2014.

[45] K. Kalimeri and C. Saitis, "Exploring multimodal biosignal features for stress detection during indoor mobility", in *Proceedings of the 18th ACM International Conference on Multimodal Interaction*, ACM, 2016, pp. 53–60.

[46] *Understanding brain waves - Neurofeedback Alliance*, [Accessed: 04- Aug- 2019]. [Online]. Available: http://neurofeedbackalliance.org/understanding-brain-waves.

[47] M. Murugappan, M. Rizon, R. Nagarajan, S. Yaacob, I. Zunaidi, and D. Hazry, "Lifting scheme for human emotion recognition using eeg", in *2008 International symposium on information technology*, IEEE, vol. 2, 2008, pp. 1–7.

[48] K. Schaaff and T. Schultz, "Towards emotion recognition from electroencephalographic signals", in *2009 3rd international conference on affective computing and intelligent interaction and workshops*, IEEE, 2009, pp. 1–6.

[49] S. Giripunje, "Design approach for eeg-based human computer interaction driver monitoring system",

[50] D. O. Bos *et al.*, "Eeg-based emotion recognition", *The Influence of Visual and Auditory Stimuli*, vol. 56, no. 3, pp. 1–17, 2006.

[51] T.-P. Jung, S. Makeig, C. Humphries, T.-W. Lee, M. J. Mckeown, V. Iragui, and T. J. Sejnowski, "Removing electroencephalographic artifacts by blind source separation", *Psychophysiology*, vol. 37, no. 2, pp. 163–178, 2000.

[52] T.-P. Jung, S. Makeig, M. Westerfield, J. Townsend, E. Courchesne, and T. J. Sejnowski, “Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects”, *Clinical Neurophysiology*, vol. 111, no. 10, pp. 1745–1758, 2000.

[53] R. Singla, B. Chambayil, A. Khosla, and J. Santosh, “Comparison of svm and ann for classification of eye events in eeg”, *Journal of Biomedical Science and Engineering*, vol. 4, no. 01, p. 62, 2011.

[54] L. Santamaria-Granados, M. Munoz-Organero, G. Ramirez-Gonzalez, E. Abdulhay, and N. Arunkumar, “Using deep convolutional neural network for emotion detection on a physiological signals dataset (amigos)”, *IEEE Access*, vol. 7, pp. 57–67, 2018.

[55] Z. Wen, R. Xu, and J. Du, “A novel convolutional neural networks for emotion recognition based on eeg signal”, in *2017 International Conference on Security, Pattern Analysis, and Cybernetics (SPAC)*, IEEE, 2017, pp. 672–677.

[56] W. Lin, C. Li, and S. Sun, “Deep convolutional neural network for emotion recognition using eeg and peripheral physiological signal”, in *International Conference on Image and Graphics*, Springer, 2017, pp. 385–394.

[57] M. Peng, C. Wang, T. Chen, and G. Liu, *Nirfacenet: A convolutional neural network for near-infrared face identification*, 2016. DOI: 10.3390/info7040061. [Online]. Available: https://www.mdpi.com/2078-2489/7/4/61.

[58] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, and H. Adeli, “Deep convolutional neural network for the automated detection and diagnosis of seizure using eeg signals”, *Computers in biology and medicine*, vol. 100, pp. 270–278, 2018.

[59] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, “Deap: A database for emotion analysis; using physiological signals”, *IEEE transactions on affective computing*, vol. 3, no. 1, pp. 18–31, 2011.

[60] Y. Kim, “Convolutional neural networks for sentence classification”, *arXiv preprint arXiv:1408.5882*, 2014.

[61] S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back, “Face recognition: A convolutional neural-network approach”, *IEEE transactions on neural networks*, vol. 8, no. 1, pp. 98–113, 1997.

[62] P. Y. Simard, D. Steinkraus, J. C. Platt, *et al.*, “Best practices for convolutional neural networks applied to visual document analysis.”, in *Icdar*, vol. 3, 2003.

[63] E. Acar, F. Hopfgartner, and S. Albayrak, “Understanding affective content of music videos through learned representations”,

[64] H. Candra, M. Yuwono, R. Chai, A. Handojoseno, I. Elamvazuthi, H. T. Nguyen, and S. Su, “Investigation of window size in classification of eeg-emotion signal with wavelet entropy and support vector machine”, in *2015 37th Annual international conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, 2015, pp. 7250–7253.

[65] S. Alhagry, A. A. Fahmy, and R. A. El-Khoribi, “Emotion recognition based on eeg using lstm recurrent neural network”, *Emotion*, vol. 8, no. 10, pp. 355–358, 2017.

[66] P. A. Wandile, N. Bawane, and M. P. Hajare, “Emotion detection from brain and audio signal”, *International Journal of scientific research and management (IJSRM)*, vol. 3, no. 8, pp. 3507–3510, 2015.

[67] P. Ackermann, C. Kohlschein, J. Á. Bitsch, K. Wehrle, and S. Jeschke, "Eeg-based automatic emotion recognition: Feature extraction, selection and classification methods", in *2016 IEEE 18th international conference on e-health networking, applications and services (Healthcom)*, IEEE, 2016, pp. 1–6.

[68] J. A. Russell, "A circumplex model of affect.", *Journal of personality and social psychology*, vol. 39, no. 6, p. 1161, 1980.

[69] X. Jie, R. Cao, and L. Li, "Emotion recognition based on the sample entropy of eeg", *Bio-medical materials and engineering*, vol. 24, no. 1, pp. 1185–1192, 2014.

[70] V. N. Vapnik, "Direct methods in statistical learning theory", in *The nature of statistical learning theory*, Springer, 2000, pp. 225–265.

[71] W.-H. Chen, C. Smith, and S. Fralick, "A fast computational algorithm for the discrete cosine transform", *IEEE Transactions on communications*, vol. 25, no. 9, pp. 1004–1009, 1977.

[72] B. Lee, "A new algorithm to compute the discrete cosine transform", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 32, no. 6, pp. 1243–1245, 1984.

[73] H. Hou, "A fast recursive algorithm for computing the discrete cosine transform", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 35, no. 10, pp. 1455–1461, 1987.

[74] M. Narasimha and A. Peterson, "On the computation of the discrete cosine transform", *IEEE Transactions on Communications*, vol. 26, no. 6, pp. 934–936, 1978.

[75] M. Lang, H. Guo, J. E. Odegard, C. S. Burrus, and R. O. Wells, “Noise reduction using an undecimated discrete wavelet transform”, *IEEE Signal Processing Letters*, vol. 3, no. 1, pp. 10–12, 1996.

[76] C. E. Heil and D. F. Walnut, “Continuous and discrete wavelet transforms”, *SIAM review*, vol. 31, no. 4, pp. 628–666, 1989.

[77] A. Jensen and A. la Cour-Harbo, *Ripples in mathematics: the discrete wavelet transform*. Springer Science & Business Media, 2001.

[78] H. Ocak, “Automatic detection of epileptic seizures in eeg using discrete wavelet transform and approximate entropy”, *Expert Systems with Applications*, vol. 36, no. 2, pp. 2027–2036, 2009.

[79] E. O. Brigham and E. O. Brigham, *The fast Fourier transform and its applications*. prentice Hall Englewood Cliffs, NJ, 1988, vol. 448.

[80] P. Welch, “The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms”, *IEEE Transactions on audio and electroacoustics*, vol. 15, no. 2, pp. 70–73, 1967.

[81] G. Bergland, “A guided tour of the fast fourier transform”, *IEEE spectrum*, vol. 6, no. 7, pp. 41–52, 1969.

[82] Z. Yin, M. Zhao, Y. Wang, J. Yang, and J. Zhang, “Recognition of emotions using multimodal physiological signals and an ensemble deep learning model”, *Computer methods and programs in biomedicine*, vol. 140, pp. 93–110, 2017.

[83] J. Atkinson and D. Campos, “Improving bci-based emotion recognition by combining eeg feature selection and kernel classifiers”, *Expert Systems with Applications*, vol. 47, pp. 35–41, 2016.

[84] J. Chen, B. Hu, P. Moore, X. Zhang, and X. Ma, "Electroencephalogram-based emotion assessment system using ontology and data mining techniques", *Applied Soft Computing*, vol. 30, pp. 663–674, 2015.

[85] X. Li, P. Zhang, D. Song, G. Yu, Y. Hou, and B. Hu, "Eeg based emotion identification using unsupervised deep feature learning", 2015.

[86] Y. Liu and O. Sourina, "Eeg-based valence level recognition for real-time applications", in *2012 International Conference on Cyberworlds*, IEEE, 2012, pp. 53–60.

[87] D. S. Naser and G. Saha, "Recognition of emotions induced by music videos using dt-cwpt", in *2013 Indian Conference on Medical Informatics and Telemedicine (ICMIT)*, IEEE, 2013, pp. 53–57.

[88] H. J. Yoon and S. Y. Chung, "Eeg-based emotion estimation using bayesian weighted-log-posterior function and perceptron convergence algorithm", *Computers in biology and medicine*, vol. 43, no. 12, pp. 2230–2237, 2013.

[89] P. S. F. Hossain, I. M. Shaikat, and F. P. George, "Emotion recognition using brian signals based on time-frequency analysis and supervised learning algorithm", PhD thesis, BRAC University, 2018.