Formation of Channel Correlation Based Significant Virtual Images from EEG Sub-bands Data to Recognize Human Emotion using Convolutional Neural Network

by

Md. Rabiul Islam

Roll No: 1603557

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Electrical and Electronic Engineering (M.Sc. Eng. in EEE)

Khulna University of Engineering & Technology

Khulna 9203, Bangladesh

July 2019

Declaration

This is to certify that the thesis work entitled "***Formation of Channel Correlation Based Significant Virtual Images from EEG Sub-bands Data to Recognize Human Emotion using Convolutional Neural Network***" has been carried out by **Md. Rabiul Islam** in the Department of **Electrical and Electronic Engineering**, Khulna University of Engineering & Technology, Khulna, Bangladesh. The above thesis work or any part of this work has not been submitted anywhere for the award of any degree or diploma.

Signature of the Supervisor

Signature of the Candidate

Certificate of Research

This is to certify that the thesis titled "***Formation of Channel Correlation Based Significant Virtual Images from EEG Sub-bands Data to Recognize Human Emotion using Convolutional Neural Network***" by **Md. Rabiul Islam,** submitted to the Department of Electrical and Electronic Engineering (EEE), Khulna University of Engineering & Technology (KUET), Khulna, Bangladesh, for the award of Master of Science (M.Sc.) in Electrical and Electronic Engineering carried out by his paper under my supervision. I certify that this thesis fulfills part of the requirements for the award of the degree. The thesis has not been submitted for any other degree elsewhere.

Prof. Dr. Mohiuddin Ahmad
Dept. of Electrical and Electronic Engineering,
Khulna University of Engineering & Technology,
Khulna-9203, Bangladesh.

Acknowledgment

All the praises and thanks to Almighty Allah for making me capable to complete my thesis work successfully despite all challenges were faced.

I would like to show the gratitude to my supervisor **Dr. Mohiuddin Ahmad**, Professor of Electrical and Electronic Engineering, KUET; who fully gave much invaluable advice, proper guidance, constant encouragement, constructive suggestion, and all kind co-operation throughout the entire process of this research work. His enriched and up-to-date experience in the field of signal processing helps me to improve my research ideas. I am also especially grateful to my departmental head **Dr. Md. Rafiqul Islam (2)**, the committee members **Dr. Bashudeb Chandra Ghosh**, **Dr. Md. Aminul Haque Akhand,** and **Dr. Md. Monirul Kabir** due to give me indispensable advice to upgrade my work. I am also indebted to the other departmental teachers for giving me continuous inspiration. The continuous support and instantaneous help of my supervisor sir and the other teachers reach me to finalize my research work. The completion of this assignment gives me much pleasure.

In addition, the most recondite goes to my parents for their ineffable support. Also, I want to express my heartfelt thanks to my wife who always encourages me to complete my work timely and give me continuous support in every sphere of my busy life. Moreover, I also grateful to my friends for their mental support and inspiration. I would also like to expand my deepest gratitude to all those who have directly and indirectly guided me in writing this thesis report.

List of Publications

1. **Md. Rabiul Islam** and Mohiuddin Ahmad, "Wavelet Analysis Based Classification of Emotion from EEG Signal," ***2nd*** ***International Conference on Electrical, Computer and Communication Engineering***, Cox's Bazar, February 2019.

2. **Md. Rabiul Islam** and Mohiuddin Ahmad, "Virtual Image from EEG to Recognize Appropriate Emotion using Convolutional Neural Network," ***1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT)***, Dhaka, Bangladesh, May 2019.

Abstract

Emotions are the most fundamental feature for non-verbal communication between human and machine. But the improvement of different human interacted media depends largely on the appropriate recognition of the human feeling that means what he or she wants just right at that moment. To communicate computer with human, to detect present feelings of the patient, to understand customer interest for shopping, to realize interest of physically handicapped people or to detect the lie of a convict; emotion recognition is a fundamental prerequisite. But due to some complexities, the proper recognition of human emotion from Electroencephalogram (EEG) has become too much challenging.

Normally, feature-based emotion recognition requires a strong effort to design the perfect feature or feature set related to the classification of emotion. To curtail the manual human effort of feature extraction, we designed a model with Convolutional Neural Network (CNN). As Electroencephalogram (EEG) is 1D data, to use CNN the 1D EEG data have to be converted into 2D significant image data. That is a challenging task. To meet up this challenge, initially, we calculated Pearson's correlation coefficients form different sub-bands of EEG to formulate a virtual image. Later, this virtual image was fed into a CNN architecture to classify emotion. We made two distinct protocols; between these, protocol-1 was to classify positive and negative emotion and protocol-2 was to classify three distinct emotions. Overall maximum accuracy of 76.52% on valence and 76.82% on arousal was obtained by using internationally authorized DEAP dataset. We observed that the Convolutional Neural Network (CNN) based method showed state-of-the-art performance for emotion classification.

Approval

This is to certify that the thesis work submitted by **Md. Rabiul Islam** entitled "***Formation of Channel Correlation Based Significant Virtual Images from EEG Sub-bands Data to Recognize Human Emotion using Convolutional Neural Network***" has been approved by the board of examiners for the partial fulfillment of the requirements for the degree of **Master of Science in Engineering** in the Department of **Electrical and Electronic Engineering**, Khulna University of Engineering & Technology, Khulna, Bangladesh in July 2019.

BOARD OF EXAMINERS

24.07.2019

1. ______________________________ Chairman (Supervisor)

Dr. Mohiuddin Ahmad
Professor, Department of EEE
Khulna University of Engineering & Technology.

24/7/19

2. ______________________________ Member

Dr. Md. Rafiqul Islam (2)
Professor & Head, Department of EEE
Khulna University of Engineering & Technology.

24.7.19

3. ______________________________ Member

Dr. Bashudeb Chandra Ghosh
Professor, Department of EEE
Khulna University of Engineering & Technology.

24/7/19

4. ______________________________ Member

Dr. Md. Aminul Haque Akhand
Professor & Head, Department of CSE
Khulna University of Engineering & Technology.

24.7.19

5. ______________________________ Member (External)

Dr. Md. Monirul Kabir
Professor, Department of EEE
Dhaka University of Engineering & Technology.

Table of Contents

List of Table

List of Figure

List of Abbreviations

Abbreviation	Stands For
2D	Two-Dimensional
3D	Three-Dimensional
Adadelta	Adaptive Delta
Adagrad	Adaptive Gradient
Adam	Adaptive Moment Estimation
AI	Artificial Intelligence
ALS	Amyotrophic Lateral Sclerosis
BCI	Brain-Computer Interface
BGD	Batch Gradient Descent
CBN	Common Bayesian Network
CNN	Convolutional Neural Network
CNS	Central Nervous System
CREMA-D	Crowd-sourced emotional multimodal actors' dataset
CSP	Common Spatial Patterns
DBN	Deep Belief Network
DE	Differential Entropy
DE	Differential Asymmetry
DEAP	Dataset for emotion analysis using EEG, physiological and video signals
DNN	deep neural network
ECG	Electrocardiogram
EEG	Electroencephalogram
ELU	Exponential Linear Unit
EMD	Empirical Mode Decomposition
EMG	Electromyogram
EOG	Electrooculogram
FC	Fully Connected
FCM	Fuzzy C Means
FFT	Fast Fourier Transform
FKM	Fuzzy k-Means
GAN	Generative Adversarial Network
GSCCA	group sparse canonical correlation analysis
HMM	Hidden Markov Model
HVHA	High valence, High arousal
HVLA	High valence, Low arousal

IADS	International Affective Digital Sounds
IAPS	International Affective Picture System
IMF	Intrinsic Mode Function
kNN	k Nearest Neighbor
LVHA	Low valence, High arousal
LVLA	Low valence, Low arousal
MEMD	Multivariate Empirical Mode Decomposition
MLP	Multi-Layer Perceptron
MND	Motor-Neuron Disease
MSE	Mean Squared Error
Nadam	Nesterov Adam optimizer
NN	Neural Network
PANA	Positive Activation Negative Activation
PCC	Pearson's Correlation Coefficients
PSD	Power Spectral Density
RASM	Rational Asymmetry
ReLU	Rectified Linear Unit
RWE	Relative Wavelet Energy
SELU	Scaled Exponential Linear Unit
SGD	Stochastic Gradient Descent
STFT	Short-Time Fourier Transform
SVM	Support Vector Machine
WE	Wavelet Energy

Chapter **1**

Introduction

Chapter Outlines:

Chapter 1

Introduction

1.1 Introduction

Emotion means the instant mental state and interaction. The degree of pleasure, displeasure, excitation, sadness, etc. can easily be extracted from emotion. However, emotion is not so simple always because it varies with respect to personality, motivation, mood, temperament, event and with the different environment of surroundings. In spite of having these types of wide variety and complexity, emotion has such unavoidable importance that attracts the researcher significantly from the last two decades [1], [2].

A human may have different types of emotion. With the advancement of Human-Computer Interface (HCI), researchers want to classify emotion properly [3]. Among numerous different emotions, every emotion can be classified shortly in three types positive, negative and neutral. The classification of emotion is a very basic process of recognizing emotions separately. But distinguishing every separate emotion is not actually possible because human actually contains an infinite number of emotions [4]. However, we tried to recognize two or three distinct types of emotion in our work. The objectives, motivation of the work, importance and research challenges are illustrated in this chapter.

1.2 Emotion Recognition from EEG Signal

The main target of our work is to appropriately recognize human emotion from the Electroencephalogram (EEG) signal. Though out looking of the human face, body movement, gesture indicates the emotional condition it is significant to extract original emotion form automatic generated brain EEG signal. Actually, any type of thought, imagination, dream, plan of human being has a meaningful and indicative impact on brain signal [5]. Therefore, it can be said that EEG is the perfect choice for the extraction of human emotion.

1.2.1 Definition of Emotion

Emotion is a feeling that implies how we act for a certain instance. For different situation, a man may have a different feeling like happy, fear, angry or bored. The portion of a person's character which indicated the present mental states, behavior, and thought can be simply termed as emotion. If someone is laughing it means that he or she in on exciting emotional states. On the contrary, one is crying that means he or she is now passing a very odd situation and feel a sad type of emotional state. It is a very interesting matter that emotion is only one thing which differentiates human from machine. Because human can show numerous states of emotion whereas machine has none. In a nutshell, the degree of pleasure, excitation, displeasure, fear, sadness, etc. can easily be extracted from emotion.

1.2.2 Classification

There is no fixed idea which can illustrate the number of total emotions. According to recent research and publication very common type human emotional states are amusement, boredom, disgust, excitement, joy, satisfaction, sympathy, romance, horror, entrancement, confusion, awe, nostalgia, fear, empathetic, calmness, anxiety, admiration, awkwardness, triumph, sadness, nostalgia, interest, envy, craving, adoration etc.

Figure 1.1: The example of common emotions (left) and corresponding real view (right).

The main eight emotions are described in psychologist Robert Plutchik's theory [6]. In agreement of this theory, emotions are fear, anger, joy, sadness, disgust, surprise, trust and anticipation. During the year 1970s, psychologist Paul Eckman analyzed six basic emotions. He proposed six emotion for all human cultures such as happiness, sadness, disgust, fear, surprise, and anger.

'Happiness' is mostly termed as a pleasant emotional state. It is characterized by feelings of, joy, gratification, contentment, satisfaction. Oppositely, unhappiness is linked with anxiety, stress, loneliness, and depression. In happy time a man wants to alive more but don't want to pass many a time in an unhappy situation. Another type of emotion 'sadness' is characterized by grief, hopelessness, disappointing and crying. Sadness people often express his emotion by a different coping mechanism like as head in his hand or passing time lonely or get detached from people. During the sad feeling vocal expression get down. For ‘fear’ there is no common environment that can lead fear in a person’s mind. Interestingly, children may not have any fear where adult people may have. During the fear time mind becomes more alert, heart rate, and respiration increase, the body tries to be shrunken. ‘Disgust’ another type of emotional state that arises in the time of unpleasant taste, smell or any awkward situation. It can be treated as a sensation that indicates something revolting. ‘Anger’ means hostility, agitation, frustration, and antagonism. In addition, whenever someone gets something strange but desirable anything then feel another emotion named ‘surprise’.

We conducted our experiment with ‘DEAP’ dataset [7]. They recorded EEG data of 32 adult participant. This dataset consists of 16 different types of emotional states and this state was selected by each participant. The following 16 emotions are shown in Table 1.1.

Table 1.1: Different emotions recorded in 'DEAP' database

Pride	Relief	Sadness	Envy
Elation	Hope	Fear	Disgust
Joy	Interest	Shame	Contempt
Satisfaction	Surprise	Guilt	Ange

Another approach of recent research is to classify emotion according to the value of valence and arousal. The description of the terms valence and arousal is illustrated in chapter 3. However, just considering the value of valence and arousal emotion can be classified into four different types shown in Fig. 1.2. Here we also classified emotion into three classes named positive, neutral and negative.

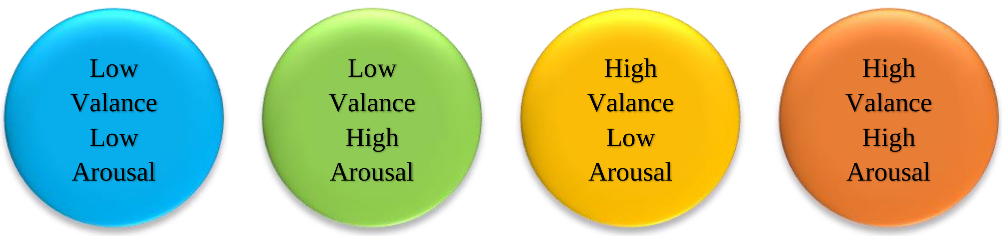


Figure 1.2: Classification of emotion according to valence & arousal.

1.2.3 Electroencephalogram (EEG) Signal

In 1875, a physician Richard Caton (1842-1926) proposed his achievement about the electrical activity of cerebral hemisphere of monkey and rabbits. The electric activity of the brain of dogs and rabbits is investigated by a Polish physiologist in 1890. In 1912 physiologist Vladimir Vladimirovich published the animal EEG and evoked potential of the mammalian [8]. In 1924, Hans Berger, a German Physiologist recorded the human EEG signal for the first time. After then with the improvement of technology, many approaches to measurement and recording are developing day by day.

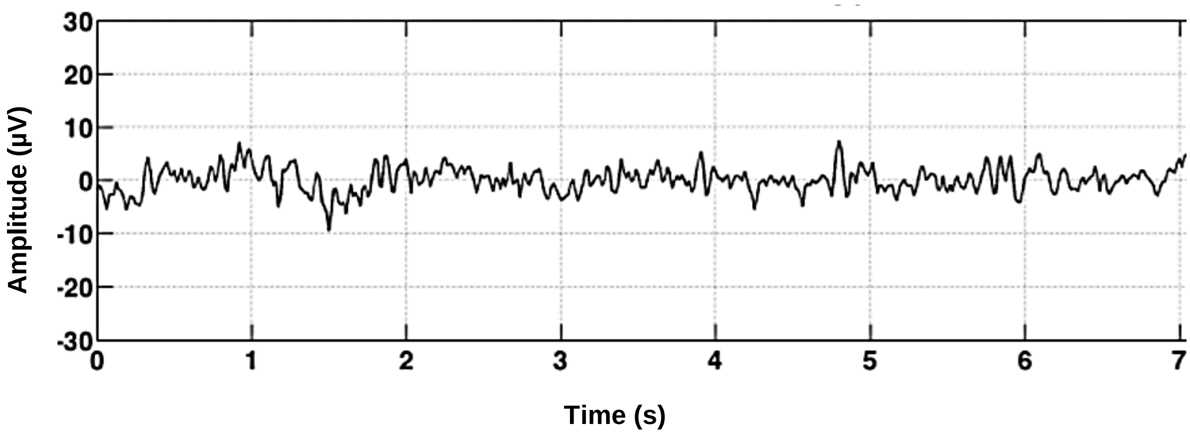


Figure 1.3: A sample of the EEG signal.

The Electroencephalogram is a waveform recording system that records the electrical brain activity from the scalp of human over a period of time. It measures the fluctuation of voltage generated from the ionic current flowing through the neurons of the brain [9]. The activity of this type of signal is very small in the range of microvolt. A sample of EEG is shown in Fig. 1.3. Clinically EEG is used to recognize epilepsy, sleep disorder, coma, brain death, anesthesia, tumors, stroke, etc.

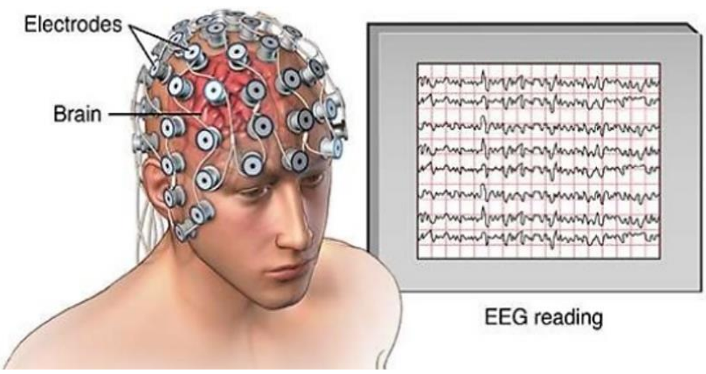


Figure 1.4: EEG signal acquisition from a human brain.

The EEG signal is recorded from the human brain and since the human head is non-linear in nature, EEG can be considered as a non-linear and non-stationary type signal. The method of recording the EEG signal is termed as EEG data acquisition. For EEG data acquisition many electrodes made connected with human head shown in Fig. 1.4.

1.2.3.1 Human Brain

For recording proper and significant EEG data one has to know about the brain autonomy of human. The brain is a surprising organ of the human body created by almighty that can control all

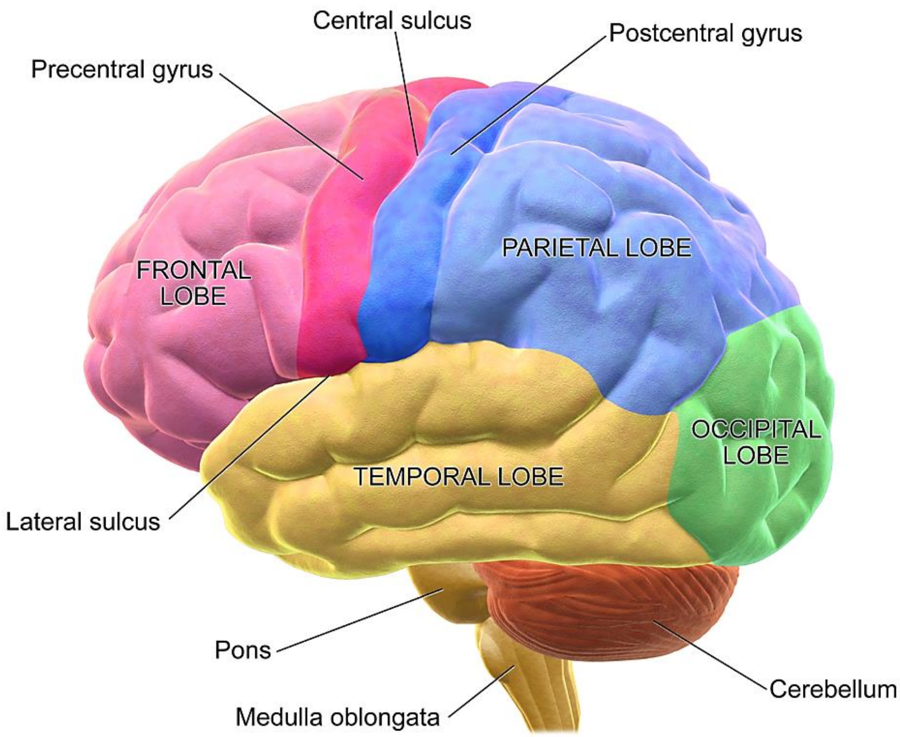


Figure 1.5: Different parts of the human brain.

of the function of the body. It gathers the information from the outside of the body and according to the information of surroundings, it can move and control any organ of the body. Smell, touch, sight, hearing, and taste can be easily sensed by the brain. The brain is the main part of the Central Nervous System (CNS). The brain has main three parts named cerebrum, cerebellum, and brainstem. The cortex is the surface of the cerebrum. The surface named cortex contains 70 billion neurons. However, the cerebrum hemisphere is composed of four different lobes namely frontal, parietal, temporal and occipital shown in Fig. 1.5.

It is important to point that no part of the lobe can do a single work alone because there is a very complex relationship among the lobes. The function related to different lobes is shown in the following Table 1.2.

Table 1.2: Activity of different lobes of the human brain

Name of Lobe	Corresponding Activity
Frontal Lobe	▪ Judgment, planning, problem-solving ▪ Personality, behavior, emotions ▪ Body movement (motor strip) ▪ Speech: speaking and writing ▪ Intelligence, concentration, self-awareness
Parietal Lobe	▪ Sense of touch, pain, temperature ▪ Interprets language, words ▪ Interprets signals from vision, hearing, motor, sensory and memory ▪ Visual and spatial perception
Temporal Lobe	▪ Understanding language (Wernicke's area) ▪ Memory ▪ Hearing ▪ Sequencing and organization
Occipital Lobe	▪ Interprets vision (color, light, movement)

At the time of EEG data acquisition, the electrode has to be connected on the scalp. The name of the electrodes comes from the name of lobe where the electrode is to be connected. Suppose 'FT' means 'Frontal' and 'Temporal' that indicates the position of this electrode is on the frontal-temporal position of the scalp. Very similarly 'PO' illustrates the location will be in Parietal-Occipital lobe.

1.2.3.2 Frequency bands of EEG signal

EEG signal can be treated as a composition of 5 sub-bands signal. Every sub-band indicates various mental states and condition. The different sub-bands name and their corresponding location, frequency range, and brain activity are shown in Table 1.3.

Delta sub-band signal mainly found in the frontal lobe. Its frequency range is up to 4 Hz but amplitude is highest. Though this band signal is significantly associated with deep sleep, it also may be present even in a working state.

The theta band is composed of 4 to 7 Hz of varying amplitude. Normally this type of band can be found in the event of drowsiness. In addition, hyperventilation and sleep also generate this banded EEG. Actually, theta banded signal can be found available in the frontal lobe. The Drowsiness, Imaginary, Enthusiastic, Fantasy are the common activity for this band.

Table 1.3: EEG sub-bands with related information

Name of Sub-bands	Range of frequency	Location	Field of activity
Delta wave	0-4 Hz	Frontal	Deep Sleep, Unconscious, Continuous attention (for babies)
Theta wave	4-12 Hz	Midline, Temporal	Drowsiness, Imaginary, Enthusiastic, Fantasy
Alpha wave	8-12 Hz	Frontal, Occipital	Closing the eye, Relaxed, Calm
Beta Wave	12-30 Hz	Frontal, symmetrically distributed on both sides.	Calm to intense to stressed, Aware of surroundings, Anxious, Thinking, Start to alert
Gamma Wave	>30 Hz	Frontal, Central, Somatosensory cortex	Cross-modal sensory processing, Alertness, Agitation, Short term memory for matching objects
Mu Wave	8-10	Sensorimotor cortex	Rest state motor neuron indication

The alpha band is ranged from 8 Hz to 12 Hz. This band was firstly named by Hans Berger at the time of his observation. This rhythm is considered as a starting point of clinical observation of EEG. The alpha band remains stable at 8Hz during the normal development of 3 years of age. For the rest year of life, it varies from 8Hz to 12Hz. This rhythm is poorly visualized for one

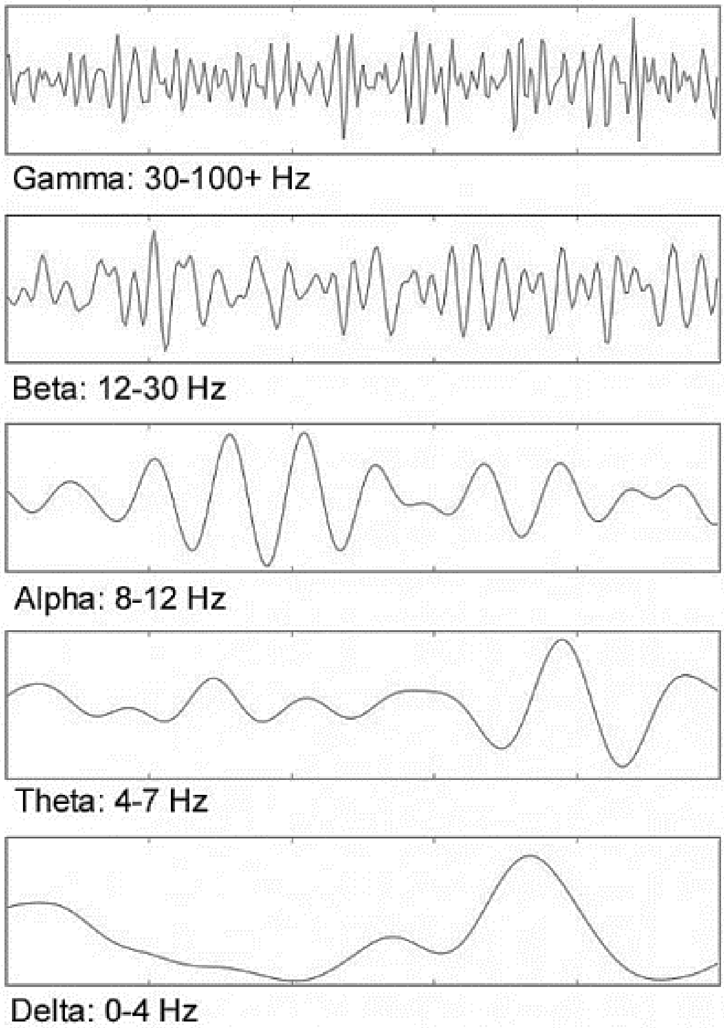


Figure 1.6: Sample waveshape of EEG frequency bands.

fourth of the normal adults. Normally the rhythm is found in the occipital lobe of the brain and at drowsiness time it shifted anteriorly. Normally, the voltage level for this band can be seen that is less than 15 microvolts.

For the beta band frequency ranges from 12Hz to 30 Hz. This type of signal can be found from the frontal lobe and distributed in two sides of the scalp. The voltage range is less than 20 microvolts is regarded as normal. But beyond 25 microvolts it is considered an abnormal condition. Beta activity is closely related to motoring action. Moreover, any kind of alertness, thinking, etc. also related to beta sub-bands signal.

The Gamma sub-band actually indicates the hyperactivity of the brain. Whenever human become so much excited or anxious for any matter it generates EEG beyond this band. Its frequency range starts from 30 Hz. Normally EEG sub-band signal greater than 30 Hz is considered as gamma sub-band.

Another special type of EEG frequency sub-band also considered in some special case. Its name is ‘mu’, frequency ranges of which lies between (8-10) Hz. It is asymmetric in type because it can be seen on one side. It indicates the synchronous firing of a neuron at rest states.

1.3 Objectives

The feature-based approaches of emotion recognition consist of the features in the time domain, frequency domain, joint time-frequency domain, etc. The Hjorth index, simple statistics, non-stationary index, etc. are calculated for the time domain. In the frequency domain, the popular methods are power spectral density (PSD), short-time Fourier transform (STFT), time-dependent energy, etc. and in time-frequency domain consists of wavelet-based features, wavelet packet features, intrinsic mode-based features, etc.

Recently, deep learning has become an indispensable part of the field of artificial intelligence to process signals and information. Numerous deep learning architecture like Deep Belief Network (DBN), Convolutional Neural Network (CNN), Deep auto-encoder, Generative Adversarial Network (GAN) achieve outstanding result compared to shallow models like SVM, MLP, k-NN, and CRF in many challenging fields like pattern recognition and image processing. In modern time for analyzing physiological signal (i.e. EEG, EOG, ECG, EMG) various deep learning-based architectures are applied successfully and attained significant results.

This research work ensures two important facts: i) formulate a significant virtual image from EEG data and ii) design a deep architecture based Convolutional Neural Network model to classify discrete emotion. Therefore, how to extract the proper feature to classify emotion and how to enter these featured images into a CNN model to predict human emotion accurately. Another key problem lies in how to construct a classifier to automatically learn the changes from the EEG multidimensional features over time and classify the changes into different emotional states. To meet the challenges, we proposed the emotion recognition model to perform the following objectives:

- To propose a novel method to construct a sequence of two-dimensional images from the one-dimensional EEG signals variation in emotions described in section 3.7 namely formation of virtual images.

- To build a CNN based predictive model described in the “section 4.4: our proposed CNN model”, to undertake the recognition of the human emotion from the EEG multidimensional feature image sequences.
- To conduct empirical research on open-source database DEAP [7] using the proposed method to demonstrate the significant improvements over current state-of-the-art approaches in this field. The improvement and the comparison of the result are written in section 5.3.

1.4 Research Challenges and Contribution

Normally feature-based emotion recognition requires a strong effort to design the perfect feature or feature set related to the classification of emotion. To curtail the manual human effort, we designed a model by using a virtual image from EEG with CNN. Here, it can be mentioned that CNN outperforms with image type 2D data whereas EEG data are the 1D type. Therefore, it is essential to convert the 1D data into significant 2D data. Here, we used the CNN model upon a virtual image formed by Pearson’s correlation coefficients (PCC) on internationally authorized DEAP dataset [7].

- In our thesis work, the main focus is to propose an effective method which can convert 1D to 2D data and of course contains the significant information simultaneously. The idea behind the work is to propose the 2D channels correlation matrix among 32 EEG channels using PCC. This formulates a grayscale type image of size 32×32.
- In addition, if we take this type of 2D PCC matrix for three most relevant sub-band (alpha, beta, gamma), it can be treated as a color image which is termed here as a virtual image. This virtual image contains very significant information which helps to classify emotion using CNN effectively.

1.5 Organization of the report

The plan of the remaining report is just like that.

- The introductory information, objectives, some common terminologies, etc. are illustrated in chapter 1. Chapter one also covers the objectives and research challenges and contribution also.

- The history of the first research of emotion and the application and importance of emotion recognition is described details in the second chapter. Moreover, the relevant research, existing methods are described in chapter II. For future researchers, the list of software to analyze the EEG signal is also given in this chapter.
- The methodology is illustrated in chapter III. In this chapter, the description of the dataset and the preprocessing, reshaping and decomposition methods are written sequentially. In next, the details procedure of formation of virtual images are depicted chronically. This chapter also covers the experimental protocol.
- The proposed CNN architecture and its details information are explained in Chapter IV. The working of CNN method and our proposed CNN model is interpreted in section 4.3 and 4.4 respectively. The details information of our optimization algorithm and loss function are also reported in this chapter.
- The result of our proposed method and discussion are reported in chapter five. Furthermore, the result of emotion recognition of the previous research and our research is compared in this chapter.
- At last, chapter VI concludes the report.

Chapter **2**

Literature Review

Chapter Outlines:

Chapter 2

Literature Review

2.1 History of research of Emotion

William James was an American philosopher and psychologist who wrote on the topic of psychology philosophy of pragmatism. Another psychologist and physician Carl Lange was a Danish citizen. Both of the researcher James (1842-1920) and Lange (1834-1900) were considered as the most potential theorists of the late 19^{th} century [10]. After completing separate research, they were capable to develop a hypothesis on the origin and nature of emotion named James-Lange theory. The concept of basic emotions was introduced by affect theory and script theory that was developed by Silvan Tomkins [11].

2.2 Importance's of Emotion Recognition

Actually, emotion is such a matter which differentiate a man from machine. Because human has a lot of emotion but the machine has none. The current mental state, though, feeling, behavior, etc. are described by emotion. It differs from person to person. In recent years, [12] owing to the outstanding development of emotion-based noninvasive Brain-Computer Interface (BCI) and the wide advancement of signal and data processing technology the analysis of EEG based emotion recognition has become a very attractive issue in the field of artificial intelligence and biomedical engineering Fig. 2.1.

Since understanding and the technology of emotion are advancing there are growing numerous branches of using automatic emotion recognition system. Recent years are the year of

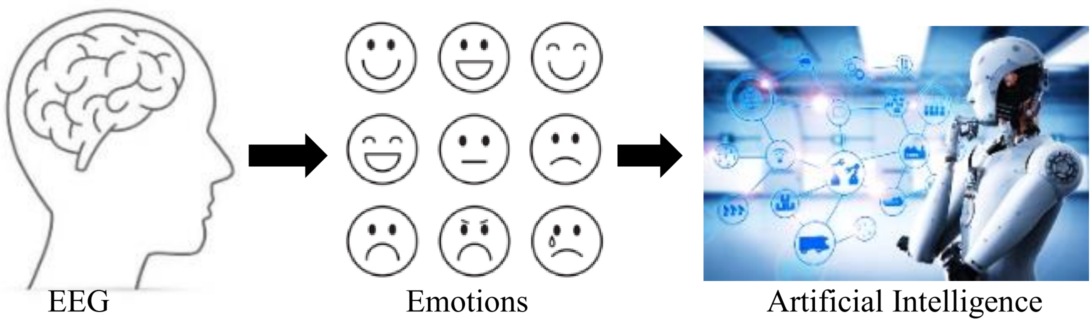


Figure 2.1: Emotion extracted from EEG is applicable in many branches of AI.

advancement and improvement of technology. Now at the era of technological development, the research of emotion recognition from EEG has become very attractive and basic research due to its non-invasive feature. Information and important data can easily be extracted from the human brain and can be analyzed for knowing the internal truth of human.

Emotion recognition has become an indispensable part of research in the field of neuroscience, computer science, cognitive science and physiology [13]. For neuroscience, the aim of the research is to explore out the neural circuit and to find out the brain mechanism. In computer science, there is various kind of application like lie detection or mental stress detection and many more described in later. In the field of cognitive science, the aim is to bring out the present mental state of human cognition.

It may be noted that cognitive science is the scientific study of intelligence or mind that examines the nature, the function, the task of mental action or process. Since human behavior is largely dependent on emotion and physiologist wants to analyze the condition of the human mind; recognition of emotion has become a hot topic in the field of physiology. Many physiological experiments demonstrated the relation between emotion with EEG signal [14] - [16].

2.2.1 Application of Emotion Recognition from EEG

The people who are physically handicapped and are not able to speak the only way to express their emotion with the EEG signal. In addition, the people who are able to speak but there remains some kind of inability like speech disorder or speech impediment they can also express their emotion using EEG signal. The persons who have the problem of mutism they can't speak and unable to express their present mental state and feeling even their present necessity of food or sleep or rest like that. Therefore, recognizing proper emotion from EEG the mental states can easily be brought out from the brain.

The famous physicist and cosmologist Stephen William Hawking (1942-2018) were affected by Motor Neuron Disease (MND) also termed as Amyotrophic Lateral Sclerosis (ALS). For his communication, a combination of several computer-aided software was developed among which most of them sense the brain signal.

Figure 2.2: Emotion from EEG is essential for handicapped people.

Today is the era of online schooling. Numerous courses and classes on different topic are available now online. In this context, the student mental attention can be extracted by recognizing proper emotion. Not only for online courses or classes but also for normal offline classes the attention and mental thought can easily be extracted from the EEG signal.

Figure 2.3: Attention of student can be extracted from emotion.

Emotion recognition or classification from EEG may be used as a partial detector of lies from the truth. Whenever it is essential to know important information truly from any person but there is no video recording, then lie detector is a fundamental prerequisite. Like as in a court a true killer always denies his guilt of murder or a guilty people always denies any type of guilt. At this time by recognizing proper emotion, a lie can be detected from emotion. Because, emotion is the reflection of present thinking, present mental behavior, etc. Therefore, it can be said that very near future the emotion recognition from EEG will be available in every court to justify the truth or lie speech of the accused.

Figure 2.4: Detection of lie or truth in court from emotion.

There remains another special application of emotion recognition from EEG. That is extraction the person's interest for shopping. The authority of many companies always tried to find out which type or how many aged persons have an interest in their product. This can also be extracted easily by recording emotion. Suppose that in a showroom of 'SAMSUNG' mobile a customer showing a mobile named 'Samsung Galaxy J2' and at this time emotion is an angry type or bore type it indicates that he is not attracted by this model.

Figure 2.5: Person interest in shopping extracted from emotion.

On the contrary, if the emotion of most of the people have a happy or positive type in the time of showing another type of phone like 'Samsung A30' that indicated this product has top popularity. Very similarly many important information's which are very important for the

Figure 2.6: Alertness level of the driver from emotion.

owner of a company can be recorded by recording the emotion at the time of shopping of customers.

Driving is a very serious issue where attention must have to at a high level at every moment. The unawareness or unconsciousness of a single moment can occur a dangerous accident which may take some lives. Emotion from EEG can be treated as a scale of attention. If the EEG of driver indicates a very high value of valance and arousal it indicates high attention, low valued valence and arousal indicate less attention. Simply neutral type of emotion is a clear indication of less alertness of a driver.

Our work is important and applicable in the fields described previously. However, the high aged people can be monitored remotely by recognizing emotion by our method for counseling and giving healthcare of patients like autism. In addition, this type of emotion recognition method can be easily applicable to determine the drowsiness of the driver during driving and the alertness of student during class. Although properly recognition of emotion is not invented yet now in a full 100 percent accuracy, the research and development section of every country tried a lot to develop a system which can be used in practical application. However, not only for this stated field but also many more fields of artificial intelligence have significant importance of emotion recognition from EEG.

2.3 Why EEG is the Best?

Multitudinous approaches have been applied to extract emotion like facial expressions [17], [18]; gesture and speech signals [19]; EEG signals [1],[12], [13], [14]; autonomous nervous signals [20] etc. For facial expression and speech signals, subjects need to express emotion explicitly. From the facial expression, only this type of emotion can be classified whose effect will change the structure of the human face. But subjects who fell happy internally but don't express it by his mouth of face can't be classified. So facial expression is not considered as a good emotion recognizer raw signal.

For speech signal subject's emotion can be classified with the help of the intensity of the voice. For that reason, this type of emotion recognition method is not applicable to mutism persons [21]. The people who are not able to speak remains out of this emotion recognition method.

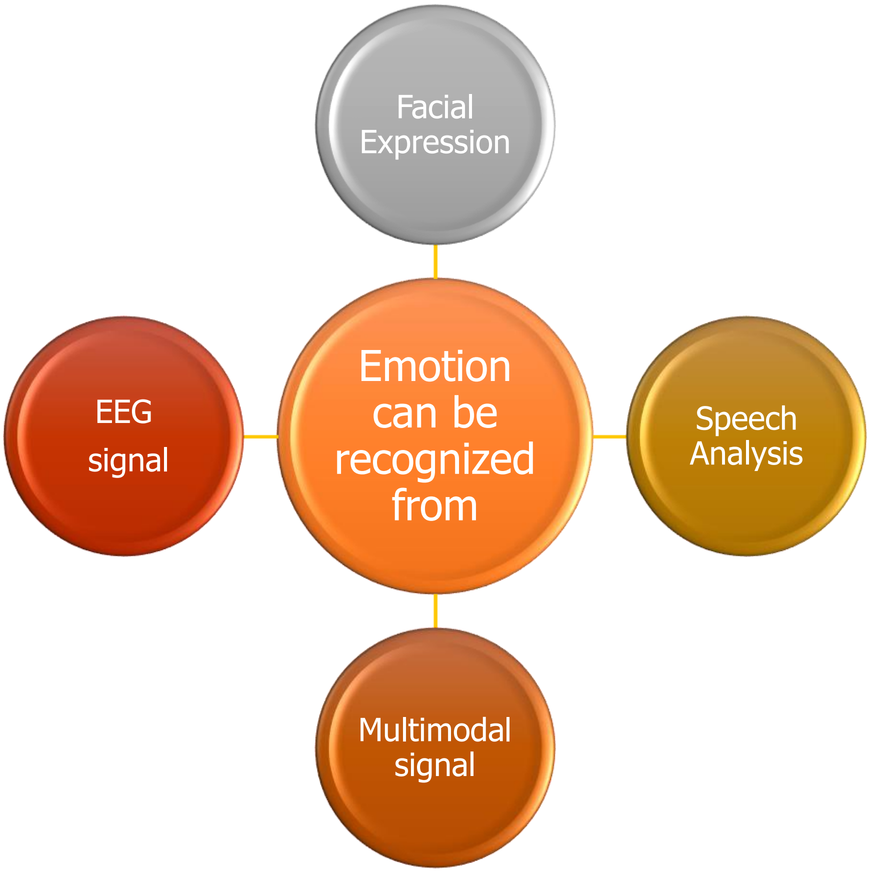


Figure 2.7: Different type of raw signal for emotion recognition.

Some researcher tried to classify emotion from gesture [22] and body movement [23]. This approach also not free from some problem. Peoples who are physically handicapped have no body language and gesture to express emotion. So, emotion recognition approach dependent upon gesture don't cove this type of people. Emotion from the autonomous nervous system is not too much applicable because of its difficult acquisition technique.

On the other hand, subjects have no way to control the automatically generated EEG signal. People who are unable to speak or unable to express their emotional stated by gesture and posture are not out of this type of emotion recognition method. Moreover, EEG signal acquisition is not too much difficult like multimodal signal acquisition. Therefore, EEG is the best and inimitable signal for emotion recognition for any kind of people at any time.

2.4 Existing Methods of Emotion Recognition

Emotion recognition from EEG has become an attractive field of research in recent years. Multiple methods and approaches are been involved by different researchers [24]. In the field of recognition and classification of emotion, the applied approaches can be classified primarily into two basic types.

- Feature-based approaches
- Deep architecture-based approaches

A complete list of existing methods is shown in the following Fig. 2.8.

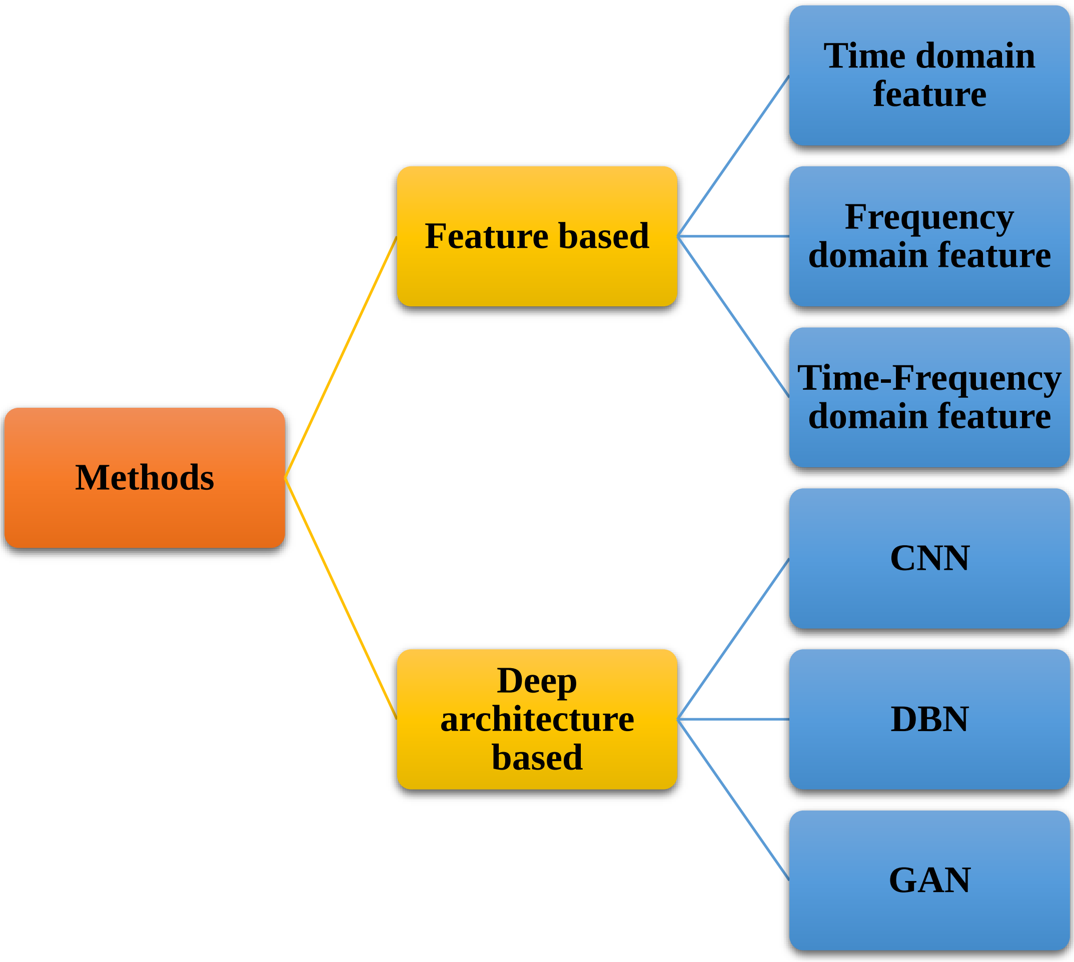


Figure 2.8: The list of existing methods of emotion recognition.

2.4.1 Feature-Based Methods

In feature-based methods many significant and related features are extracted from the raw EEG data. Then the extracted features are fed into a classification model for recognition of emotion. The feature-based method consists of time-domain feature, frequency-domain feature and joint time-frequency domain feature.

2.4.1.1 Time Domain Features

Simple statistics includes the features namely mean, median, standard deviation, mode, variance, minimum, maximum, etc. are examined for the frequency-domain feature. These are called simple statistics. A few researchers calculated the non-stationary index and developed a system which worked on the base of these index as the feature set [25]. Some researchers found a better result with taking Hjorth index as a feature [26] -[28]. The Hjorth index performs well corresponding to simple statistical type feature. The time-domain features are sensitive for a time but not worked accurately in anti-noise platform. So, the performance of time-domain features is not satisfactory at all.

2.4.1.2 Frequency Domain Features

Since EEG is a non-stationary type signal so time-domain features are not so meaningful for classification or recognition of this type of signal. Frequency domain feature contains more relevant information for this type of signal. In the frequency domain, the popular methods are Power Spectral Density (PSD), Fast Fourier Transform (FFT), Short Time Fourier Transform (STFT) [29], etc. In the field of emotion recognition, the spectrum analysis is also a popular analysis using Fourier transform [30]. Some scholars tried to classify emotion using Eigenvector method which is suitable for transient and stationary signal [31]. A few types of research [32], [33] are also available which cover the autoregressive model which give good frequency resolution but the spectral estimation is difficult. Actually, autoregressive analysis is perfect for this type of signal which has sharp spectral features [31]. The main advantages of frequency domain are that it has comparatively high anti-noise ability. The problem is that the calculation complexity is higher, complex and not cost-effective.

2.4.1.3 Time-Frequency Domain Features

According to the research in [31], the Fast Fourier Transform and Auto-Regressive model fall a victim of slow computing and inability to analyze the non-stationary type signal. Frequency domain features have comparatively good performance in anti-noise ability but joint time-frequency domain features perform more than this. The joint time-frequency domain consists of wavelet transform [12], [34] Basically, EEG is a signal whose type is non-stationary and non-linear. The analysis of this type of signal is challenging and complicated. Recently, the wavelet transform has become a very popular method of analysis due to its good performance both in time and frequency domain [34]. Again, wavelet transform can be classified into two types i) Continuous Wavelet Transform ii) Discrete Wavelet Transform. We also made a classification using wavelet analysis but it is not including in this report.

2.4.2 Deep Architecture-based Methods

A recent development in the field of machine learning attracts researcher to solve many complicated problems by using a neural network. Moreover, the deep learning architectures are capable of solving a particular task easily though it is affected by many factors and high complexity. Therefore, deep learning has emerged a great impact on signal and image processing [35]. Some common deep learning architectures are

- Convolutional Neural Network (CNN)
- Deep Belief Network (DBN)
- Gaussian Adversarial Network (GAN)
- Deep auto-encoder etc.

2.5 Relevant Research

EEG signal is non-linear, non-stationary and temporal asymmetry type signal, computation and analysis of this type of signal are too much challenging and difficult. In general, the conventional feature extraction method considered a number of feature or feature set to train the recognition system. Whereas deep learning eliminates the difficulties of a selection of proper and significant features. It allows the machine to automatically learn the feature from the input data set and transfer its learning to the classifier.

In the last recent years, there remains a trend of research of emotion recognition by selecting proper features and shallow machine learning algorithm like Support Vector Machin (SVM), k Nearest Neighbor (kNN), Decision Tree or Multi-Layer Perceptron (MLP).

Hadjidimitriou et al. experimented on basis of the spectrogram, Zhao Atlas Marks transform and Hilbert-Huang Spectrum to estimate the liking and familiarity ratings [36]. Li and Lu proposed an emotion recognition method taking high importance on gamma-band EEG signal [37]. In this experiment, the emotions are recorded showing pictures of a smile and cry facial expression. The proposed Common Spatial Patterns (CSP) with Linear SVM to differentiate emotions namely sadness and happiness. Soleymani et al. developed emotion recognition method which is based on user-independent analysis and achieved a good result in terms of accuracy like 68.5% for valence and 76.4% for arousal at the classification of three class [38]. Daun et al. investigated the performance of emotion recognition using the combination of Differential Entropy (DE) and symmetrical electrodes and compare results with the traditional frequency-domain features. Then concluded that for emotion recognition from EEG the most relevant feature is Differential Entropy (DE) [39]. Lu et al. advanced of using the combination of EEG and eye movement data to classify emotion. In their multimodal emotion recognition research, they investigated the relation of 16 different type eye movements and tried to correlate with positive, negative and neutral emotion [40]. Knyazev et al. suggested to classify emotion considering the gender variation with facial expression and event-related synchronization [41]. Zheng et al. [42] investigated stable patterns of EEG using machine learning and systematically evaluated the performance of different methods of feature extraction, selection, and smoothing. Atkinson et al. [43] suggested valence arousal-based method which was actually the combination of mutual information-based feature selection methods and kernel classifier. The proposed EEG based Brain-Computer Interface to explore a set of emotion type and incorporates additional features relevant for signal preprocessing and classification task. Mert et al. [44] proposed the advanced properties of Empirical Mode Decomposition (EMD) and its multivariate extension. They investigated several time-frequency domains approaches like as power ratio, power spectral density, entropy, Hjorth parameter and correlation features on the multichannel Intrinsic Mode Function (IMF) extracted by Multivariate Empirical Mode Decomposition (MEMD). In the paper of Katsigiannis et al. [45] proposed a multichannel EEG and ECG signals recorded at the time of affect elicitation showing the audio-visual type stimuli. They recorded the database consisting of the value of valence, arousal, dominance from the portable, wearable, low-cost EEG and ECG data of 23 participants.

From another point of view mainly emotion classification task can be classified into two basic categories: one is based on supervised learning methods and the other is based on unsupervised learning methods. The supervised learning method means the system is firstly trained by some given input data with its output level whereas the unsupervised learning methods are given only the input data without a corresponding level. In unsupervised machine learning, the system classifies based on the different approaches of the clustering algorithm. The researcher He et al. [46] used a supervised learning method named Common Bayesian Network (CBN) to differentiate multiclass EEG signals. Another group of researchers Velchev et al. [47] used Support Vector Machine a supervised learning-based algorithm to classify emotion automatically by selecting Hjorth parameter on theta, alpha, beta and gamma taken from the certain channel with the best performance of 80%. Chen et al. also used a supervised learning-based method named SVM and adopt EEG connectivity between electrodes and tried to identify persons affective states with stimuli like music video and images [48]. Some researchers used another supervised learning algorithm like Hidden Markov Model (HMM) to recognize emotions [49]. The unsupervised learning methods include Fuzzy C Means (FCM), transfer learning [50], self-organizing map, K means, etc. Murugappan et al. [51] used Fuzzy C-Means and Fuzzy k-Means (FKM) clustering methods on a reduced number of 24 biosensors from 64 biosensors using the wavelet transform. Khosrowabadi et al. [52] used an unsupervised learning type method namely self-organizing map and investigated the performance of emotion recognition collecting 8 channels EEG data from 26 right-handed subjects.

In addition, with this type of research recently deep learning is attracted by most of the scholars. Martin et al. [53] proposed unsupervised learning architecture named Deep Belief Network (DBN) to reduce the complexity and necessity of multimodal sleep data. They also recover the fact of time consumption in the classification stage. Li et al. worked using differential entropy with DBN and compared with five baselines improvement of 11.5% and 24.4% [54]. Martinez et al. used the skin conductance and blood volume pulse signal with an effective convolutional neural network to classify four emotional states [55]. Suwicha et al. proposed a deep learning network and used principal component analysis to alleviate overfitting. They investigated better performance compared to SVM and naive Bayes classifier [56]. Tripathy et al. [57] also used DNN and CNN over the DEAP dataset to classify emotional states into two and three classes. They conclude the Neural Network (NN) could be the robust classifier for brain signal. Yanagimoto & Sugimoto (2016) used CNN to classify two classes of emotional states namely positive and negative. The authors in [58] proposed a group sparse canonical correlation analysis (GSCCA) for channel selection and feature extraction. A deep neural network (DNN) has been proposed for channel selection based on the distribution of mean absolute weights in [59]. The researcher in [60] proposed a method of using channel correlation but they didn’t use the emotion relevant sub-band data. Another researcher used the deep and convolutional neural network on mean, median, standard deviation and much other time-domain feature [57], but time domain features are not so significant actually. The author in [61] calculates PCC of multichannel EEG data taking four sub-bands in consideration. But since emotion is highly related with beta and gamma sub-band and moderately related with alpha sub-band and very poorly related with theta sub-band; we took only beta gamma and alpha sub-bands in consideration in spite of considering four sub-bands, which makes our information more significant.

2.6 Modeling of Emotion

Several researchers model emotions in different aspects of view. Simply emotion can be classified into two categories, one is a positive emotion and other is the negative emotion. A few argued only positive and negative emotions are not sufficient. Emotion may be not positive nor negative. This category can be termed as a neutral type of emotion. According to this opinion there are three types of emotion in total i) positive emotion, ii) negative emotion and iii) neutral emotion. For modeling emotions there remains a different model as follows.

- Circumplex Model
- Vector Model
- Plutchik,s Model [57]
- Positive Activation Negative Activation (PANA) Model
- PAD Model

Among these models, the Circumplex Model is more popular. The circumplex model of emotion indicates that every emotion is located distinctly in a circular two-dimensional space. From another point of view, according to dimensionality, the emotion model can also be categorized into two types.

- 2D Model of Emotion
- 3D Model of Emotion

The 2D (Two-Dimensional) model classify emotions based on two-dimensional data one is Valene and another is arousal. On the contrary, the 3D (Three-Dimensional) model deals with valence, arousal and with dominance. We classified emotion according to Russell's Circumplex model of emotion which is 2D in dimensionality.

James Russell was the developer of Russell's Circumplex Model of Affect in 1980 [62]. This model consists of two axes shown in Fig. 2.9. The horizontal axis represents how much pleasure the subjects are. From left to right it scaled as unpleasant to a pleasant emotional state. The vertical axis is for indicating the activation of subjects. The downside represents the less activation and upward represents more activation.

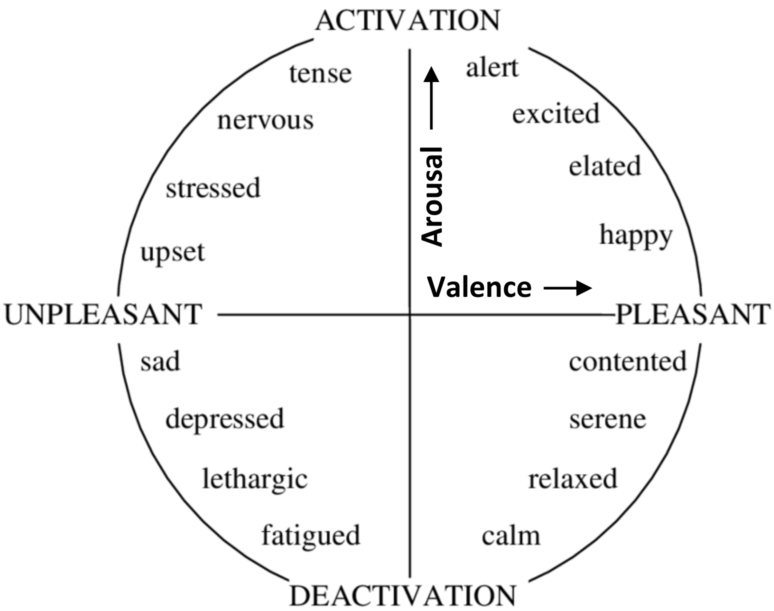


Figure 2.9: 2D Russell's circumplex model of affect for emotion classification.

According to the theme of Russell’s circumplex model, some researchers classified emotion based on the value of valence and arousal. Regarding this low and high value of these two-dimension emotions are classified into four categories.

- High valence, High arousal (HVHA)
- High valence, Low arousal (HVLA)
- Low valence, High arousal (LVHA)
- Low valence, Low arousal (LVLA)

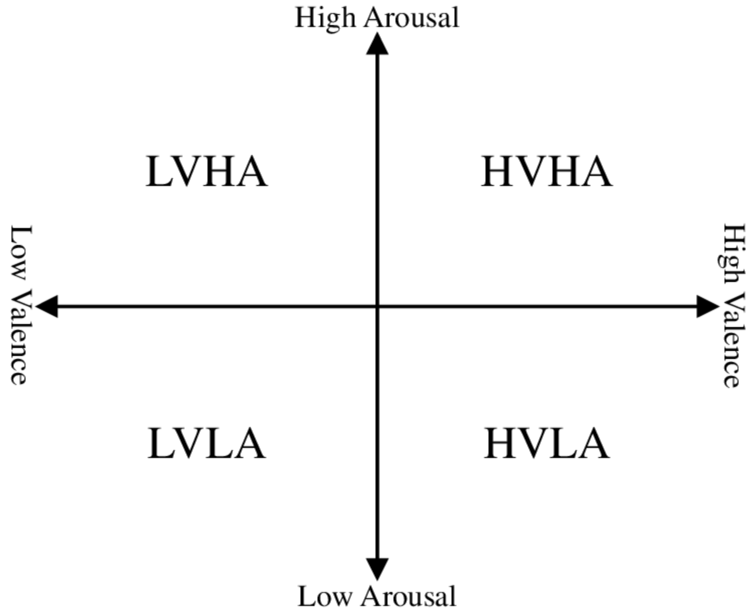


Figure 2.10: Categories of emotion according to the values of valence and arousal.

The two-dimensional circular space is divided into four categories which indicates the four quadrants of emotions actually shown in Fig. 2.10.

The 3D Model of emotion not only consider the level valence and arousal but also consider the values of dominance. In 3D space, the emotions are distinctly located. Suppose for Surprise the level of valence, arousal, and dominance all are high.

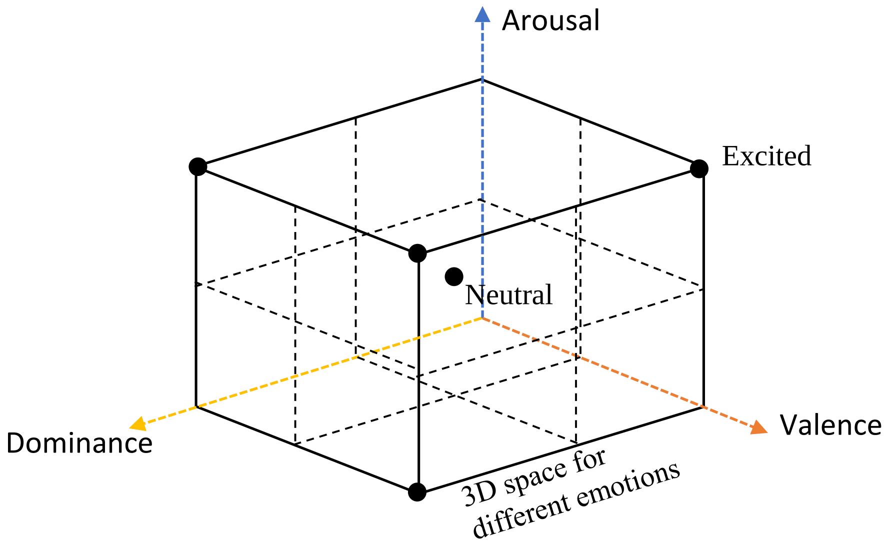


Figure 2.11: The 3D model of emotion.

For the high value of valence and arousal but the low value of dominance indicates excitation. The following Fig. 2.11 represents a 3D model of emotion where every point of 3D space indicates various emotions.

2.7 Software for EEG Signal Analysis

Research-based on EEG signal is becoming attractive day by day. For an introductory analysis, some researcher falls in a trap in finding the EEG signal analyzing software. Several software is available to analyze the EEG signal, that is listed in Table 2.1.

Table 2.1: List of software for EEG signal analysis

Name of the software	Additional Information
BESA Research	BESA GmbH
BIOPAC ACQKnowledge	BIOPAC Inc.
Brainstorm	University of Southern California, Los Angeles, CA
BrainVision Analyzer	Brain Products GmbH
BrainVoyager	Brain Innovation B.V.
Cartool Software	Biomedical Imaging Center Lausanne, Geneve.
Curry 7	Compumedics Limited
EEGLAB	Swartz Center for Computational Neuroscience, UCSD
LORETA	A key Institute of Brain-Min Research, Zurich

Chapter **3**

Methodology

Chapter Outlines:

Chapter 3

Methodology

3.1 Introduction

With the advancement of a different branch of Artificial Intelligence (AI), the human brain-controlled automation system is developing drastically. In previous, the human brain was just a physiological matter. No controlling or automation from the outer side of the body was related to the human brain. On the contrary, the present trend of biological and automation research tries to do everything according to our thinking or imagination. Just from this type of aspect, we tried to extract our real emotion so that it can be used in various branch of AI regarding necessity.

Our main purpose is to record the appropriate emotion. For extracting the real emotion of human being EEG signal is the fundamental prerequisite. Here the EEG signal is decomposed firstly into relevant sub-bands signal. Then formed significant virtual images to train a Convolutional Neural Network (CNN). CNN with virtual images from EEG performed at an expected level to recognize emotion.

3.2 Brief Overview of the Proposed Method

As our main objective is to reach on emotion from EEG, the EEG signal is the first and essential requirement. The EEG data acquisition from our local environment would be possible but this type of data was not internationally approved and validated. In addition, the comparison would not be possible for self-recorded data. As a result, we worked with DEAP dataset which is internationally authorized.

Since the raw EEG data may contain the EOG artifacts or power line noise, firstly the preprocessing is performed on the raw data to clarify the noises and unwanted signal. The dataset contains the EEG data of 32 participants. For every participant or subjects are shown 40 different emotional videos as stimuli. For every participant and every emotional video 40 channels, EEG data are recorded. Among them, we considered only 32 EEG channel data for further processing. Then EEG data are decomposed into five sub-bands. From the sub-bands, only three relevant sub-bands are taken for the formation of the significant virtual image. Pearson’s Correlation Coefficients of 32 channels data are arranged properly to formulate our desired virtual images. Then the virtual images with labeled emotions are taken as the input on a CNN architecture-based classifier.

Figure 3.1: The simplified block diagram of the proposed method.

After processing from the classifier, the classified emotion can be found as an output. The overall simplified process is shown in the following Fig. 3.1.

3.3 Dataset Description

This section illustrates the ins and outs of EEG data recording. For recording the EEG data, the factor that varies the number of electrodes, electrode placement system on the scalp, types of stimuli, frequency of recording and the device of signal acquisition. Electrodes are sometimes termed as a channel. That means the number of channels will be equal to the number of electrodes on the scalp. The electrode placement system is described in the next sub-section. Stimuli are the exciter of the human brain according to which the variation of EEG signal will occur. Images, videos, audio, gesture, posture or any outward environment may be treated as stimuli. Here in our working dataset, the video data was the stimuli. Frequency of recording should neither in such a large value that file size will become very large nor a minimum value that will be unable to contain necessary information. It should maintain a medium value. For example, for DEAP recording the initial recording was done in 512Hz. But after preprocessing it was down to 128Hz. The data acquisition methods and the information of public datasets are described in the following two sub-sections.

3.3.1 EEG Data Acquisition System

For recording EEG data some common acquisition systems are Biosemi ActiveTwo, Emotiv Epoc+ headset, EEG module of Neuroscan, etc. Among the system, the Biosemi ActiveTwo system is more familiar and popular to researchers which are shown in the following Fig. 3.2. About 37.1% of all researchers used this method till 2017 [1].

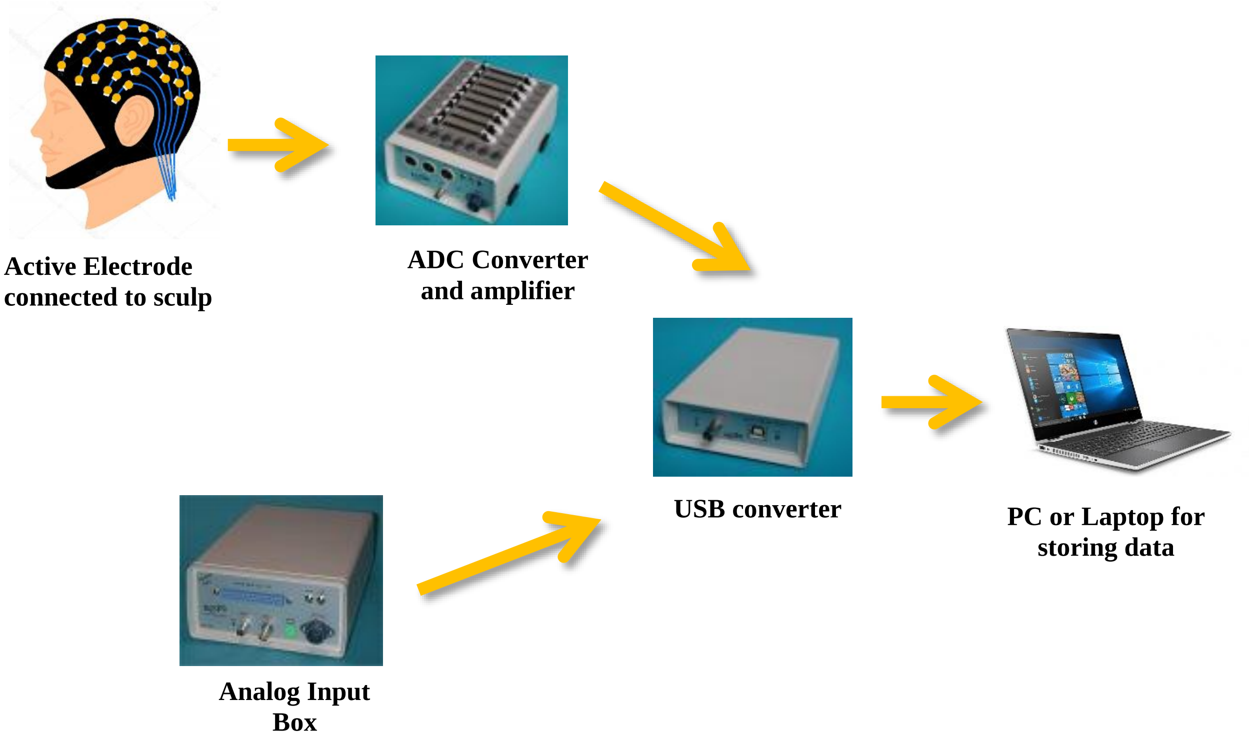


Figure 3.2: Biosemi ActiveTwo EEG data acquisition system.

Figure 3.3: Emotiv Epoc EEG data recording equipment.

Another popular data acquisition method is of using Emotiv Epoc shown in Fig. 3.3. About 16.1% used the signal where the acquisition equipment was Emotiv Epoc headset. It is wearable, portable, wireless in addition low cost. These features make it attractive to recent researchers. Furthermore, its portability and wireless feature are too convenient to use.

3.3.2 Available dataset for Emotion Recognition

For comparable and next-level research the first requirement is to collect or record EEG dataset. Whenever research will be done on the same raw data then the performance or accuracy of the research can be compared. As a result, some scholars build a research-level data and made it available internationally without any cost. Among them, some common and popular dataset are listed in the following Table 3.1.

Table 3.1: Publicly available dataset for Emotion Recognition

Name	**Stimuli**	**Year**	**Short Notes**
Dataset for emotion analysis using EEG, physiological and video signals (DEAP) [7], [63]	Video	2012	A multimodal dataset consists of EEG and peripheral physiological signal
International Affective Picture System (IAPS) [64], [66]	Color Photo	2008	Provides emotional stimuli for emotion and attention
International Affective Digital Sounds (IADS) [65], [66]	Sound	2007	Provides ratings of pleasure, arousal, dominance for acoustic stimuli
Crowd-sourced emotional multimodal actors dataset (CREMA-D) [67]	Audio, Visual, Audio-visual	2014	Consists of facial and vocal emotional expression in sentences spoken
(CREST-ESP) [68]	Speech	2001	The main target is to collect a dataset of spontaneous and expressive speeches, also determine the acoustic variables reflecting emotion and attitudes
MAHNOB-HCI [59]	Video, Image	2011	A multimodal database where eye gaze, video, audio peripheral and physiological signals are synchronized.
Reading-Leeds Database [70]	Speech	1994	Dataset of genuine emotional natural or near-natural speeches
Belfast Database [71]	laboratory-based emotion induction tasks	2012	Dataset to recognize emotion from facial and vocal signal regarding the emotional task.

3.3.3 Details Information of DEAP Dataset

Our method is conducted by 'DEAP' database [7]. 'DEAP' is a widely useable free dataset for emotion analysis which contains EEG with the peripheral physiological signal. The dataset is formed with two main parts. Firstly, the ratings from an online self-assessment where 120 one-minute extracts of music videos were each rated by 14-16 volunteers based on arousal, valence, and dominance. Secondly, the participant ratings, physiological recordings and face video of an experiment where 32 volunteers watched a subset of 40 of the above music videos. EEG and physiological signals were recorded and each participant also rated the videos as above. For 22 participants frontal face video was also recorded.

Table 3.2: The summary of the 'DEAP' dataset

No. of participants	32 (16 female)
No. of video	40 (1 min long)
No. of distinct emotion	15
Rating scale	Valence, arousal, dominance, liking, familiarity
Rating value	1-9, except familiarity (1-5)
Recorded signals	32- channel 512 Hz EEG; physiological signals; Face video (no. 22)
EEG acquisition medium	Biosemi ActiveTwo

'DEAP' recorded EEG data of 32 participants using BioSemi ActiveTwo system. Among this 32 participant, 16 are male and 16 are female whose age ranges from 19 to 37 years. Each of the participants is watched 40 music videos of an expert musician having a 60-second duration of each. A total 48 channel including 32 EEG channels, where 12 peripheral channels, 3 unused channels, and 1 status channel recorded the raw data.

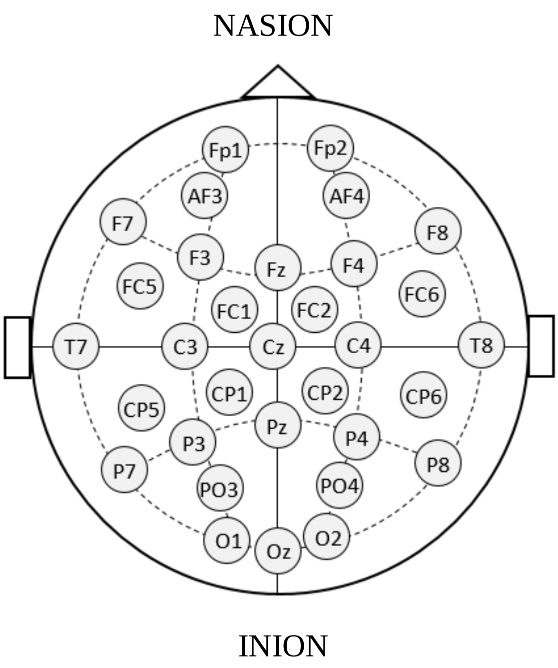


Figure 3.4: 32 EEG electrodes used by 'DEAP'.

The electrodes are placed according to the rule of the international 10/20 system which is shown in Fig. 3.4.

Table 3.3: The number and name of electrodes used by 'DEAP'

No. of Channel	Name of Channel	No. of Channel	Name of Channel
1	Fp1	17	Fp2
2	AF3	18	AF4
3	F3	19	Fz
4	F7	20	F4
5	FC5	21	F8
6	FC1	22	FC6
7	C3	23	FC2
8	T7	24	Cz
9	CP5	25	C4
10	CP1	26	T8
11	P3	27	CP6
12	P7	28	CP2
13	PO3	29	P4
14	O1	30	P8
15	Oz	31	PO4
16	Pz	32	O2

After watching the video all participant/subject are inquired for video rating in terms of valence, arousal, dominance, familiarity, and liking ranges between 1 to 9 (only linking ranges from 1 to 5). After mapping the subjects ratting four different emotions are classified according to the following rule using Russell's circumplex model.

The name of 32 channels or electrode associated with 'DEAP' are Fp1, F3, FC1, AF3, F7, FC5, T7, C3, CP1, CP5, O1, Oz, O2, P7, P3, Pz, PO3, PO4, P4, P8, CP6, CP2, C4, T8, F4, F8, AF4, Fp2, FC6, FC2, Cz and Fz. The number and name of the channels are shown in the following Table 3.3. The available data is 3d in the format like video×channel×data = 40×40×8064. It indicates for every video and every channel or electrode the number of the data point is 8064.

The full files the made available conditionally for researchers contains seven different important files. The name and the contents of the files are given in the following table for easy illustration in Table 3.4.

Table 3.4: The contents of the file available on the website of 'DEAP'

Filename	Format	Part	Contents
Online-ratings	xls, csv, ods spreadsheet	Online self-assessment	All individual ratings from the online self-assessment.
Video-list	xls, csv, ods spreadsheet	Both parts	Names/YouTube links of the music videos used in the online self-assessment and the experiment + stats of the individual ratings from the online self-assessment.
Participant-ratings	xls, csv, ods spreadsheet	Experiment	All rating participants gave to the videos during the experiment.
Participant-questionnaire	xls, csv, ods spreadsheet	Experiment	The answers participants gave to the questionnaire before the experiment.
Face_video	Zip file	Experiment	The frontal face video recordings from the experiment for participants 1-22.
Data_original	Zip file	Experiment	The original unprocessed physiological data recordings from the experiment in BioSemi .bdf format
Data-preprocessed	Zip file for Python and Matlab	Experiment	The preprocessed (downsampling, EOG removal, filtering, segmenting, etc.) physiological data recordings from the experiment in Matlab and Python(numpy) formats

3.4 Pre-processing

Our analyzed data is 'DEAP' dataset. 'DEAP' records the raw EEG data at 512 Hz. The EEG signal is very low amplitude signal. However, the extraction of information from this type of signal to too complex. Moreover, EOG and another noise signal may hamper the effectiveness of the original signal if it is not pre-processed. Firstly, the raw data were downsampled to 128 Hz from 512 Hz. Without this down sampling, the size of the EEG data had become extremely large that was not informative actually. Afterward, the EOG artifacts were removed. The signal was passed through a bandpass frequency filter of 4 to 45 Hz subsequently to avoid the un-necessary higher frequency signal. Eventually, the data were averaged to the common reference. This preprocessed data was downloaded from 'DEAP' website [63]. For further processing, the pre-processed data were used.

3.5 Reshaping

In the emotion recognition task, the downloaded data shape format was like that video×channel×data = 40×40×8064. It illustrates there remain 40 different videos. The length of each video is 60s and every video is reserved for a single emotion. Then 40 channels mean 40 electrodes data were available in this data set. Before reshaping the full data, the dimension is like the first part of Fig. 3.5.

Among them, only the first 32 electrodes are for EEG signal. Therefore, we only extract these 32 channels data. Since the length of a video is 60s and the sampling frequency is

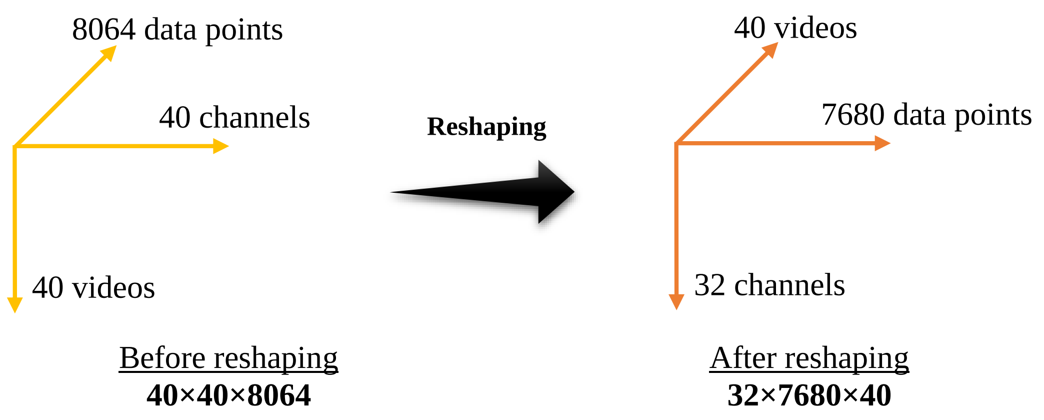


Figure 3.5: Change of dimension during reshaping.

128Hz, the number of data point must have to be only 7680. But here in the data, there remains 3s pre-trial baseline data which made an additional 384 data point in first. Thus, make the total data point 8064. So, at this point, the first 384 pre-trial baseline data were removed for our further processing as it doesn't contain any information relating to emotion. Up to this processing, the number video becomes 40, the number of developed into 32 and the number of data point reduced to 7680.

In Fig. 3.6, the pictorial view of unshaped and reshaped data was shown. The data reshaping was a very convenient arrangement to extract significant EEG data for every subject.

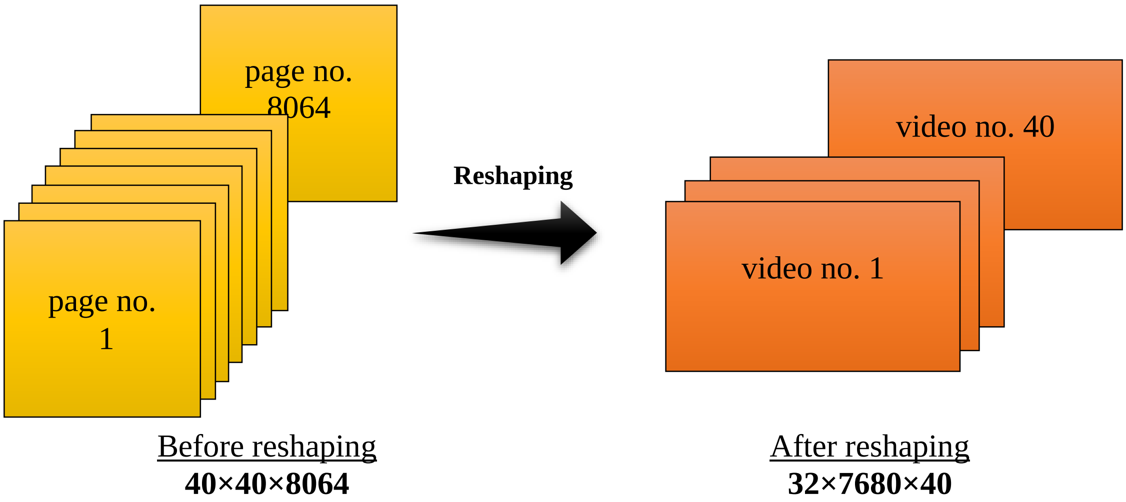


Figure 3.6: Pictorial view of reshaped data.

Using the '*permute ( )*' function on MATLAB R2017a the previous data was reshaped and only the corresponding data were extracted for supplementary processing of emotion recognition.

3.6 Decomposition

The last format of the reshaped data in the previous section was channel×data×video is equal to 32×7680×40. Here in this part of our work the EEG signals were decomposed into the 5 sub-bands namely delta *(1 Hz < f < 4 Hz)*, theta *(4 Hz < f < 8 Hz)*, alpha *(8 Hz < f < 12 Hz)*, beta *(12 Hz < f < 30 Hz)* and gamma *(30 Hz < f < 60 Hz)*. The brain activity of these sub-bands is described details in section *1.2.2.2*. Since alpha, beta and gamma subbands are more relevant with emotion, we considered only these three sub-bands to formulate a virtual image. The original EEG signal and decomposed alpha, beta and gamma sub-band signals are shown in Fig. 3.7-3.12.

For easy illustration and understanding for a single participant there shown two different emotional videos. For instance, Fig. 3.7 and Fig. 3.8 represent the EEG data for the same participant whose number is 5, But signals of Fig. 3.7 are for emotional video of number 1 whereas signals of Fig. 3.8 are for video number 27. Similarly, Fig. 3.9 and Fig. 3.10 show the signals of participant number 17. Later the signals of participant number 32 are given in Fig.

3.11 and Fig. 3.12. These figures are just for illustration of the difference of main EEG band with changing participant and also for the representation of sub-bands data of corresponding composite signal.

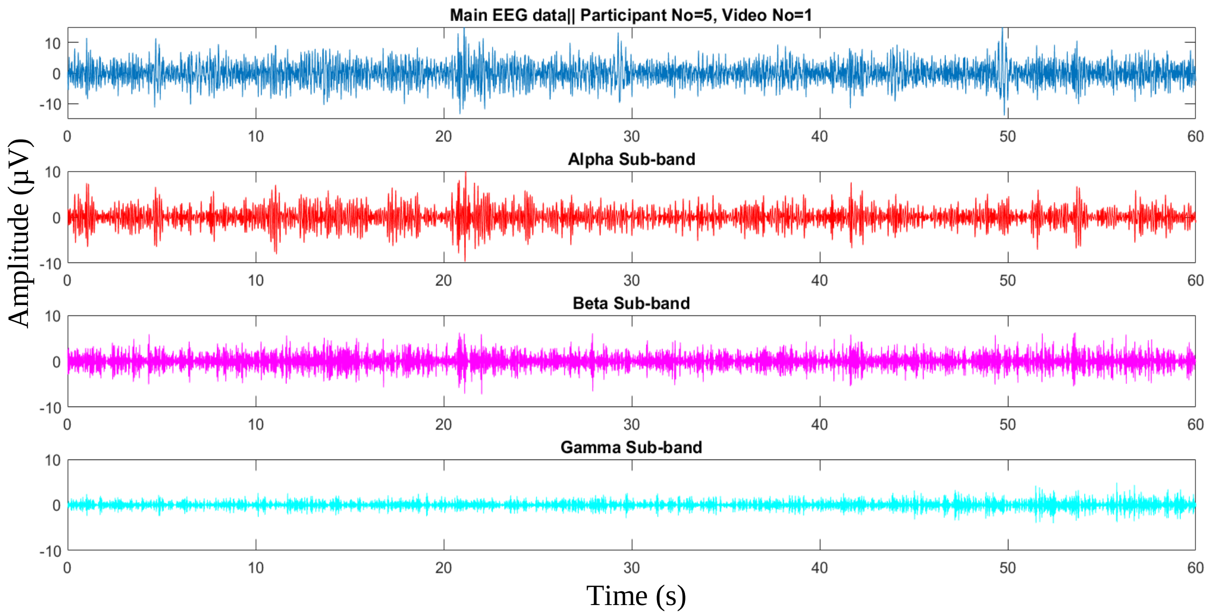


Figure 3.7: Main and Sub-bands EEG signals of participant 5 for emotional video no 1.

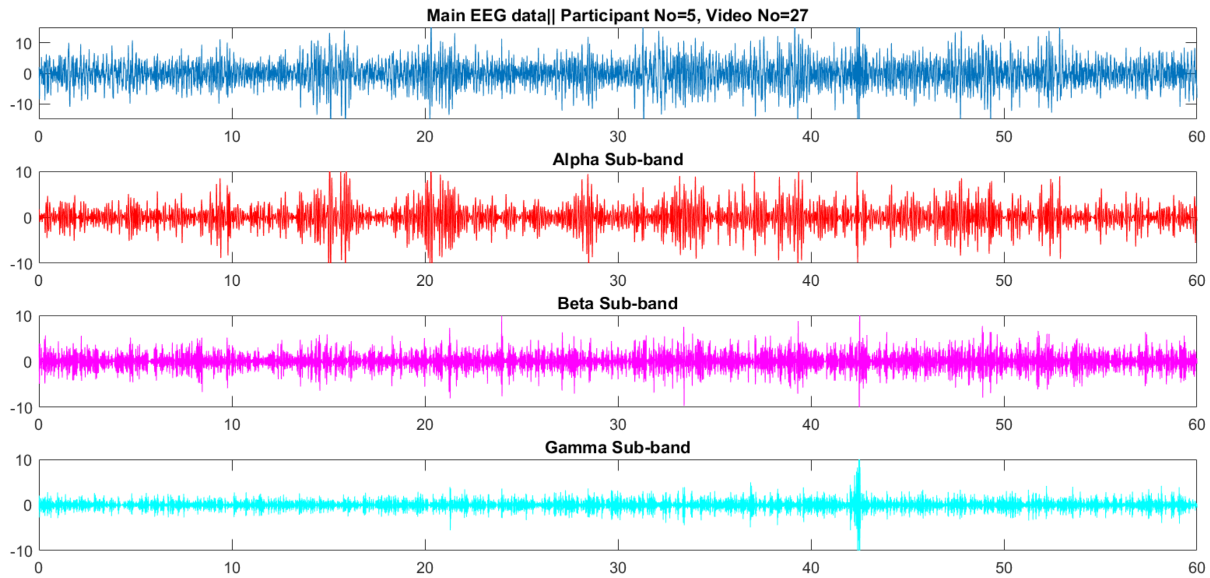


Figure 3.8: Main and Sub-bands EEG signals of participant 5 for emotional video no 27.

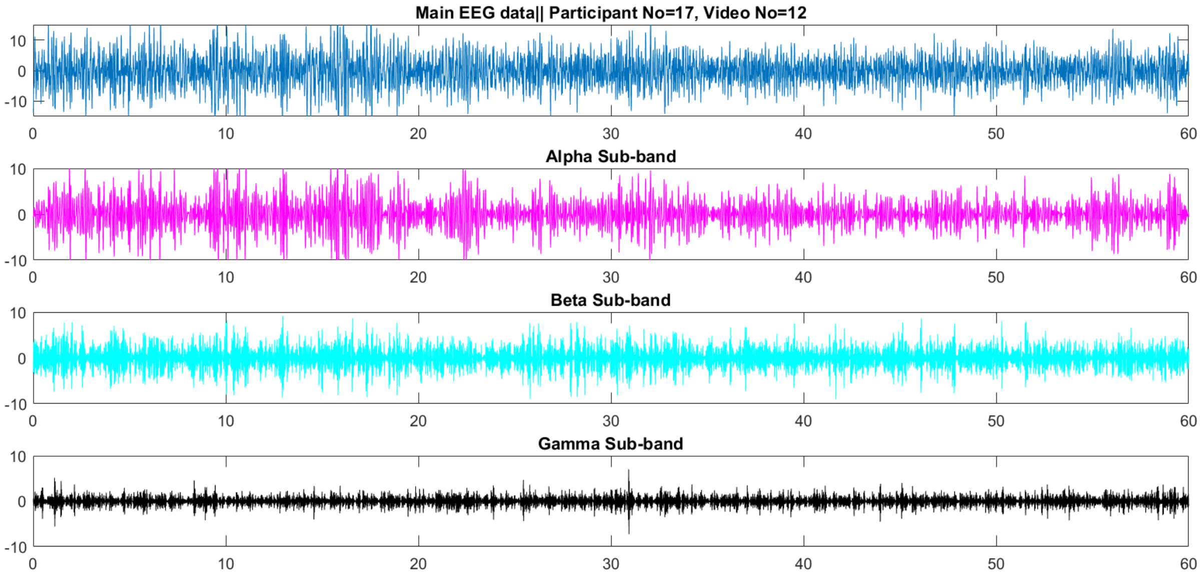


Figure 3.9: Main and Sub-bands EEG signals of participant 17 for emotional video no 12.

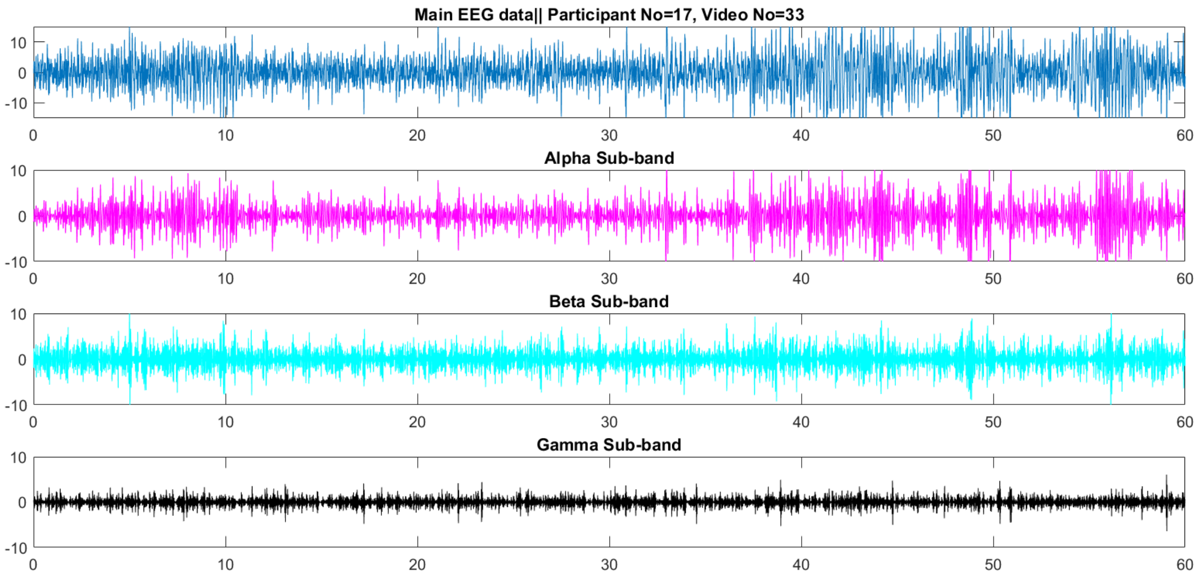


Figure 3.10: Main and Sub-bands EEG signals of participant 17 for emotional video no 33.

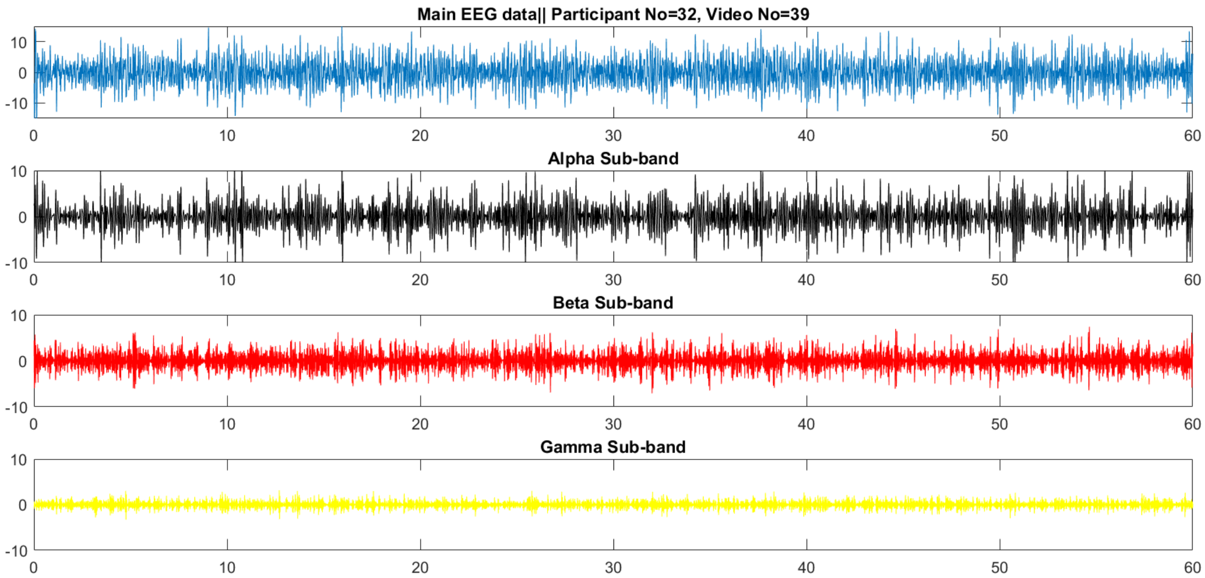


Figure 3.11: Main and Sub-bands EEG signals of participant 32 for emotional video no 39.

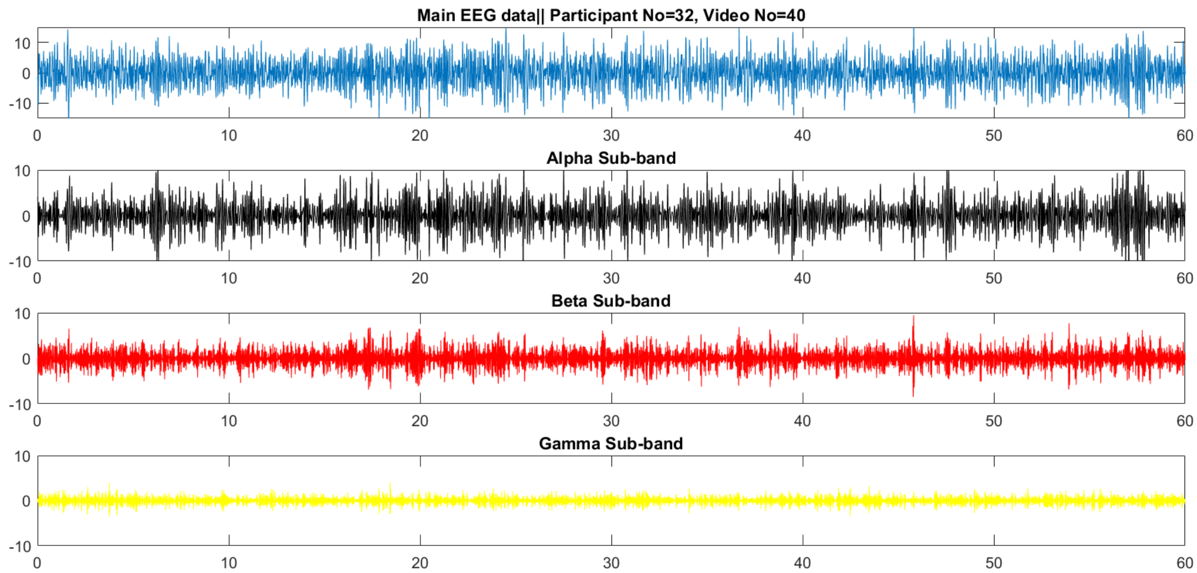


Figure 3.12: Main and Sub-bands EEG signals of participant 32 for emotional video no 40.

3.7 Formation of Virtual Images

The final target is to formulate virtual images which contain significant information for the recognition of emotion. In this section firstly the data segmentation and later the calculation of Pearson's Correlation Coefficient are described in the two different sub-sections.

3.7.1 Data Segmentation

As far we know the length of one EEG signal is the 60s. Therefore, the number of total data point is 7680. Before calculation the PCC of different channels we need to be segmented the data. The importance of this segmentation is that without the segmentation the number of training data remains limited. As CNN required a lot of training data for an exclusive performance, the EEG data have to be segmented. Here the EEG signal is segmented into 20 parts that mean every segment contains data points of 3s. The segmented part of a single channel EEG is shown in the following Fig. 3.13. So, for a single participant, the number of the total segmented partition will be 20 times 32 times 40 i.e. 25,600.

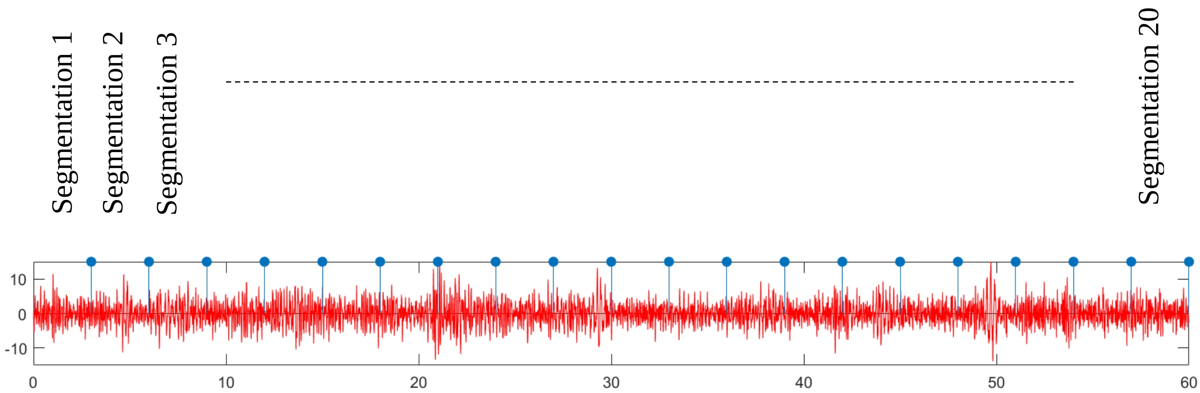


Figure 3.13: Every segment content the data of 3s; thus, a total of 20 segments in the 60s.

3.7.2 Pearson's Correlation Coefficients (PCC)

For emotion recognition, many scholars used numerous features like Power Spectral Density (PSD), Differential Entropy (DE), Differential Asymmetry (DASM), Rational Asymmetry (RASM), Wavelet Energy (WE), Relative Wavelet Energy (RWE), etc. We used Pearson's Correlation Coefficient (PCC) of a similar and consecutive point to build a matrix of cross-correlation.

Pearson's Correlation Coefficient or bivariate correlation is a measure of linear correlation between two one-dimensional variables. Mathematically, it is the ratio of covariance of two arrays to the product of their standard deviation. Its value ranges from -1 to +1. Where, +1 indicates 100% positive correlation, 0 indicates no correlation, -1 means 100% correlation but anti or negative correlation.

Suppose that, two different one-dimensional variables are x and y, where $x=(x_1, x_2, ..., x_n)$ and $y=(y_1, y_2, ..., y_n)$. Then the correlation between these two data sequences will be

$$\rho_{xy} = \frac{\mathrm{cov}(x, y)}{\sigma_x \sigma_y} \tag{3.1}$$

Where, ρ_{xy} and $cov(x,y)$ represents the PCC and covariance respectively between x and y, σ indicates the standard deviation. However, In details, the equation can be written as follows.

$$\rho_{xy} = \frac{n\sum_{i=1}^{n}(x_i y_i) - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{n\sum_{i=1}^{n} x_i^{\,2} - \left(\sum_{i=1}^{n} x_i\right)^2}\sqrt{n\sum_{i=1}^{n} y_i^{\,2} - \left(\sum_{i=1}^{n} x_i\right)^2}} \tag{3.2}$$

Here we calculated the correlation between channel to channel for every part of segmentation. The proper explanation of the calculation of correlation is described in the next section.

3.7.3 No. of Participant, Video, Segmentation and Virtual Images

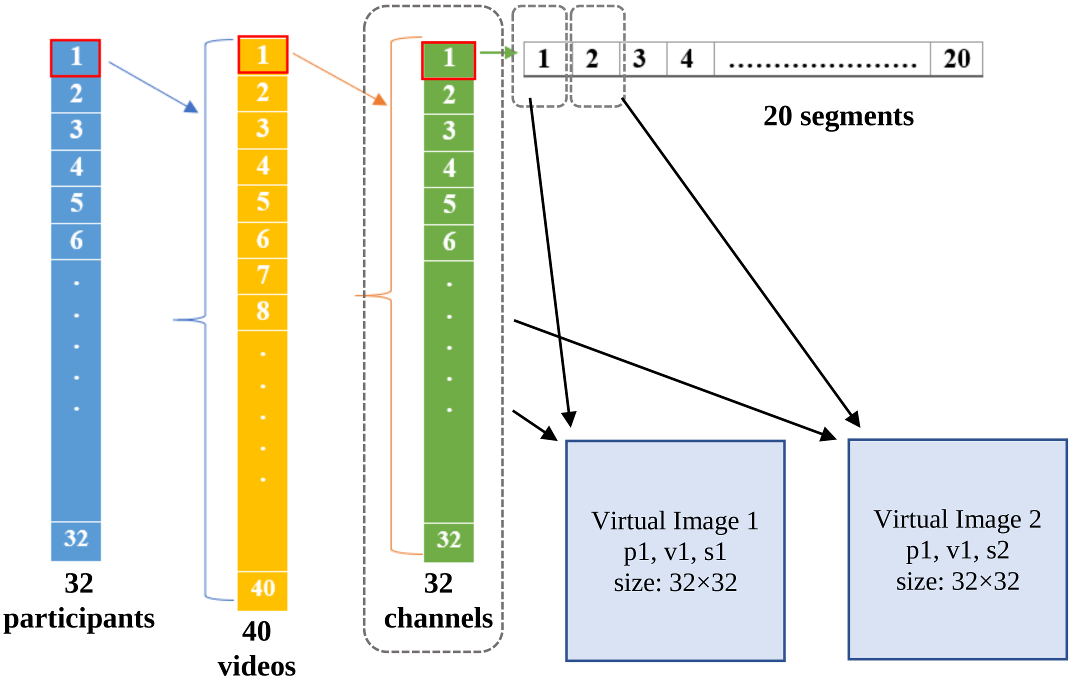


Figure 3.14: Numbering of virtual images.

The virtual image number 1 is formulated from the data of participant no1, video number 1, segment number 1 by using PCC of 32 channel. Very similarly virtual image number 2 is for p1, v1, s2 which indicates participant 2, video 1 and segmentation 2. Thus, as a whole, for a single video total, 20 virtual images can be found. And for 40 videos total of 800 virtual images can be generated. So, for 32 participants the number of total virtual images will be 32×800 = 2,5600.

3.7.4 Virtual Images for Emotion Recognition

Our main target is to design a deep architecture-based emotion recognition method. This is why we used the Convolutional Neural Network on virtual Images. The virtual images contain the Pearson's Correlation Coefficient features of two consecutive channels dataset. To formulate the virtual images firstly the PCC between the combination of every two channels are calculated. Afterward, if we arrange the coefficients in a matrix form, then an array of data can be found. In our work as we considered 32 EEG channels data, the size of the PCC matrix will be 32×32 for any single segmentation.

For example, the PCC matrix of alpha sub-band of the 1st participant for the 1st video and for the 1st segmentation is shown in Fig. 3.15. It may be noted that, since the correlation between two same series of data is 1, the diagonal data of the matrix indicated so.

Columns 1 through 16

1.0000	0.8813	0.6558	0.7971	0.4374	0.2144	-0.1724	0.1791	-0.4891	-0.7263	-0.7468	-0.4167	-0.7365	-0.6301	-0.6211	-0.8501
0.8813	1.0000	0.9021	0.7493	0.4660	0.5082	0.0805	0.2226	-0.3487	-0.5326	-0.6802	-0.4470	-0.6545	-0.6291	-0.7205	-0.7394
0.6558	0.9021	1.0000	0.6632	0.5123	0.6569	0.3522	0.3657	-0.0818	-0.2631	-0.4617	-0.2714	-0.4210	-0.4311	-0.6435	-0.5590
0.7971	0.7493	0.6632	1.0000	0.7331	0.2041	0.1673	0.5685	-0.1143	-0.5413	-0.5032	-0.0785	-0.4745	-0.3685	-0.5341	-0.8159
0.4374	0.4660	0.5123	0.7331	1.0000	0.4822	0.4724	0.6614	0.2841	-0.2286	0.0016	0.3367	-0.2025	-0.1863	-0.4449	-0.6320
0.2144	0.5082	0.6569	0.2041	0.4822	1.0000	0.4618	0.1454	0.0623	0.0952	-0.0901	-0.1275	-0.2710	-0.3764	-0.5582	-0.2509
-0.1724	0.0805	0.3522	0.1673	0.4724	0.4618	1.0000	0.6476	0.7711	0.5759	0.4109	0.4574	0.3378	0.1610	-0.2227	0.0806
0.1791	0.2226	0.3657	0.5685	0.6614	0.1454	0.6476	1.0000	0.5720	0.0901	0.1300	0.5090	0.1367	0.1125	-0.2219	-0.3212
-0.4891	-0.3487	-0.0818	-0.1143	0.2841	0.0623	0.7711	0.5720	1.0000	0.6949	0.8137	0.8239	0.7300	0.5610	0.1906	0.3246
-0.7263	-0.5326	-0.2631	-0.5413	-0.2286	0.0952	0.5759	0.0901	0.6949	1.0000	0.7353	0.4301	0.6710	0.4461	0.2550	0.7662
-0.7468	-0.6802	-0.4617	-0.5032	0.0016	-0.0901	0.4109	0.1300	0.8137	0.7353	1.0000	0.7897	0.8539	0.6872	0.4748	0.5977
-0.4167	-0.4470	-0.2714	-0.0785	0.3367	-0.1275	0.4574	0.5090	0.8239	0.4301	0.7897	1.0000	0.6740	0.6441	0.3209	0.1632
-0.7365	-0.6545	-0.4210	-0.4745	-0.2025	-0.2710	0.3378	0.1367	0.7300	0.6710	0.8539	0.6740	1.0000	0.9003	0.7086	0.6879
-0.6301	-0.6291	-0.4311	-0.3685	-0.1863	-0.3764	0.1610	0.1125	0.5610	0.4461	0.6872	0.6441	0.9003	1.0000	0.8540	0.5629
-0.6211	-0.7205	-0.6435	-0.5341	-0.4449	-0.5582	-0.2227	-0.2219	0.1906	0.2550	0.4748	0.3209	0.7086	0.8540	1.0000	0.6435
-0.8501	-0.7394	-0.5590	-0.8159	-0.6320	-0.2509	0.0806	-0.3212	0.3246	0.7662	0.5977	0.1632	0.6879	0.5629	0.6435	1.0000
0.8757	0.7714	0.5036	0.6255	0.1202	0.0281	-0.3077	0.0245	-0.5936	-0.7454	-0.8415	-0.5784	-0.7476	-0.6673	-0.5855	-0.7513
0.7235	0.6251	0.3413	0.3941	0.1358	0.2404	-0.4468	-0.2428	-0.6946	-0.7124	-0.7286	-0.5791	-0.8180	-0.7650	-0.6243	-0.7113
0.5684	0.7997	0.7821	0.3744	0.1680	0.6911	0.0848	-0.0247	-0.3718	-0.2748	-0.6274	-0.5685	-0.5998	-0.6442	-0.6915	-0.4571
0.6429	0.6331	0.4404	0.3574	-0.0754	0.2224	-0.2701	-0.1442	-0.6007	-0.5197	-0.7845	-0.6887	-0.7306	-0.7159	-0.6173	-0.5502
0.6679	0.4505	0.1521	0.4230	-0.0223	-0.1634	-0.5082	-0.1728	-0.6868	-0.7466	-0.7669	-0.5655	-0.7444	-0.6308	-0.4529	-0.6610
0.4930	0.2713	-0.0462	0.2059	-0.0929	-0.1099	-0.5735	-0.3509	-0.7366	-0.6904	-0.6726	-0.5555	-0.7718	-0.6876	-0.4432	-0.5635
0.3226	0.4836	0.4385	0.0907	-0.1185	0.5472	0.0280	-0.1482	-0.4127	-0.1034	-0.5408	-0.6110	-0.5084	-0.6033	-0.5816	-0.1908
-0.3488	-0.0988	0.0515	-0.3920	-0.1549	0.5443	0.3632	-0.1333	0.1853	0.5921	0.2063	-0.0559	0.0819	-0.0622	-0.1468	0.3945
-0.0856	-0.1346	-0.2821	-0.3058	-0.2869	0.0326	-0.3805	-0.5508	-0.5058	-0.1082	-0.2484	-0.4417	-0.4710	-0.5452	-0.2974	0.0552
0.3016	0.0185	-0.2735	0.0982	-0.0402	-0.3439	-0.6872	-0.3467	-0.6320	-0.6555	-0.4174	-0.3287	-0.5655	-0.4408	-0.1433	-0.4173
-0.2562	-0.3420	-0.4547	-0.4429	-0.5504	-0.3974	-0.5611	-0.6204	-0.4791	-0.1139	-0.1695	-0.3931	-0.2176	-0.1956	0.1332	0.2674
-0.5228	-0.4185	-0.3512	-0.6749	-0.6427	0.0076	-0.0646	-0.4638	-0.0993	0.5415	0.1076	-0.2940	0.0878	-0.0404	0.1267	0.7181
-0.6026	-0.6995	-0.7233	-0.7373	-0.5528	-0.3833	-0.4909	-0.6566	-0.2103	0.2018	0.2843	-0.0757	0.2024	0.2020	0.5153	0.6301
-0.1054	-0.3313	-0.5015	-0.2979	-0.5236	-0.6772	-0.7104	-0.4987	-0.4969	-0.3431	-0.2247	-0.2979	-0.1319	0.0309	0.4259	0.1294
-0.6148	-0.7358	-0.7379	-0.7154	-0.6144	-0.5495	-0.4775	-0.5723	-0.1260	0.1919	0.3219	-0.0054	0.3913	0.4566	0.7800	0.7029
-0.5082	-0.6667	-0.6648	-0.5206	-0.5207	-0.6588	-0.4562	-0.3769	-0.0867	0.0651	0.2468	0.0798	0.4268	0.5933	0.8824	0.5718

Columns 17 through 32

0.8757	0.7235	0.5684	0.6429	0.6679	0.4930	0.3226	-0.3488	-0.0856	0.3016	-0.2562	-0.5228	-0.6026	-0.1054	-0.6148	-0.5082
0.7714	0.6251	0.7997	0.6331	0.4505	0.2713	0.4836	-0.0988	-0.1346	0.0185	-0.3420	-0.4185	-0.6995	-0.3313	-0.7358	-0.6667
0.5036	0.3413	0.7821	0.4404	0.1521	-0.0462	0.4385	0.0515	-0.2821	-0.2735	-0.4547	-0.3512	-0.7233	-0.5015	-0.7379	-0.6648
0.6255	0.3941	0.3744	0.3574	0.4230	0.2059	0.0907	-0.3920	-0.3058	0.0982	-0.4429	-0.6749	-0.7373	-0.2979	-0.7154	-0.5206
0.1202	0.1358	0.1680	-0.0754	-0.0223	-0.0929	-0.1185	-0.1549	-0.2869	-0.0402	-0.5504	-0.6427	-0.5528	-0.5236	-0.6144	-0.5207
0.0281	0.2404	0.6911	0.2224	-0.1634	-0.1099	0.5472	0.5443	0.0326	-0.3439	-0.3974	0.0076	-0.3833	-0.6772	-0.5495	-0.6588
-0.3077	-0.4468	0.0848	-0.2701	-0.5082	-0.5735	0.0280	0.3632	-0.3805	-0.6872	-0.5611	-0.0646	-0.4909	-0.7104	-0.4775	-0.4562
0.0245	-0.2428	-0.0247	-0.1442	-0.1728	-0.3509	-0.1482	-0.1333	-0.5508	-0.3467	-0.6204	-0.4638	-0.6566	-0.4987	-0.5723	-0.3769
-0.5936	-0.6946	-0.3718	-0.6007	-0.6868	-0.7366	-0.4127	0.1853	-0.5058	-0.6320	-0.4791	-0.0993	-0.2103	-0.4969	-0.1260	-0.0867
-0.7454	-0.7124	-0.2748	-0.5197	-0.7466	-0.6904	-0.1034	0.5921	-0.1082	-0.6555	-0.1139	0.5415	0.2018	-0.3431	0.1919	0.0651
-0.8415	-0.7286	-0.6274	-0.7845	-0.7669	-0.6726	-0.5408	0.2063	-0.2484	-0.4174	-0.1695	0.1076	0.2843	-0.2247	0.3219	0.2468
-0.5784	-0.5791	-0.5685	-0.6887	-0.5655	-0.5555	-0.6110	-0.0559	-0.4417	-0.3287	-0.3931	-0.2940	-0.0757	-0.2979	-0.0054	0.0798
-0.7476	-0.8180	-0.5998	-0.7306	-0.7444	-0.7718	-0.5084	0.0819	-0.4710	-0.5655	-0.2176	0.0878	0.2024	-0.1319	0.3913	0.4268
-0.6673	-0.7650	-0.6442	-0.7159	-0.6308	-0.6876	-0.6033	-0.0622	-0.5452	-0.4408	-0.1956	-0.0404	0.2020	0.0309	0.4566	0.5933
-0.5855	-0.6243	-0.6915	-0.6173	-0.4529	-0.4432	-0.5816	-0.1468	-0.2974	-0.1433	0.1332	0.1267	0.5153	0.4259	0.7800	0.8824
-0.7513	-0.7113	-0.4571	-0.5502	-0.6610	-0.5635	-0.1908	0.3945	0.0552	-0.4173	0.2674	0.7181	0.6301	0.1294	0.7029	0.5718
1.0000	0.8157	0.5833	0.8366	0.8349	0.6702	0.4421	-0.3563	0.0311	0.3574	-0.0999	-0.3745	-0.5311	0.0417	-0.5391	-0.4462
0.8157	1.0000	0.6086	0.8245	0.8072	0.8139	0.5113	-0.1623	0.3086	0.5001	-0.0088	-0.2691	-0.2733	0.0223	-0.4205	-0.4507
0.5833	0.6086	1.0000	0.7035	0.3521	0.2762	0.7831	0.3265	0.0817	-0.1341	-0.2283	-0.0206	-0.5307	-0.4032	-0.6404	-0.6805
0.8366	0.8245	0.7035	1.0000	0.7494	0.7064	0.7324	-0.0396	0.1959	0.2206	-0.0427	-0.1078	-0.4369	-0.0269	-0.4949	-0.4924
0.8349	0.8072	0.3521	0.7494	1.0000	0.8866	0.2936	-0.4033	0.2712	0.6407	0.1334	-0.2745	-0.2421	0.2523	-0.3103	-0.2540
0.6702	0.8139	0.2762	0.7064	0.8866	1.0000	0.3739	-0.2606	0.5682	0.7652	0.3588	-0.1233	-0.0109	0.3344	-0.1694	-0.2149
0.4421	0.5113	0.7831	0.7324	0.2936	0.3739	1.0000	0.5172	0.3279	-0.1314	-0.0469	0.2820	-0.3206	-0.3221	-0.4484	-0.5588
-0.3563	-0.1623	0.3265	-0.0396	-0.4033	-0.2606	0.5172	1.0000	0.2226	-0.4588	-0.0719	0.6369	0.0452	-0.4339	-0.1003	-0.2801
0.0311	0.3086	0.0817	0.1959	0.2712	0.5682	0.3279	0.2226	1.0000	0.5656	0.7474	0.5187	0.5271	0.2731	0.1975	-0.0401
0.3574	0.5001	-0.1341	0.2206	0.6407	0.7652	-0.1314	-0.4588	0.5656	1.0000	0.6010	-0.1640	0.3247	0.6244	0.1580	0.1179
-0.0999	-0.0088	-0.2283	-0.0427	0.1334	0.3588	-0.0469	-0.0719	0.7474	0.6010	1.0000	0.4975	0.7412	0.7333	0.5767	0.4133
-0.3745	-0.2691	-0.0206	-0.1078	-0.2745	-0.1233	0.2820	0.6369	0.5187	-0.1640	0.4975	1.0000	0.5885	0.1174	0.4553	0.2049
-0.5311	-0.2733	-0.5307	-0.4369	-0.2421	-0.0109	-0.3206	0.0452	0.5271	0.3247	0.7412	0.5885	1.0000	0.6026	0.9009	0.7048
0.0417	0.0223	-0.4032	-0.0269	0.2523	0.3344	-0.3221	-0.4339	0.2731	0.6244	0.7333	0.1174	0.6026	1.0000	0.6752	0.6827
-0.5391	-0.4205	-0.6404	-0.4949	-0.3103	-0.1694	-0.4484	-0.1003	0.1975	0.1580	0.5767	0.4553	0.9009	0.6752	1.0000	0.9124
-0.4462	-0.4507	-0.6805	-0.4924	-0.2540	-0.2149	-0.5588	-0.2801	-0.0401	0.1179	0.4133	0.2049	0.7048	0.6827	0.9124	1.0000

Figure 3.15: 32×32 size image data of alpha-band; p1, v1, s1.

Then after getting the correlation matrix of three sub-bands (alpha, beta, and gamma) we concatenated the data and got the matrix of size 32×32 like as Fig. 3.16. The first pixel of this matrix indicates the PCC of alpha sub-banded data of channel 1 with channel 1. The next pixel indicates the same type of value between channel 1 and channel 2. The rule is the same for the rest of all pixels. For a single participant, we got 2400 images containing PCC matrix.

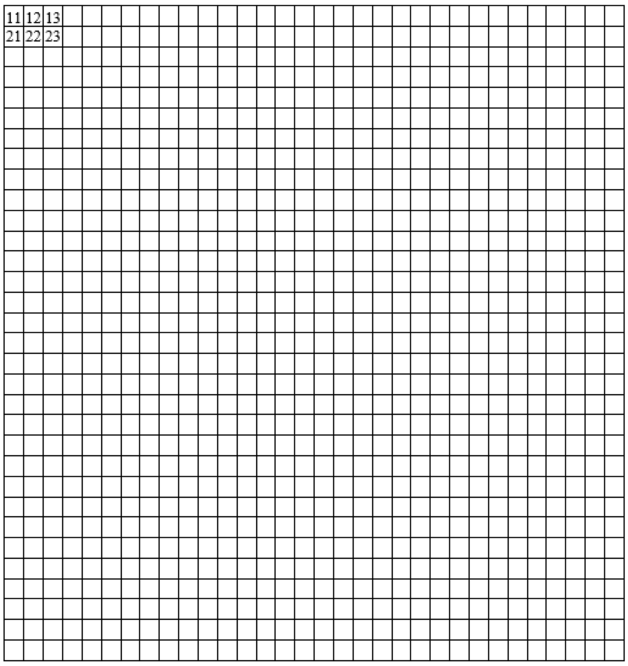


Figure 3.16: The general format of the average valued matrix of size 32×32.

Then the average valued matrix can be plot as a virtual color image using MATLAB software. Some examples of this type of virtual color images we formulated are shown in the following Fig. 3.17-3.19.

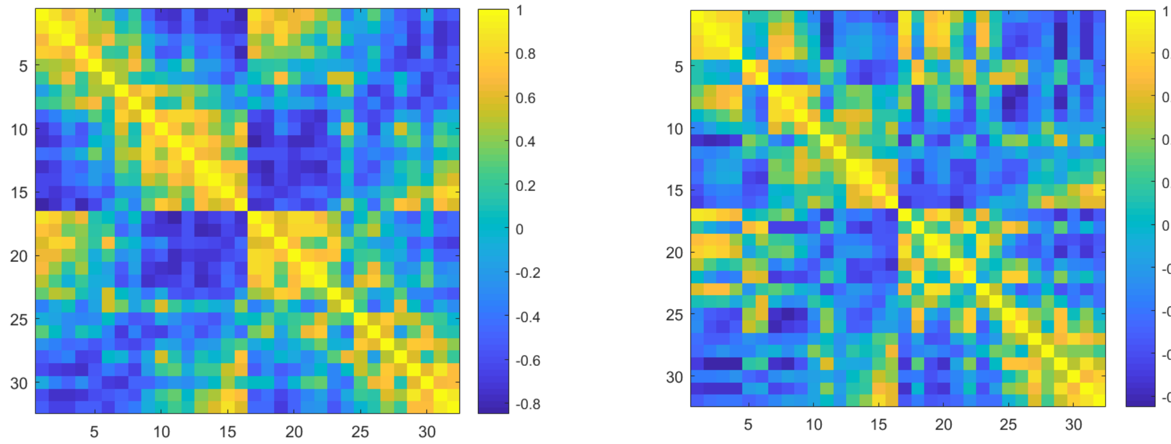


Figure 3.17: Virtual images of p1, v1; (a): segmentation 1, (b) segmentation 2.

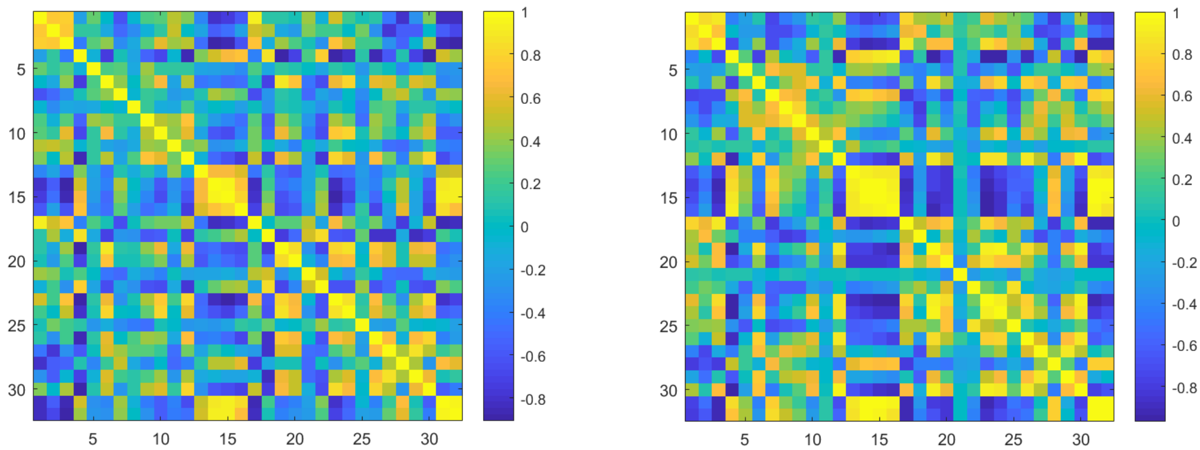


Figure 3.18: Virtual images of p17, s11; (a) video 1, (b) video 30.

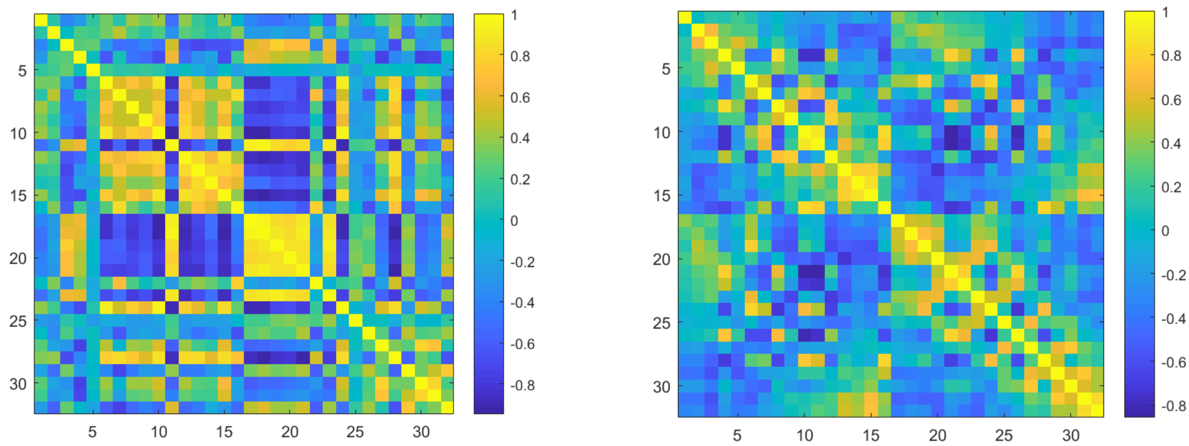


Figure 3.19: Virtual images of v20, s7; (a) participant 15, (b) participant 31.

Fig. 3.17 represents the virtual images of participant 1, video 1 and for segmentation 1 & 2. In Fig. 3.18, the participant and the segmentation remain same; here the change of virtual image according to video is expressed. Similarly, in Fig. 3.19, the number of videos and the segmentation remains fixed (video number 20, segmentation number 7) whereas the variation of the participant is illustrated. In our experiment, we need to classify emotion in two and three categories. As a result, the more the variation of the virtual image with the variation of video number, the more accuracy can be found.

However, to illustrate the variety of virtual images for any single participant, the images of a random number of video and segmentation are important. In Fig. 3.20, the random virtual images of participant number 24 are shown.

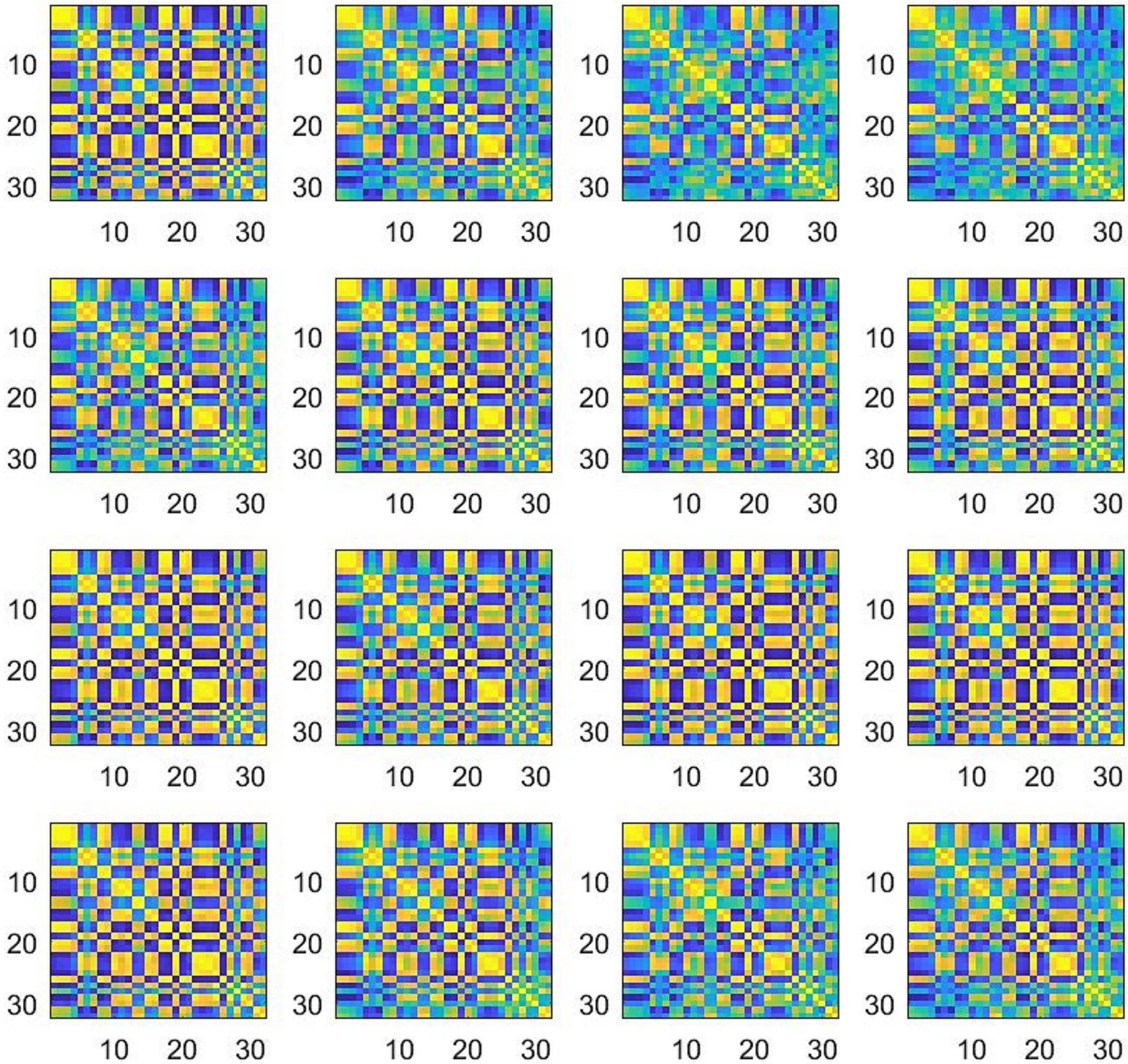


Figure 3.20: Random virtual images of participant number 24.

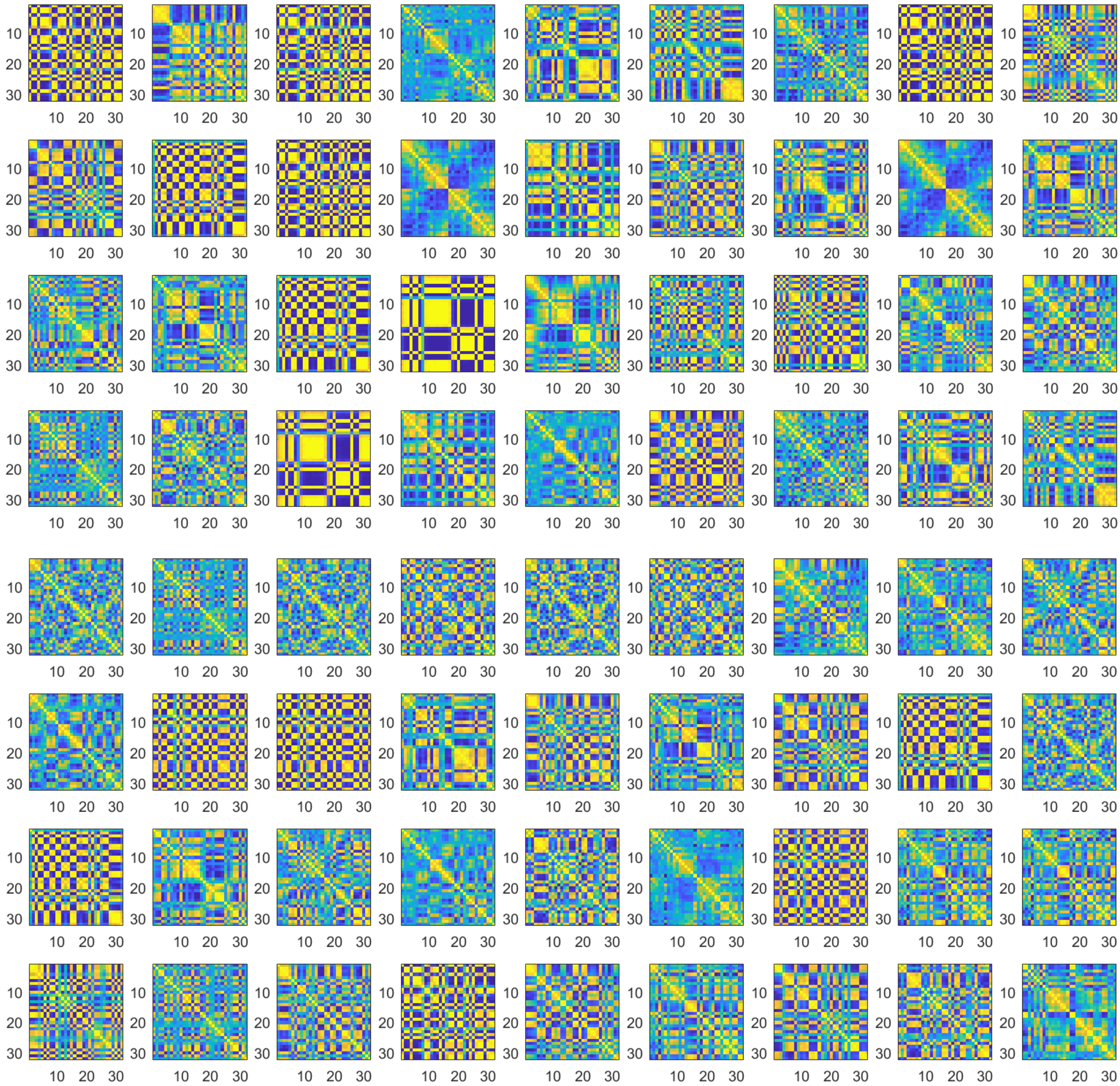


Figure 3.21: Some random virtual images [1 to 25600].

3.8 Experiment Protocol

For a subject, for a single trial video, and for a single channel there remain 8064 (63s ×128Hz) data points. As CNN required a large amount of dataset for training, to operate successfully we segmented the data to increase the quantity of virtual image. For doing this after removing the data of first 3 seconds we took only the last 60s data and segmented it into 20 samples each contains 384 data points of 3 seconds. Thus, for only one participant, we got 800 samples. Thus, we had 25,600 samples of the virtual image in total for all subject. In our work, we have made 2 different protocol.

- Protocol-1: Here, we just classified two discrete emotions positive and negative. For each participant, we considered the valence and arousal rating. From the rating, our model predicts 0 to less than 5 as low and from 5 to 9 as high. We classified as low as negative and high as a positive emotion.
- Protocol-2: Here emotion is classified into three categories positive, negative and neutral. For this classification, we considered 0 to 2.99 as low, 3 to 5.99 as medium and 6 to 9 as high. Thus, low values of valence and arousal indicate the negative emotion, the medium value indicates neutral emotion and high values indicate positive emotions.

The overall pictorial view from starting to the generation of virtual images are shown in Fig. 3.22. After generating the 25600 virtual images from single sub-bands we made a dataset of 76800 virtual images for three sub-bands. These virtual images are fed into a Convolutional Neural Network model to classify emotion into two and three classes. The description of CNN based architecture is explained in the next chapter.

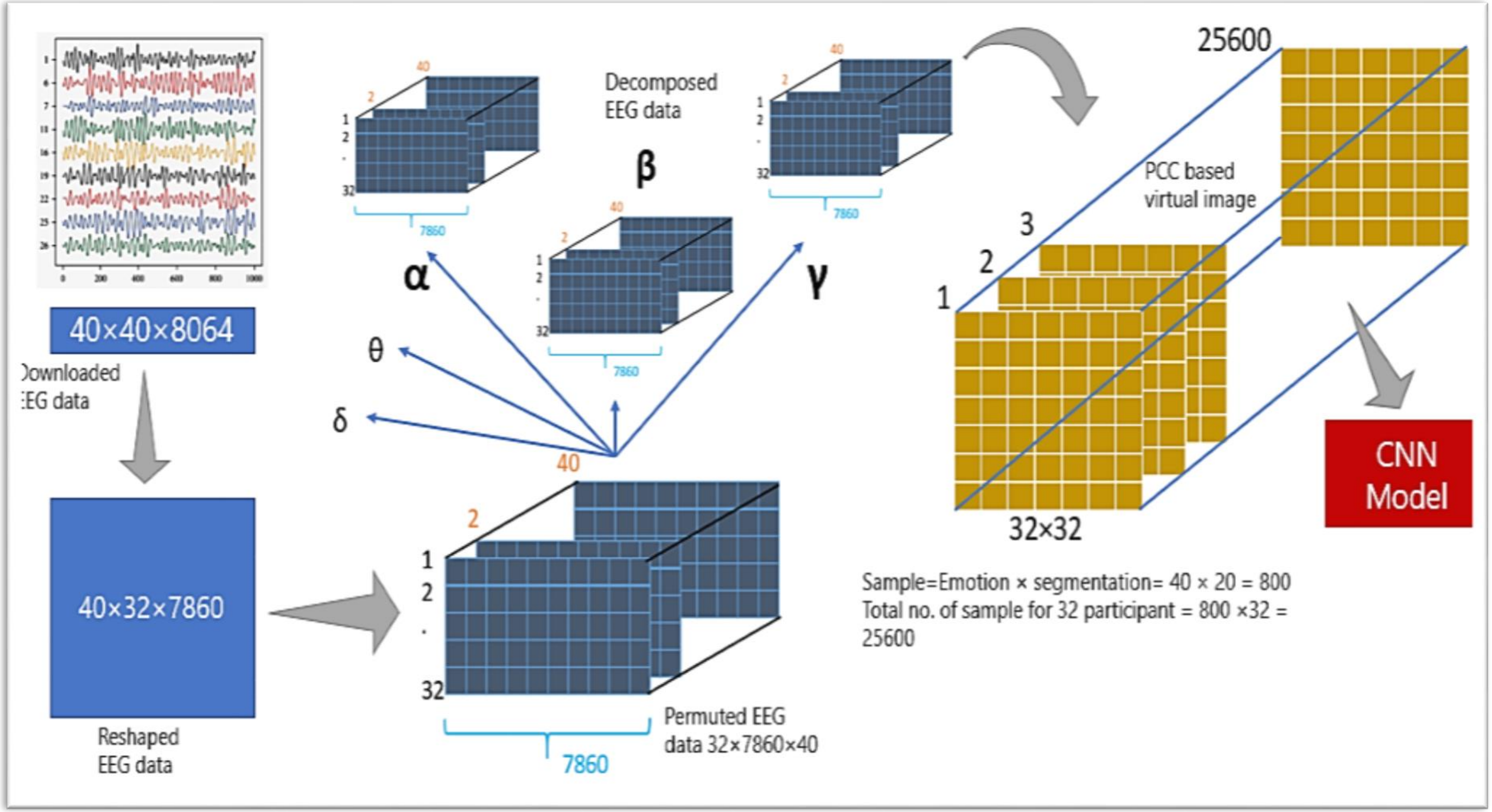


Figure 3.22: Pictorial flowchart of all processing.

Chapter **4**

CNN Architecture

Chapter Outlines:

Chapter 4

CNN Architecture

4.1 Introduction

Emotion Recognition from EEG is the fundamental goal of our work. To achieve this goal the downloaded EEG data are preprocessed, reshaped and permuted properly according to our necessity. Then the processed data are used to formulate the PCC array. From this PCC array, we generated the virtual images to train the CNN model to classify emotion accurately. This chapter is for the proper explanation of our proposed CNN model. The details information of the building blocks, different layers, optimization algorithm, loss function, etc. of our CNN model are clarified in this section of this report.

4.2 Convolutional Neural Network

Convolutional Neural Network (CNN) or simply ConvNet is a branch of Deep Neural Networks, which proven outstanding performance on image classification and computer vision. The main structure of CNN is very much similar to the connectivity pattern of the neuron of the human brain. One of the main advantages of CNN is that it requires a little bit preprocessing or sometimes needs no preprocessing like other traditional classification algorithms. It has the ability to learn a large number of features automatically from the training dataset and uses this specific feature to predict the test data. The ConvNet takes an input image, assign and upgrade the weights and biases on various point of views and able to classify one from other.

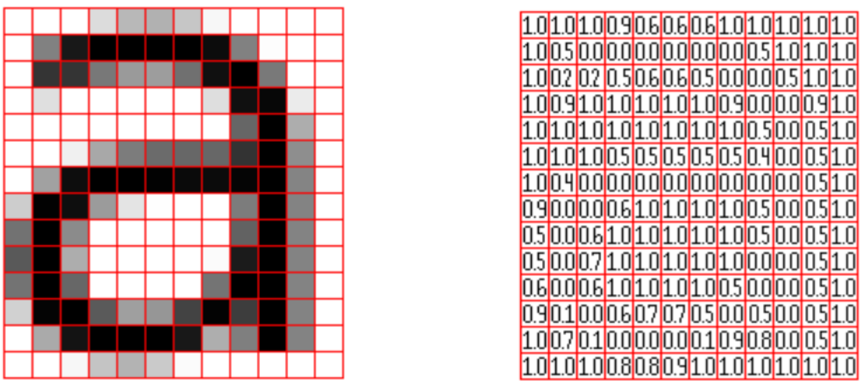


Figure 4.1: Human vision (on left) and computer (on right) vision of an image.

The human brain visualizes an image like the left part of Fig. 4.1 and recognizes easily within a moment. Whereas the computer gets only some digital data from an image like the right part of Fig. 4.1. For recognizing any object or text or any other element from the image it is difficult for a computer. This challenging task can be done easily by using CNN architecture.

Artificial Intelligence has been growing in apex position for minifying the gap of defensibility between human and machine. In the world of AI, the CNN performs state of the art success in object detection, recognition, and classification; image and video recognition, image to text conversion, handwritten detection, natural language processing, robotic vision, driving assistance for a driverless car, media recreation, etc. CNN actually works with 2D image data but capable of working with 1D and 3D data.

In the present world, the image size is increasing day by day. For instance, a single-color image of size 250×250 contains the pixels of 250×250×3=187,500.

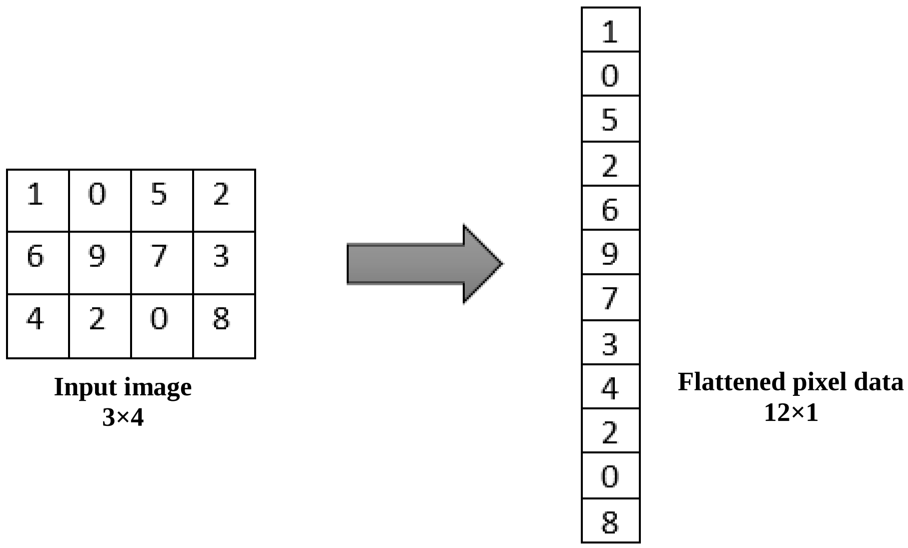


Figure 4.2: Flattening pixels in feed-forward Neural Network.

Therefore, in a traditional feed-forward neural network the number of input feature for this size input images will be 187,500; a large number. Considering a common number of hidden layers 1,024; the neural network has 187,500×1024=192,000,000 nodes only for the first layer. One hundred and ninety-two millions of nodes in just one layer for an image size of 250×250. So, the computational complexity increases from an intolerable extent. Another problem is that what will be for larger-size photo. Another important issue is that the object position may have been changed. A Convolutional Neural Network can mitigate these problems.

4.3 Working of CNN Model

Like another neural network, the CNN architecture consists of an input layer, hidden layers, and an output layer. There remain some building blocks in the hidden layer.

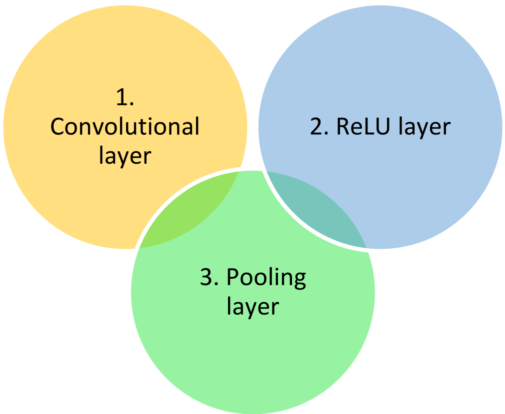


Figure 4.3: Elements of a hidden layer of a CNN.

A set of a convolutional layer, ReLU layer, and pooling layer forms the hidden layer shown in Fig. 4.3. A complete CNN also contains Fully Connected (FC) layer like flatten and dense layer. The Convolutional layer is the main layer of a Convolutional Neural Network architecture.

The first operation of a CNN is making convolution with the image and a filter. In the neural network discussion, the filter and kernel are exactly the same terms. The kernel moves through the image and made a new convolved image. The direction of movement of the kernel in the whole image is shown in the following Fig. 4.4.

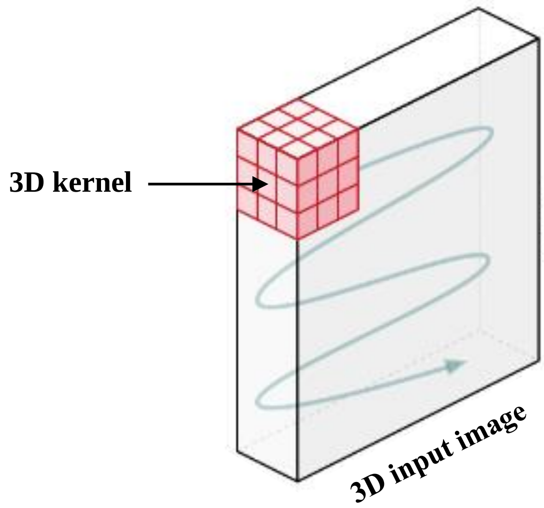


Figure 4.4: The movement of the kernel during convolution.

A color image can be treated as a 3-dimensional matrix. The red, green and blue are the 1st, 2nd and 3rd pages of the image volume. Similarly, the 3D color image will be convolved with a 3D kernel. An example of convolution operation is showing below for illustration its operation easily. Here, in Fig. 4.5, a single layer padding (i.e. p=1) is added with a 5×5 image and then a convolution operation is made with this padded image and 3×3 filter.

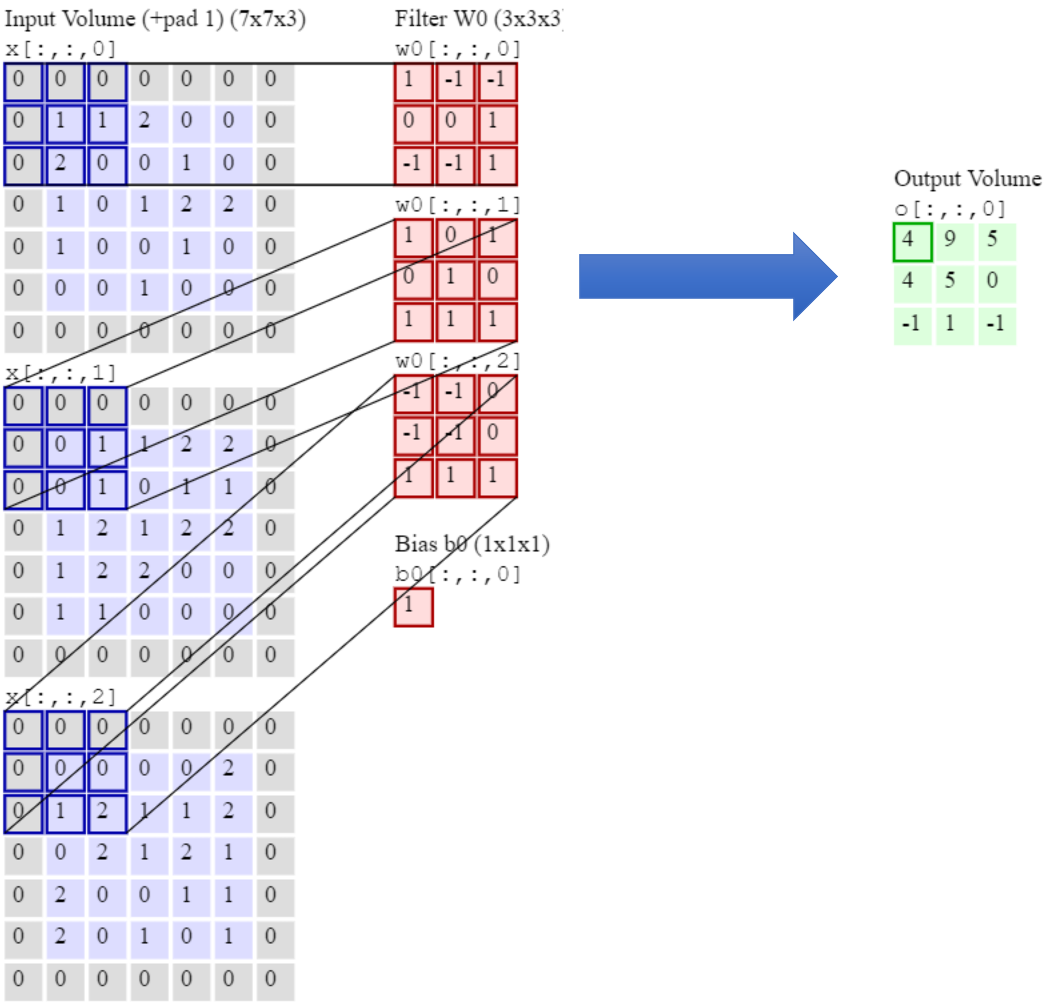


Figure 4.5: Convolution between an image and a filter (step 1). [adopted from 'skymind.ai' webpage]

Using the stride two the convolution operation will again be done shown in Fig. 4.6 and thus calculate the 2nd value of output volume. Here the 2nd value is 9.

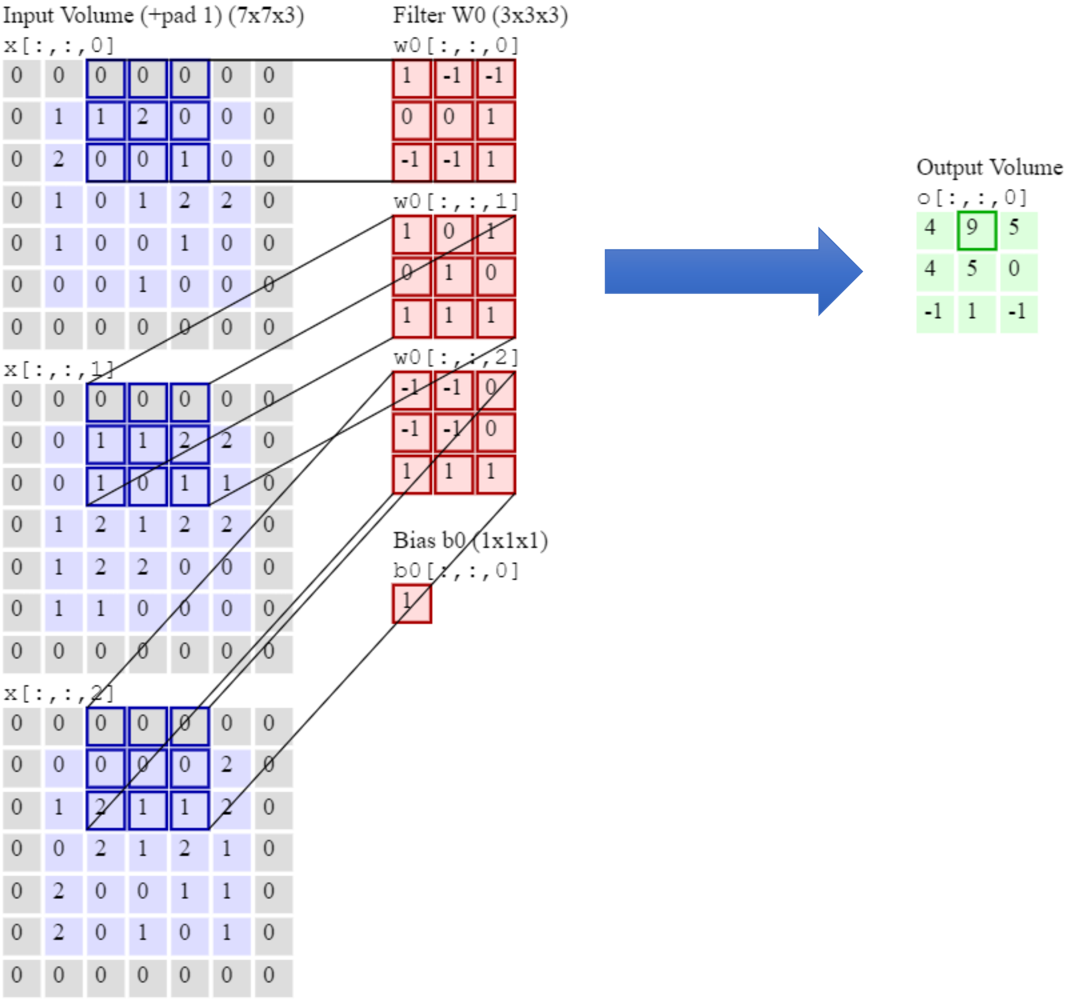


Figure 4.6: Convolution between an image and a filter (step 2). [adopted from 'skymind.ai' webpage]

Like as the previous procedure the 3rd value of the output volume can be calculated. This operation with its corresponding figure is given in Fig. 4.7.

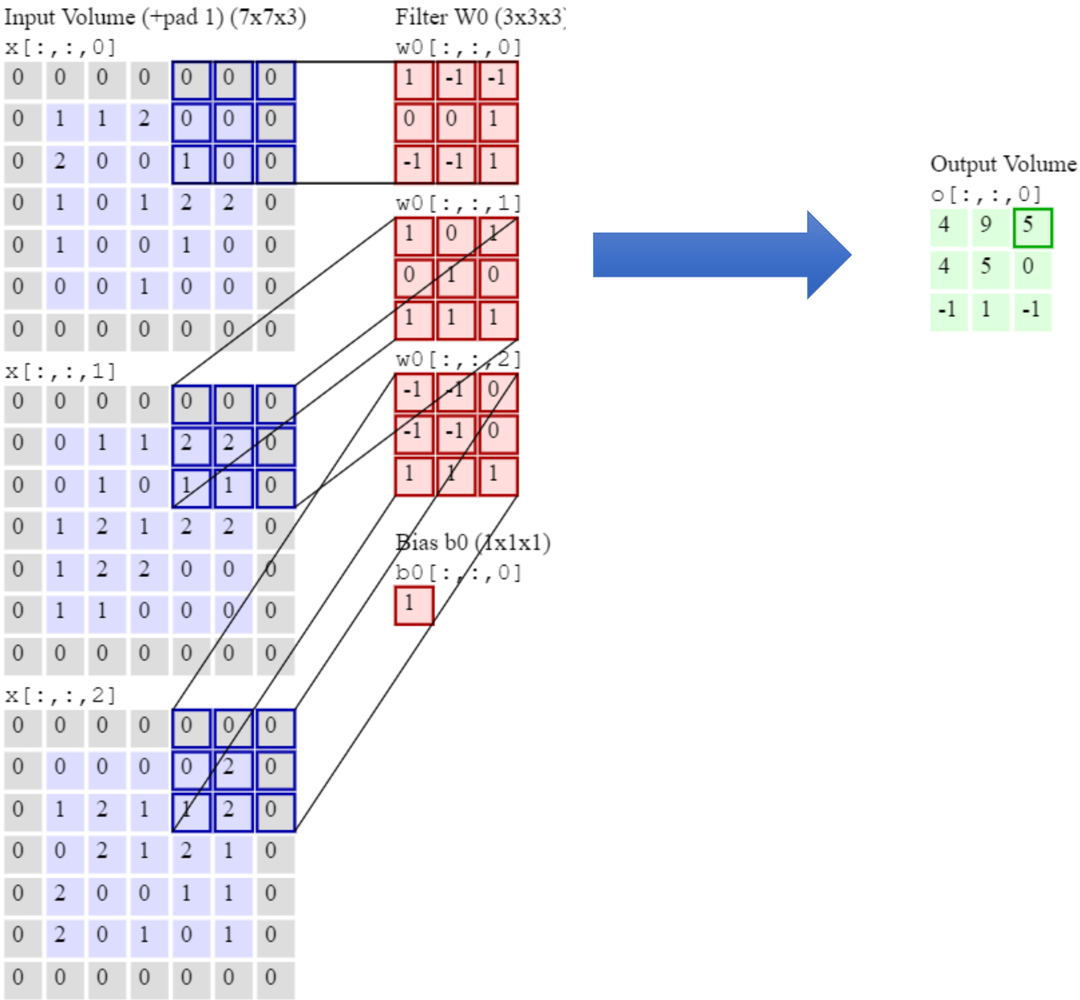


Figure 4.7: Convolution between an image and a filter (step 3). [adopted from 'skymind.ai' webpage]

If a similar process is maintained for the full image, a total of 9 steps have to be completed. Then a 3 by 3 output image can be found. This 3×3 output image is then passing through any layer of the activation function (suppose ReLU) and lastly passes through a pooling layer (suppose max pooling). Later, the final output of the 1st set of convolutions, ReLU and pooling layer is taken as an input of the 2nd set of layers. Thus, the hidden layer of CNN is structured and operated.

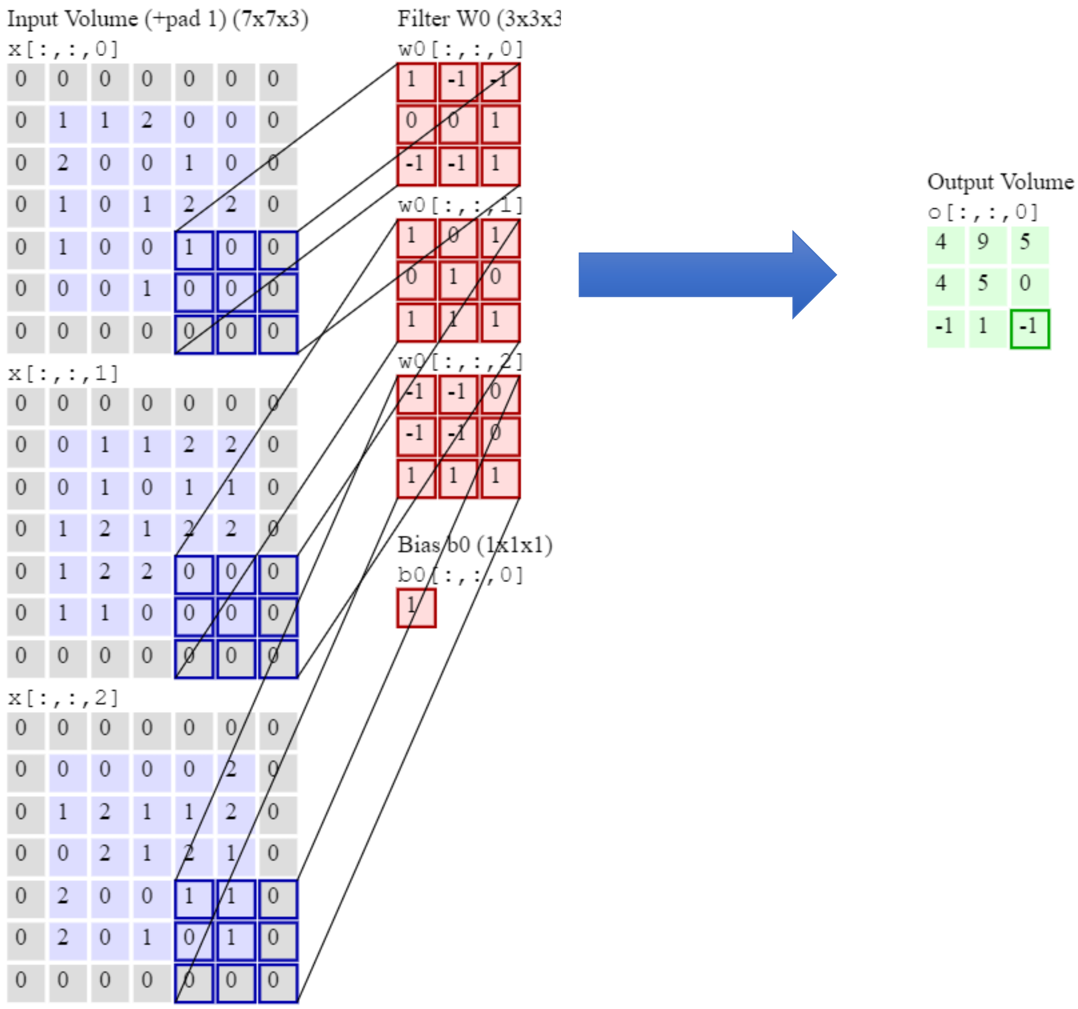


Figure 4.8: Convolution between an image and a filter (step 9). [adopted from 'skymind.ai' webpage]

4.4 Our Proposed CNN Model

Our Proposed CNN model architecture from the virtual image to classified emotion is shown in the figure below in Fig. 4.9. In our architecture, there remain three sets of Convolution, ReLU, and Pooling layer. Afterward, we used a flatten layer. Lastly, the dense layer is connected. Here we used a dropout of 0.25 for reducing more complexity.

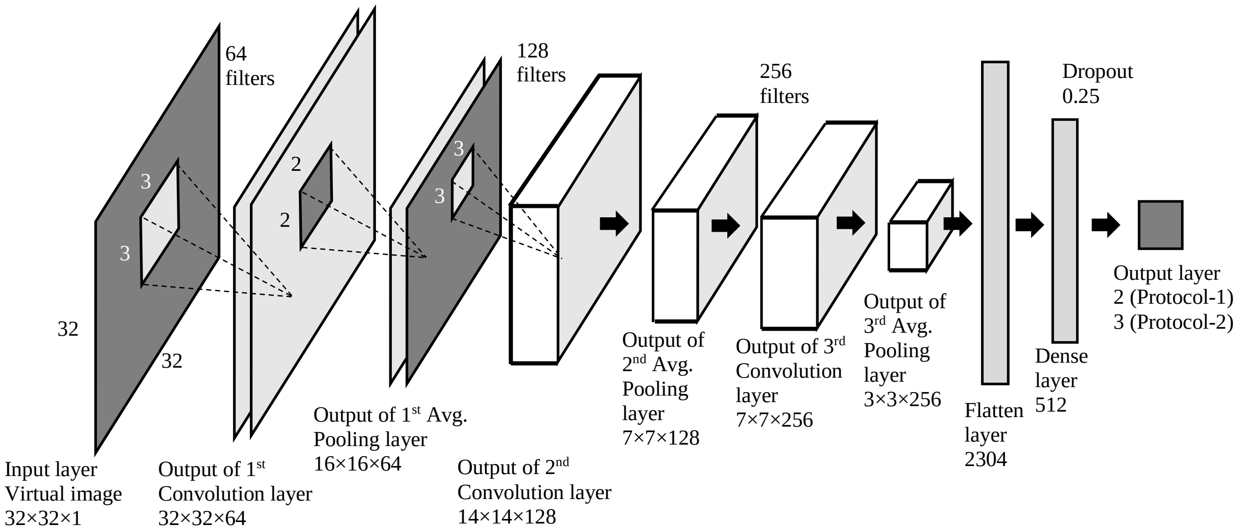


Figure 4.9: The proposed CNN model.

4.5 Convolutional Layer

The convolutional layer is the main layer of any CNN network. This layer performs matrix multiplication between a certain portion of input images and a specific shape of kernel. The size of a kernel was small compared with the image but has a large volume. In any CNN network, many filters are passing through a single image and create a map of edge. This process can be easily explained by understanding the similarity of moving a magnifying glass over a single image to find something. The CNN network passes different arrangement of the filter in horizontal then vertical order one by one and learns a lot of features automatically by continuing this same procedure. In spite of focusing a single pixel at once, this type of architecture focuses on several nearest pixels of a portion. The sample of a simple convolution is shown in Fig. 4.10.

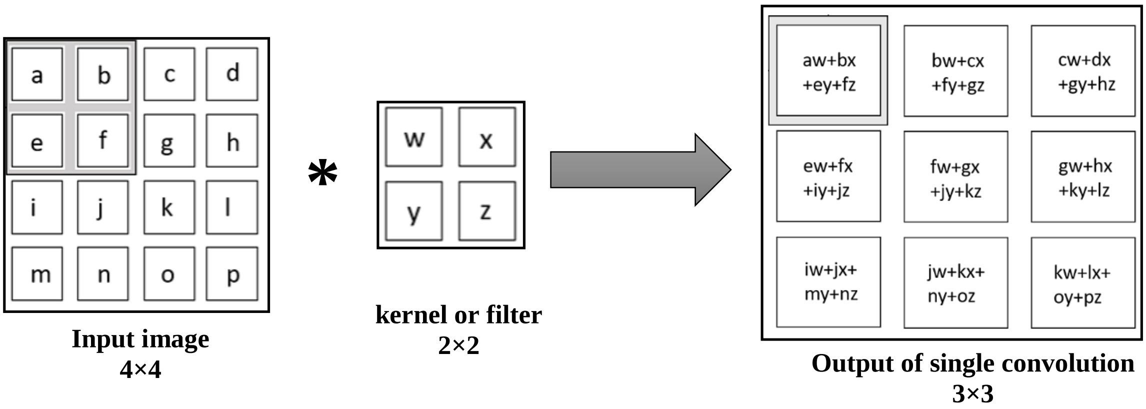


Figure 4.10: A sample of convolutional operation.

For any single layer of convolution if the size of the input image and the filter is expressed as like as (4.1) and (4.2), respectively; the dimension of the output after completing a convolutional neural network will be *nH×nW×nC.*

$$size\_of\_input = n_H^{[l-1]} \times n_W^{[l-1]} \times n_C^{[l-1]} \tag{4.1}$$

$$size\_of\_filter = f_H^{[l]} \times f_W^{[l]} \times n_C^{[l-1]} \tag{4.2}$$

$$size\_of\_weight = \left(f_H^{[l]} \times f_W^{[l]} \times n_C^{[l-1]}\right) \times n_C^{[l]} \tag{4.3}$$

$$size\_of\_bias = 1 \times 1 \times 1 \times n_C^{[l]} \tag{4.4}$$

Where, (4.3) and (4.4) represents the weight and bias, respectively. The output layer at level *l* can be calculated as follows,

$$n_H = \frac{n_H^{[l-1]} + 2p^{[l]} - f^{[l]}}{S^{[l]}} + 1 \tag{4.5}$$

$$size\_of\_output = n_H^{[l]} \times n_W^{[l]} \times n_C^{[l]} \tag{4.6}$$

The calculation of n_w is exactly similar to n_H like equation (4.5). Where *p[l]* indicates the dimension of padding and *S[l]* indicates the value of stride at level *l*. Simply, according to pure mathematics, the convolution operation in between two signals *h(t)* and *k(t)* can be expressed as follows.

$$(h * k)(t) = \int_{-\infty}^{\infty} h(t-\tau)k(\tau)d\tau \tag{4.7}$$

We used 64, 128 and 256 filters of size 3×3 and of stride 1 in the 1st, 2nd, and 3rd convolutional layer respectively.

4.5.1 Padding

In convolutional neural network according to padding, there remain two types of convolution. Padding is an additional layer of zero which ensures the importance of the corner pixel of image data. Since we don't want to lose the information from the pixels of a corner, we used the same convolution.

- Valid convolution: No padding.
- Same convolution: Padding is present.

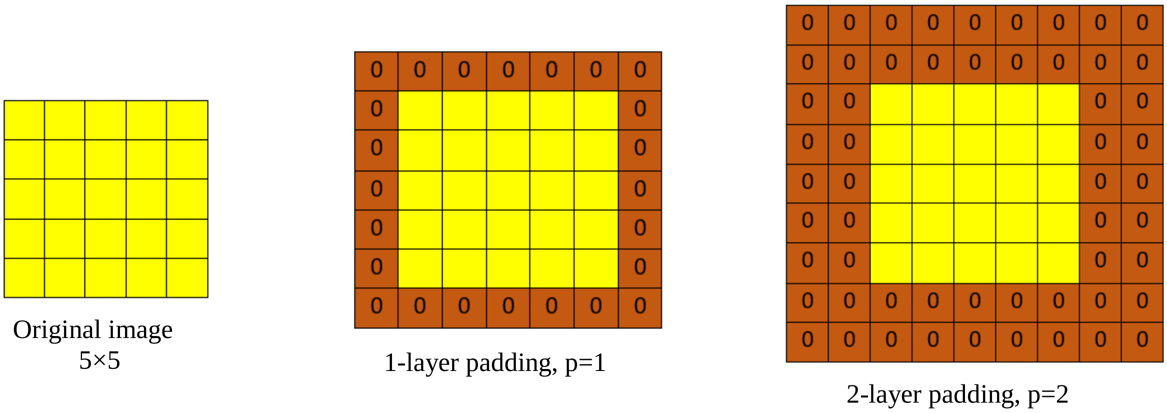


Figure 4.11: Layer of padding.

4.5.2 Stride

Another important terminology is stride. It indicates the number of pixels jumping off the filter on the input image. For instances, for the previous image the stride 2 means, 2nd filter move 2-pixel right for the next operation (Fig. 4.12).

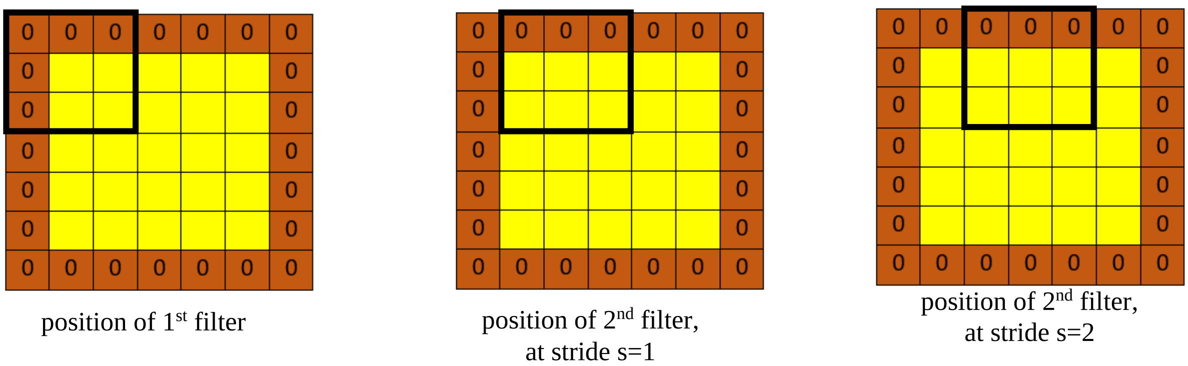


Figure 4.12: Easy illustration of stride.

4.6 Activation Functions

The activation function decides if the neuron would fire or not. There remains a different activation function like the following list. We used 'ReLU', 'sigmoid' and 'softmax' activation function in the different stages of the emotion recognition algorithm.

- Rectified Linear Unit (ReLU)
- Exponential Linear Unit (ELU)
- Scaled Exponential Linear Unit (SELU)
- hyperbolic tangent (thah)
- Maxout
- softmax
- softsign
- softplus
- sigmoid

4.6.1 ReLU Activation Function

Here, we used ReLU as activation function which is the best fit with CNN architecture and faster. The ReLU has a positive feature compared to 'sigmoid' and 'tanh' that never get saturated with a large value of x. The ReLU activation function has a great advantage than 'sigmoid' and 'tanh' that it is more reliable and accelerates the convergence by six times. As a result, the ReLU activation function has become very attractive in the last few years. The equation of this activation function and its graphical representation is given below.

$$f(x) = max\ (0, x) \qquad (4.8)$$

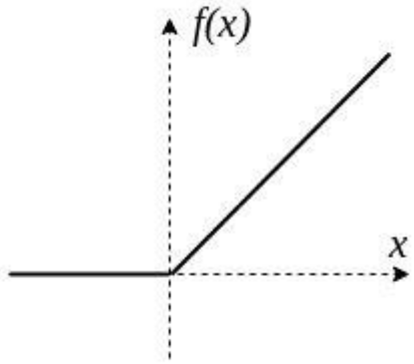


Figure 4.13: The ReLU activation function.

4.6.2 The ‘sigmoid’ Activation Function

In protocol 1, we made a binary classification of emotion. For doing this, we used ‘sigmoid’ activation function in the last layer. The ‘sigmoid’ is a smooth non-linear activation function and looking like *S* shape. It returns the probabilities of a class and since the probability’s ranges from 0 to 1, its range is also the same. For binary classification, ‘sigmoid’ activation function is popularly used.

The equation of ‘sigmoid’ activation function is $\sigma(z) = \frac{1}{1+e^{-z}}$ (4.9)

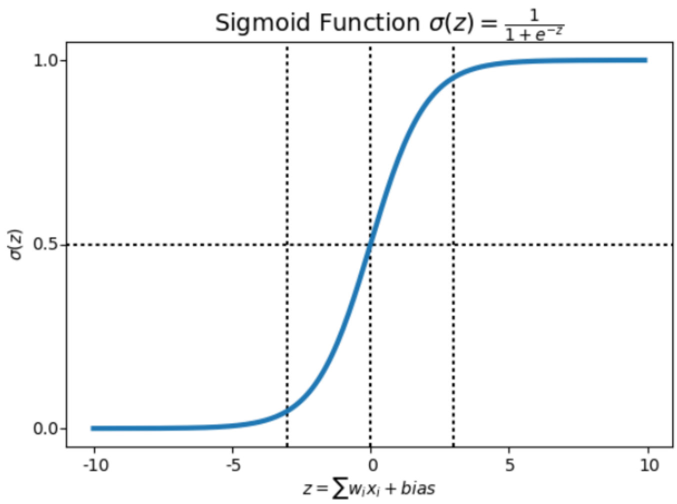


Figure 4.14: The 'sigmoid' activation function.

4.6.2 The ‘softmax’ Activation Function

Protocol 2 is not a binary classification because here we classified emotion into three class. Therefore, ‘sigmoid’ is not suitable for this purpose. For multi-class classification ‘softmax’ activation function is perfect. It generates the probability ranges from 0 to 1. The sum of the probability of all class will be 1. For classification, the ‘softmax’ activation function returns the probability of every class and lastly, it targets the class which belongs to the largest probability.

The ‘softmax’ function is defined by the following equation

$$h_\theta(x) = \frac{1}{1 + \exp(-\theta^T x)} \tag{4.10}$$

Here h_θ is the scalar output of softmax. the range of $h_\theta(x) \in R$ *and* $0 < h_\theta(x) < 1$. The θ and x are the vectors of weights and input values respectively.

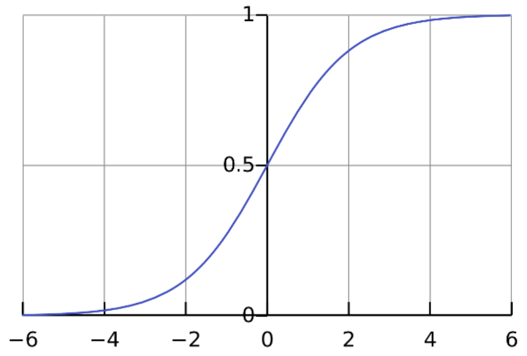


Figure 4.15: The 'softmax' activation function.

It may be noted that the ‘sigmoid’ function is used for the two-class logistic regression, whereas the ‘softmax’ function is used for the multiclass logistic regression. In the two-class logistic regression, the predicted probabilities are as follows, using the sigmoid function:

$$\Pr(Y_i = 0) = \frac{e^{-\beta.X_i}}{1 + e^{-\beta_0.X_i}} \tag{4.11}$$

$$\Pr(Y_i = 1) = 1 - \Pr(Y_i = 0) = \frac{1}{1 + e^{-\beta.X_i}} \tag{4.12}$$

In the multiclass logistic regression, with KK classes, the predicted probabilities are as follows, using the softmax function:

$$\Pr(Y_i = k) = \frac{e^{\beta_k.X_i}}{\sum_{0 \le c \le k} e^{\beta_c.X_i}} \tag{4.13}$$

The ‘softmax’ activation function is actually considered as an extension of sigmoid function to the multiclass case. The logistic regression with k=2 classes can be explained as follows. When *β=−(β0−β1)*, the same probabilities can be found into the two-class logistic regression using the sigmoid function.

$$\Pr(Y_i = 0) = \frac{e^{\beta_0.X_i}}{\sum_{0 \le c \le k} e^{\beta_c.X_i}} = \frac{e^{\beta_0.X_i}}{e^{\beta_0.X_i} + e^{\beta_1.X_i}} = \frac{e^{(\beta_0-\beta_1).X_i}}{e^{(\beta_0-\beta_1).X_i} + 1} = \frac{e^{-\beta.X_i}}{1 + e^{-\beta.X_i}} \tag{4.14}$$

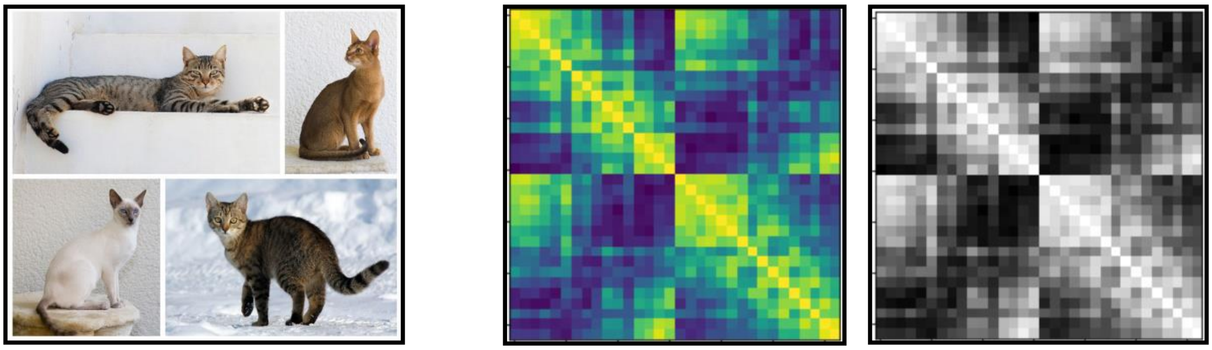
$$\Pr(Y_i = 1) = \frac{e^{\beta_1.X_i}}{\sum_{0 \le c \le k} e^{\beta_c.X_i}} = \frac{e^{\beta_1.X_i}}{e^{\beta_0.X_i} + e^{\beta_1.X_i}} = \frac{1}{e^{(\beta_0-\beta_1).X_i} + 1} = \frac{1}{1 + e^{-\beta.X_i}} \tag{4.15}$$

4.7 Pooling Layer

To reduce the amount of computational complexity and weight pooling is essential. By pooling the spatial size of the image can be reduced. There are two types of pooling layer.

- Max pooling
- Average pooling

The most popular pooling is max pooling. In our experiments, the PCC based virtual image is a texture type image. The classification from any natural image is not similar to the classification from texture type image. To extract information from natural image max pooling is suitable.



Natural Image: Max pooling is perfect

Texture Type Image: Average pooling is perfect

Figure 4.16: Natural vs texture type image.

But for texture type image it contains no particular pattern. As a result, the algorithm works well by average pooling. Here we used a 2×2 matrix with stride 1 for pooling.

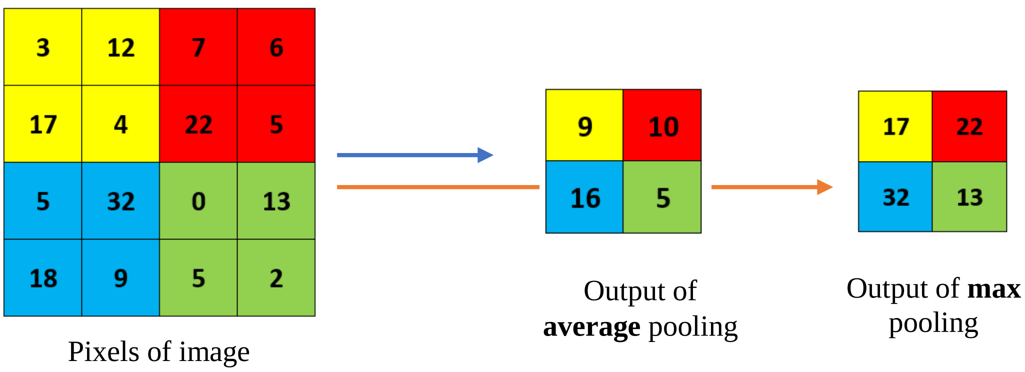


Figure 4.17: Average and max-pooling.

4.8 Optimization Algorithm

In a neural network, to update the model parameter like weight and bias values there needs an optimization algorithm. The optimization algorithm runs a procedure of finding the optimum or satisfactory solution. The main function of an optimization algorithm is to minimize the loss or error of a neural network model. Some popular examples of the optimization algorithm are

- Stochastic Gradient Descent (SGD)
- Batch Gradient Descent (BGD)
- Nadam (Nesterov Adam optimizer)
- Adagrad (Adaptive Gradient)
- Adadelta (Adaptive Delta)
- Adam (Adaptive Moment Estimation)
- Adamax
- RMSprop

The Adaptive Moment Estimation (Adam) gradient descent algorithm was used to optimize the neural network. 'Adam' can be thought of as a combination of heuristics of 'RMSprop' and 'SGD' with momentum. Hence, we used 'Adam' optimizer in our experiment. For 'Adam' the update rule of the parameter is like below. For each parameter w_j; If we represent;

η= Initial learning rate;

g_t= Gradient at time t along w_j;

V_t=Exponential average of the gradient along w_j;

S_t= Exponential average of the square of the gradient along w_j;

β_1, β_2= Hyperparameters;

Then,

$$V_t = \beta_1 * V_{t-1} - (1 - \beta_1) * g_t \tag{4.16}$$

$$S_t = \beta_2 * S_{t-1} - (1 - \beta_2) * g_t^2 \tag{4.17}$$

$$\Delta w_t = -\eta \frac{V_t}{\sqrt{S_t + \epsilon}} * g_t \tag{4.18}$$

$$w_{t+1} = w_{t+1} + \Delta w_t \tag{4.19}$$

Again, it can be written that,

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\hat{v}_t} + \in} \hat{m}_t \tag{4.20}$$

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t} \tag{4.21}$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t} \tag{4.22}$$

Where m_t and v_t indicates the estimates of the 1st moment (i.e. mean) and the 2nd moment (i.e. un-centered variance) of the gradient respectively.

Here for 'Adam' optimizer the best fit value of decay rate $\beta_1 = 0.9$, $\beta_2 = 0.999$ and $\epsilon = 10^{-8}$.

4.9 Loss Function

A loss function is used to optimize the parameter values in a neural network. There remain several loss functions in machine learning architecture. Bur the configuration of a network is an indicator for selecting appropriate loss function. Every neural network learns from the training data and classifies the output according to the desire of the program. But the challenge is to determine the exact values of weights and biases for a certain network. It may be noted that the main target of a perfect neural network is not to increase the accuracy; the main target is to reduce the loss. The optimization algorithm always updates the values of weights and biases as to reduce the value of the loss.

According to Alpaydin (2004) [72] "*The function we want to minimize or maximize is called the objective function or criterion. When we are minimizing it, we may also call it the cost function, loss function, or error function*."

The authors Reed et al. [73] gave the opinion that "*It is important, therefore, that the function faithfully represents our design goals. If we choose a poor error function and obtain unsatisfactory results, the fault is ours for badly specifying the goal of the search*."

Table 4.1: Network architecture and loss function

Types of Problem	Configuration of the output layer	Loss or Cost function
Regression	One node with a 'sigmoid' activation function.	Mean Squared Error (MSE).
Binary Classification	One node with a 'sigmoid' activation function.	Binary Cross-Entropy, also referred to as Logarithmic loss
Multi-Class Classification	One node for each class using the 'softmax' activation function.	Categorical Cross-Entropy

In our research work in the first protocol, we tried to classify positive and negative emotion. So, in this protocol, the type of classification is binary classification. As a result, in this architecture, we used a binary classification. On the other hand, in protocol 2, three types of emotion are classified. So, this is a multi-class classification problem and here we used categorical cross-entropy as the loss or cost function.

- In binary classification, where the number of classes *M=2*, cross-entropy can be calculated as calculating the quantity, $-[y\,log(p) + (1-y)\,log(1-p)]$
- If *M>2* (i.e. multiclass classification), Here *M=3*; we calculate a separate loss for each class label per observation and sum the result. Therefore, the loss will be equal to the following quantity $-\sum_{c=1}^{M} y_{o,c}\log(P_{o,c})$.

Where,

M= Class level,

y= binary indicator *(0,1)* if class label *c* is the correct classification for observation *o*,

P= predicted probability observation *o* is of class *c*.

4.10 Summary of CNN model

In our proposed CNN model, we used three sets of convolutional, ReLU and pooling layer. Lastly, flatten layer, dropout layer and dense layer are connected sequentially. We used the '**Google-Colab**' user interface to execute the code. The 'Google-Colab' is a cloud computing-based interface where python code can be executed like a Jupiter notebook interface. The summary of our CNN model is shown in Table 4.2 and Table 4.3 respectively. It may be noted that the two architecture is almost the same except the output shape of the last dense layer. As we classified 2 and 3 different emotion in protocol 1 and 2 respectively, the shape of the last dense layer is 2 and 3 respectively.

Table 4.2: The Summary of proposed CNN model for protocol-1

Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 32, 32, 64)	640
average_pooling2d_1 (Average)	(None, 16, 16, 64)	0
conv2d_2 (Conv2D)	(None, 14, 14, 128)	73856
average_pooling2d_2 (Average)	(None, 7, 7, 128)	0
conv2d_3 (Conv2D)	(None, 7, 7, 256)	295168
average_pooling2d_3 (Average)	(None, 3, 3, 256)	0
flatten_1 (Flatten)	(None, 2304)	0
dense_1 (Dense)	(None, 512)	1180160
dropout_1 (Dropout)	(None, 512)	0
dense_2 (Dense)	(None, 2)	1026
		Total params: 1,550,850
		Trainable params: 1,550,850
		Non-trainable params: 0

Table 4.3: The Summary of proposed CNN model for protocol-2

Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 32, 32, 64)	640
average_pooling2d_1 (Average)	(None, 16, 16, 64)	0
conv2d_2 (Conv2D)	(None, 14, 14, 128)	73856
average_pooling2d_2 (Average)	(None, 7, 7, 128)	0
conv2d_3 (Conv2D)	(None, 7, 7, 256)	295168
average_pooling2d_3 (Average)	(None, 3, 3, 256)	0
flatten_1 (Flatten)	(None, 2304)	0
dense_1 (Dense)	(None, 512)	1180160
dropout_1 (Dropout)	(None, 512)	0
dense_2 (Dense)	(None, 3)	1539
		Total params: 1,551,363
		Trainable params: 1,551,363
		Non-trainable params: 0

Chapter **5**

Result and Discussion

Chapter Outlines:

Chapter 5

Result and Discussion

5.1 Introduction

Deep learning has emerged the potential attraction of scholars in recent years. In our experiment, we used the 'DEAP' dataset of EEG signal to classify positive, negative emotion and positive negative and neutral type emotion. The emotion-related EEG and peripheral signal are firstly converted into PCC based virtual images. Afterward, the virtual images are feed into a CNN based classification algorithm. In protocol-1, only positive and negative emotion are classified. In the 2nd protocol, the positive, negative and neutral emotion are classified. Both of the two protocols consist of two different tasks (i) valence recognition task and (ii) arousal recognition task. After considering the three sub-band data we used 25600 data for three sub-band. Therefore, there remains a 25600×4=102,400 labeled virtual image in total. The data points are split like the ration training data 70%, validation data 20% and testing data 10%. The classification accuracy is given in details in the next section.

5.2 Result of Classification

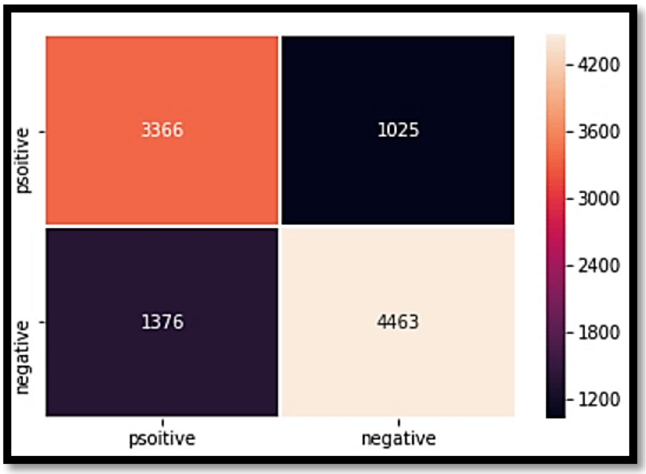
We classified emotion by classifying the level of valence and arousal. This classification is done by using the logistic regression on the two indicators of emotional state namely valence and arousal. The overall result of our experiment can be described into two following sections.

5.2.1 Protocol-1

In the 'DEAP' dataset there remains EEG data of 32 participants. Every participant is shown 40 different emotional audio-visual stimuli and recorded the data of valence, arousal, dominance, liking, and familiarity. From the 'DEAP' dataset of EEG, the rating of valence is firstly labeled into two categories. The valence value from 0 to 4.99 is categorized as low valence and 5.00 to 9.00 is categorized as high valence. The similar scaling is maintained into for arousal classification.

5.2.1.1 Valence Classification Task of Protocol-1

We classified positive and negative emotion using valence firstly. The confusion matrix is shown in the following Fig. 5.1. It indicates the result of test data that means 10% of all data.



3366 32.9%	1025 10.01%	76.65% 23.23%
1376 13.45%	4463 43.62%	76.43% 23.57%
70.98% 29.02%	81.32% 18.68%	76.52% 23.48%

Figure 5.1: Confusion matrix of two class valence classification.

The model accuracy and loss curve are shown in Fig. 5.2 and Fig. 5.3 respectively.

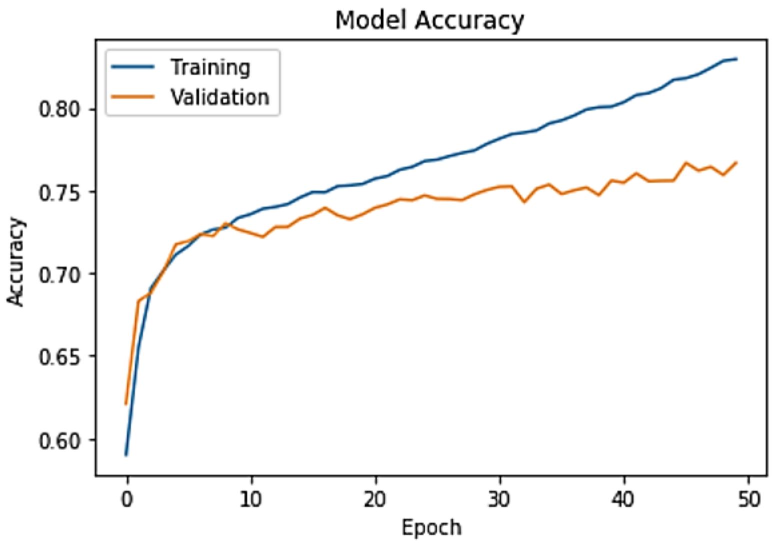


Figure 5.2: Model accuracy of two class valence classification.

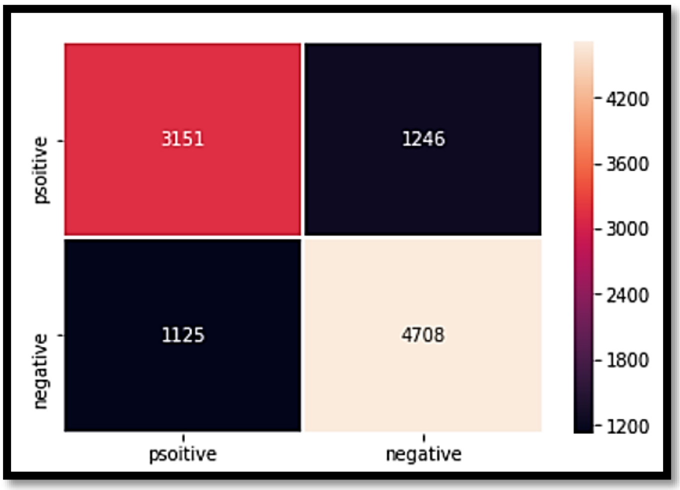
Figure 5.3: Model loss of two class valence classification.

Table 5.1: Classification Report of two class valence classification

Classification Report				
	Precision	**Recall**	**F1 score**	**support**
0	0.71	0.77	0.74	4391
1	0.81	0.76	0.79	5839
Accuracy			0.77	10230
Macro avg	0.76	0.77	0.76	10230
Weighted avg	0.77	0.77	0.77	10230
Total Classification Accuracy = **76.52%**				

5.2.1.2 Arousal Classification Task of Protocol-1

After classifying the valence level next we classified the arousal data for emotion recognition.



3151 30.81%	1246 12.17%	71.67% 28.33%
1125 10.99%	4708 46.02%	80.71% 19.29%
73.69% 26.31%	79.07% 20.93%	76.82% 23.18%

Figure 5.4: Confusion matrix of two class arousal classification.

The confusion matrix of arousal recognition is given in Fig. 5.4.

The model accuracy and loss curve are shown in Fig. 5.5 and Fig. 5.6 respectively.

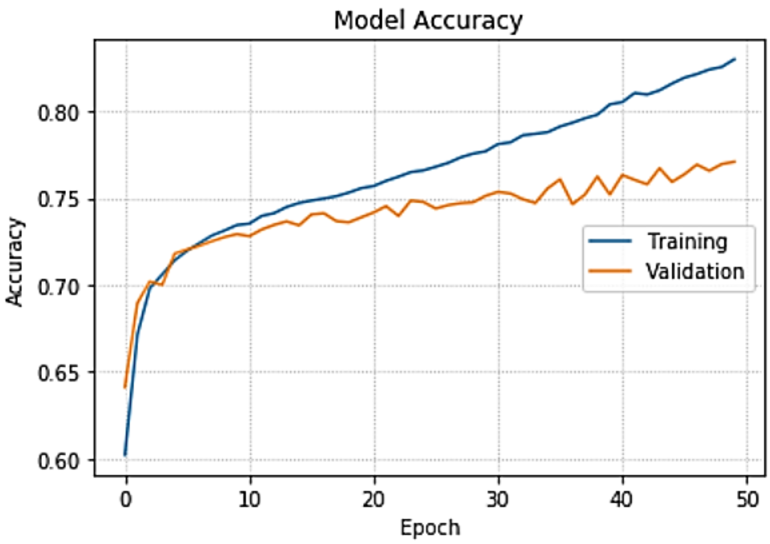


Figure 5.5: Model accuracy of two class arousal classification.

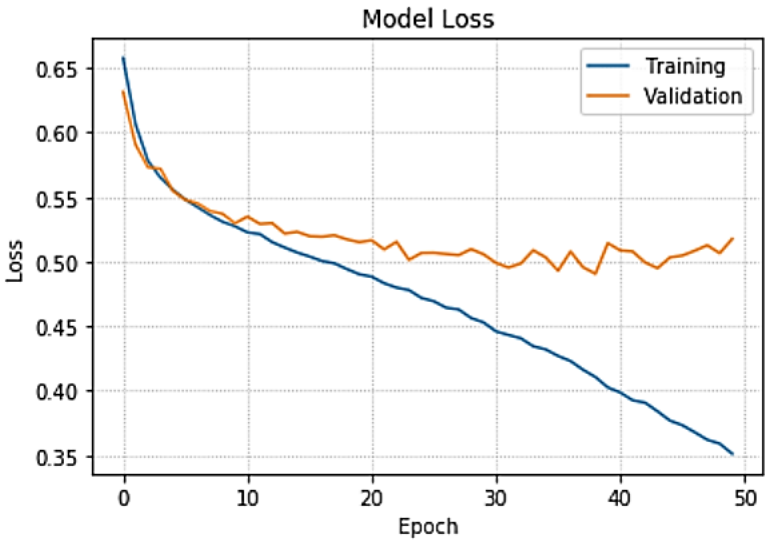


Figure 5.6: Model loss of two class arousal classification.

Table 5.2: Classification Report of two class arousal classification

Classification Report				
	Precision	**Recall**	**F1 score**	**support**
0	0.74	0.72	0.73	4397
1	0.79	0.81	0.80	5833
Accuracy			0.77	10230
Macro avg	0.76	0.76	0.76	10230
Weighted avg	0.77	0.77	0.77	10230
Total Classification Accuracy = **76.82%**				

5.2.2 Protocol-2

In this protocol-2 the three types of emotion are recognized. Here we tried to recognize positive, negative and neutral type emotion. The scaling was based on the value of valence and arousal. The total dataset label was firstly classified into three categories low, medium and high. Here low means it values of the level is 0 to 2.99, medium indicates the value from 3 to 5.99 and 6 to 9 are indicated as high value. These three levels indicate the positive, neutral and negative type emotion respectively.

5.2.2.1 Valence Classification Task of Protocol-2

The 'DEAP' EEG data are preprocessed and relabeled according to the rule of three categories division written in the previous section. The confusion matrix shows the overall accuracy is

1030 10.06%	370 3.61%	269 2.63%	61.71% 38.29%
436 4.27%	2565 25.07%	1120 10.95%	62.24% 37.78%
242 2.36%	755 7.38%	3443 33.65%	77.54% 22.45%
60.30% 39.70%	69.51% 30.48%	71.25% 28.75%	68.80% 31.20%

Figure 5.7: Confusion matrix of three class valence classification.

The model accuracy curve is shown in the following Fig. 5.8.

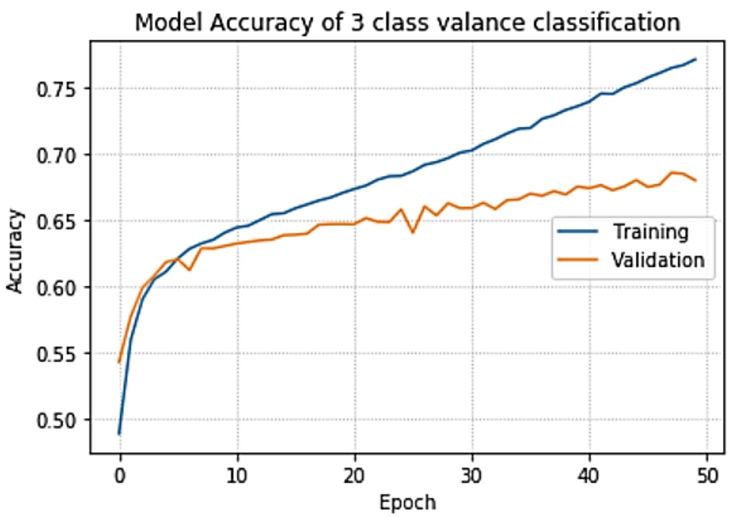


Figure 5.8: Model accuracy curve of three class valence classification

The model loss curve was like the following Fig. 5.9.

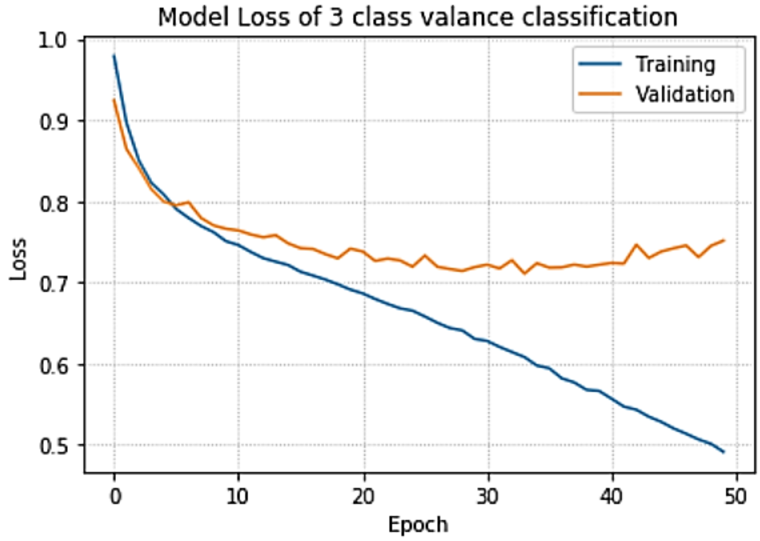


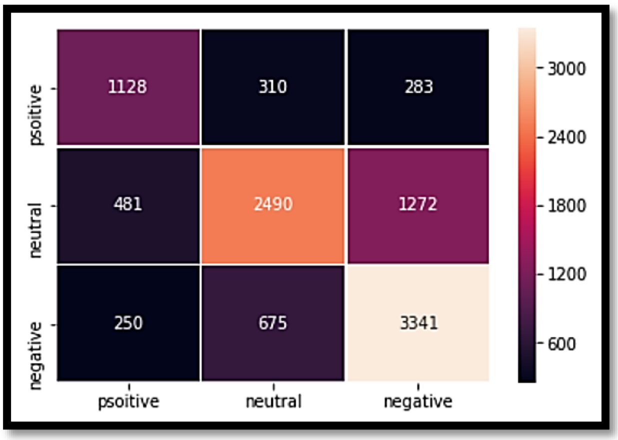
Figure 5.9: Model loss curve of three class valence classification.

Table 5.3: Classification Report of three class valence classification

Classification Report				
	Precision	**Recall**	**F1 score**	**support**
0	0.60	0.62	0.61	1669
1	0.70	0.62	0.66	4121
2	0.71	0.78	0.74	4440
Accuracy			0.69	10230
Macro avg	0.67	0.67	0.67	10230
Weighted avg	0.69	0.69	0.69	10230
Total Classification Accuracy = **68.80%**				

5.2.2.1 Arousal Classification Task of Protocol-2

In this section, the arousal data for emotional EEG data are categorized into three class. The confusion matrix is given below.



1128 11.02%	310 3.03%	283 2.77%	65.54% 34.45%
481 4.70%	2490 24.34%	1272 12.43%	58.68% 41.31%
250 2.44%	675 6.59%	3341 32.65%	78.32% 21.68%
60.67% 39.32%	71.65% 28.34%	68.23% 31.76%	68.03% 31.97%

Figure 5.10: Confusion matrix of three class arousal classification.

The model accuracy and loss curve are shown in Fig. 5.11 and Fig. 5.12 respectively.

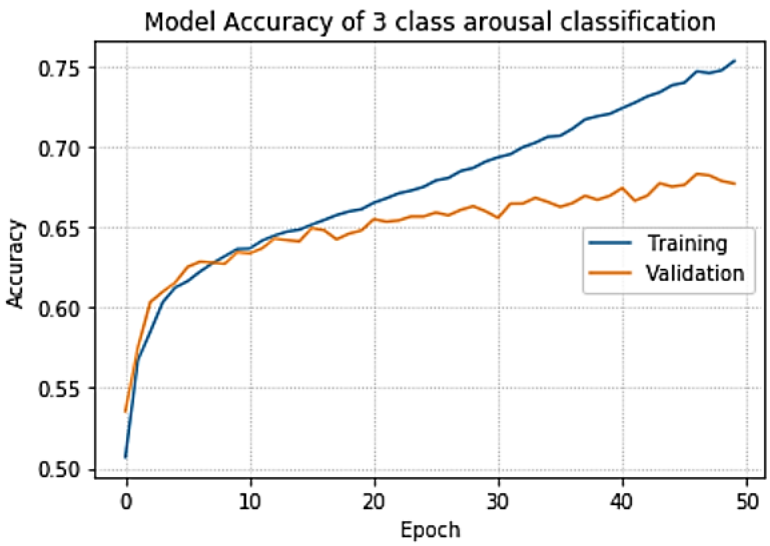


Figure 5.11: Model accuracy curve of three class arousal classification.

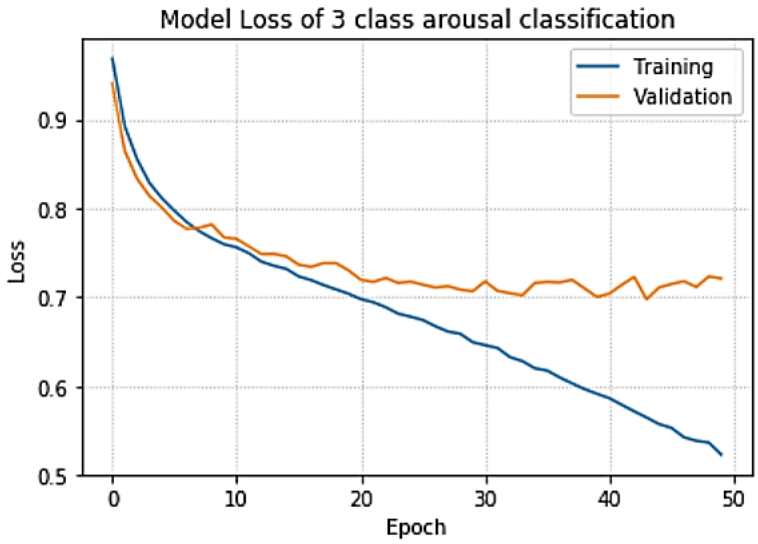


Figure 5.12: Model loss of three class arousal classification.

Table 5.4: Classification Report of three class arousal classification

Classification Report				
	Precision	**Recall**	**F1 score**	**support**
0	0.61	0.66	0.63	1721
1	0.72	0.59	0.65	4243
2	0.68	0.78	0.73	4266
Accuracy			0.68	10230
Macro avg	0.67	0.68	0.67	10230
Weighted avg	0.68	0.68	0.68	10230
Total Classification Accuracy = **68.03%**				

5.3 Comparison of Results

After all, the overall accuracy of our proposed model is shown in the following table.

Table 5.5: The overall system accuracy in percent

Protocol	No. of emotion	Accuracy (%)	
		Valence	Arousal
Protocol-1	2	76.52%	76.82%
Protocol-2	3	68.80%	68.03%

From the above Table 5.5, it is clear that the maximum accuracy of 76.82% is found in the 2-class arousal classification task. Another important fact that the percentage accuracy of protocol-1 is comparatively higher than the accuracy of protocol-2. The cause is that whenever the number of class to be categorized will be increased then the loss will also increase. As a result, the accuracy will decrease.

Due to the wide range of variation in a number of channels, no of emotion, extracted feature, an applied method for classification, comparison of accuracy in this type of research is not actually significant. Despite this, comparison Table 5.6 is shown below. It may be noted that Koelstra et al. [58] used the PSD feature with Naive Bayes classifier with 32 EEG channels and achieved 62% overall accuracy.

Table 5.6: Comparison of accuracy of different emotion recognition methods

Author Name		Year	No. of emotion	Accuracy (%)	
				Valance	Arousal
Koelstra et al [7]		2012	4	57.6	62.00
Naser and Saha [74]		2013	2	64.3	66.20
Bahari et al. [75]		2013	3	58.05	64.56
Zhuang et al. [76]		2014	4	70.90	67.10
Jirayucharoensak et al. [56]		2014	9	49.52	46.03
Chen et al. [48]		2015	2	76.17	73.59
Tripathi et al. [57]		2017	2	81.41	73.36
Our Method	Protocol-1	2019	2	76.52	76.82
	Protocol-2	2019	3	68.80	68.03

The classification accuracy is largely dependent on the no. of output class. The more the no. of emotion have to be classified the less the classification accuracy would be found. However, in this work, our contribution is to propose a CNN based method to recognize emotion from multichannel EEG data, which reduce the trouble of human to manually extract the significant feature.

5.4 Limitations

There remains some limitation of our proposed method of the emotion recognition system. In future, we will continue our research to overcome these limitations.

- As our experiment is done with the Convolutional Neural Network and the performance of classification is largely dependent with the available number of training data, it will be better if we got raw EEG dataset of participant more than 32.
- Only one feature (i.e. PCC) is considered as a feature matrix. More feature matrix may give a more accurate result. But it would be more time consuming and increase the computational complexity.
- This recognition system requires some time during the training period. However, it works within a short time during testing and validation after training once.

Chapter **6**

Conclusion and Future Work

Chapter Outlines:

Chapter 6

Conclusion and Future Work

6.1 Conclusion

In this research work, a Convolutional Neural Network was implemented on a virtual image data to classify emotion. The EEG data were firstly reshaped and decomposed and then PCC was calculated to build a featured virtual image. The first objective of our work was to convert the one-dimensional data to significant image type two-dimensional data. It was completed successfully and its explanation is given in the third chapter. For this conversion, we calculated the Pearson's Correlation Coefficient (PCC) of different channels. Later a CNN architecture was proposed to recognize positive, negative and neutral emotions. Here we used three sets of convolutional ReLU and pooling layer in a suitable format. The convolutional layer is a core layer of this CNN layer. We used many training images as CNN outperforms with many training data. The recognition of an emotional task is completed by following the rule of classification using logistic regression. There was done two different protocols and, in both protocols, we checked the accuracy and model loss. We achieved 76.52% and 68.80% accuracy in valence classification and 76.82% and 68.03% accuracy in arousal classification for two and three classification protocol respectively. To enjoy the advantages of emotion recognition in practical field more accuracy and real time operation compatibility are the fundamental prerequisite. As a consequence, more research and studies are important in this field and in the field of channel reduction, significant feature extraction, deep network optimization etc.

Our achievement was satisfactory, but further improvement will be possible by taking a large set of data for training the model. But undoubtedly, our experiment provides the important information that Convolutional Neural Network with significant input data is a very effective and potential network in the research domain associated with human brain signal.

6.2 Future Work

The recognition of emotion is actually a challenging and difficult task. The practical implementation of this field is too much challenging. Many branches of artificial intelligence competitive company and researchers tried to apply the emotion classification in the practical field. To use practically the recognition accuracy, have to be more than that. The more suitable feature set has to be extracted. One limitation of our research is that it contains the EEG signal from 'DEAP' dataset. This type of data acquisition system is very bulky and not so flexible to use in the practical field. Moreover, the increment of a number of participants and the volume of the dataset will help to improve accuracy in the valence and arousal classification of emotion.

REFERENCES

1. Alarcao, S.M., and Fonseca, M.J., 2017, "Emotions recognition using EEG signals: a survey", *IEEE Transactions on Affective Computing*.
2. Vinola, C. and Vimaladevi, K., 2015, "A survey on human emotion recognition approaches, databases and applications", *ELCVIA: electronic letters on computer vision and image analysis*, pp.24-44.
3. Bethel, C.L., Salomon, K., Murphy, R.R. and Burke, J.L., 2007, August, "Survey of psychophysiology measurements applied to human-robot interaction", In *RO-MAN 2007-The 16th IEEE International Symposium on Robot and Human Interactive Communication*, pp. 732-737.
4. Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W. and Taylor, J.G., 2001, "Emotion recognition in human-computer interaction", *IEEE Signal Processing Magazine*, *18*(1), pp.32-80.
5. Fox, E., 2008, *Emotion science: cognitive and neuroscientific approaches to understanding human emotions*, Palgrave Macmillan.
6. Plutchik, R., 1997. The circumplex as a general model of the structure of emotions and personality.
7. Koelstra, S., Muhl, C., Soleymani, M., Lee, J.S., Yazdani, A., Ebrahimi, T., Pun, T., Nijholt, A. and Patras, I., 2012, "Deap: A database for emotion analysis; using physiological signals", *IEEE transactions on affective computing*, *3*(1), pp.18-31.
8. Niedermeyer, E. and da Silva, F.L. eds., 2005, "*Electroencephalography: basic principles, clinical applications and related fields"*, Lippincott Williams & Wilkins.
9. Pravdich-Neminsky, W., 1912. Ein versuch der registrierung der elektrischen gehirnerscheinungen. *Zentralbl Physiol*, *27*, pp.951-960.
10. Cherry K. "What Is the James-Lange Theory of Emotion?". Retrieved 30 April 2012.
11. The Tomkins Institute. "Applied Studies in Motivation, Emotion, and Cognition". Archived from the original on 19 March 2012. Retrieved 30 April 2012.
12. M. R. Islam and M. Ahmad, "Wavelet Analysis Based Classification of Emotion from EEG Signal," *2nd International Conference on Electrical, Computer and Communication Engineering (ECCE 2019)*, Cox's Bazar, Bangladesh, February 2019.
13. Zheng, W.L. and Lu, B.L., 2015. Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks. *IEEE Transactions on Autonomous Mental Development*, *7*(3), pp.162-175.

14. Mathersul, D., Williams, L.M., Hopkinson, P.J. and Kemp, A.H., 2008. Investigating models of affect: Relationships among EEG alpha asymmetry, depression and anxiety. *Emotion*, *8*(4), p.560.
15. Knyazev, G.G., Slobodskoj-Plusnin, J.Y. and Bocharov, A.V., 2010. Gender differences in implicit and explicit processing of emotional facial expressions as revealed by event-related theta synchronization. *Emotion*, *10*(5), p.678.
16. Sammler, D., Grigutsch, M., Fritz, T. and Koelsch, S., 2007. Music and emotion: electrophysiological correlates of the processing of pleasant and unpleasant music. *Psychophysiology*, *44*(2), pp.293-304.
17. Barsoum, E., Zhang, C., Ferrer, C.C. and Zhang, Z., 2016, October. Training deep networks for facial expression recognition with crowd-sourced label distribution. In *Proceedings of the 18th ACM International Conference on Multimodal Interaction* (pp. 279-283). ACM.
18. Soleymani, M., Asghari-Esfeden, S., Fu, Y. and Pantic, M., 2015. Analysis of EEG signals and facial expressions for continuous emotion detection. *IEEE Transactions on Affective Computing*, *7*(1), pp.17-28.
19. Kadiri, S.R., Gangamohan, P., Gangashetty, S.V. and Yegnanarayana, B., 2015. Analysis of excitation source features of speech for emotion recognition. In *Sixteenth Annual Conference of the International Speech Communication Association*.
20. Valderas, M.T., Bolea, J., Laguna, P., Vallverdú, M. and Bailón, R., 2015, August. Human emotion recognition using heart rate variability analysis with spectral bands based on respiration. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 6134-6137). IEEE.
21. Sethu, V., Epps, J. and Ambikairajah, E., 2015. Speech based emotion recognition. In *Speech and Audio Processing for Coding, Enhancement and Recognition* (pp. 197-228). Springer, New York, NY.
22. Piana, S., Staglianò, A., Odone, F. and Camurri, A., 2016. Adaptive body gesture representation for automatic emotion recognition. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, *6*(1), p.6.
23. Vaskinn, A., Sundet, K., Østefjells, T., Nymo, K., Melle, I. and Ueland, T., 2016. Reading emotions from body movement: a generalized impairment in schizophrenia. *Frontiers in Psychology*, *6*, p.2058.
24. Jenke, R., Peer, A. and Buss, M., 2014. Feature extraction and selection for emotion recognition from EEG. *IEEE Transactions on Affective Computing*, *5*(3), pp.327-339.

25. Zhuang, N., Zeng, Y., Tong, L., Zhang, C., Zhang, H. and Yan, B., 2017. Emotion recognition from EEG signals using multidimensional information in EMD domain. *BioMed research international*, *2017*.
26. Oh, S.H., Lee, Y.R. and Kim, H.N., 2014. A novel EEG feature extraction method using Hjorth parameter. *International Journal of Electronics and Electrical Engineering*, *2*(2), pp.106-110.
27. Mehmood, R.M. and Lee, H.J., 2015. Towards emotion recognition of EEG brain signals using Hjorth parameters and SVM. *Adv. Sci. Technol. Lett., Biosci. Med. Res.*, *91*, pp.24-27.
28. Mehmood, R.M. and Lee, H.J., 2015. EEG based emotion recognition from human brain using Hjorth parameters and SVM. *International Journal of Bio-Science and Bio-Technology*, *7*(3), pp.23-32.
29. Lee, G., Kwon, M., Sri, S.K. and Lee, M., 2014. Emotion recognition based on 3D fuzzy visual and EEG features in movie clips. *Neurocomputing*, *144*, pp.560-568.
30. Schaaff, K. and Schultz, T., 2009, September. Towards emotion recognition from electroencephalographic signals. In *2009 3rd international conference on affective computing and intelligent interaction and workshops* (pp. 1-6). IEEE.
31. Al-Fahoum, A.S. and Al-Fraihat, A.A., 2014. Methods of EEG signal features extraction using linear analysis in frequency and time-frequency domains. *ISRN Neuroscience*, *2014*.
32. Hatamikia, S., Maghooli, K. and Nasrabadi, A.M., 2014. The emotion recognition system based on autoregressive model and sequential forward feature selection of electroencephalogram signals. *Journal of medical signals and sensors*, *4*(3), p.194.
33. Zhang, Y., Ji, X. and Zhang, S., 2016. An approach to EEG-based emotion recognition using combined feature extraction method. *Neuroscience Letters*, *633*, pp.152-157.
34. Mohammadi, Z., Frounchi, J. and Amiri, M., 2017. Wavelet-based emotion recognition system using EEG signal. *Neural Computing and Applications*, *28*(8), pp.1985-1990.
35. Tang, H., Liu, W., Zheng, W.L. and Lu, B.L., 2017, November. Multimodal emotion recognition using deep neural networks. In *International Conference on Neural Information Processing* (pp. 811-819). Springer, Cham.
36. Hadjidimitriou, S. K., and Hadjileontiadis, L.J., 2013. EEG-based classification of music appraisal responses using time-frequency analysis and familiarity ratings. *IEEE Transactions on Affective Computing*, *4*(2), pp.161-172.

37. Li, M. and Lu, B.L., 2009, September. Emotion classification based on gamma-band EEG. In *2009 Annual International Conference of the IEEE Engineering in medicine and biology society* (pp. 1223-1226). IEEE.
38. Soleymani, M., Pantic, M. and Pun, T., 2011. Multimodal emotion recognition in response to videos. *IEEE transactions on affective computing*, *3*(2), pp.211-223.
39. Duan, R.N., Zhu, J.Y. and Lu, B.L., 2013, November. Differential entropy feature for EEG-based emotion classification. In *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)* (pp. 81-84). IEEE.
40. Lu, Y., Zheng, W.L., Li, B. and Lu, B.L., 2015, June. Combining eye movements and EEG to enhance emotion recognition. In *Twenty-Fourth International Joint Conference on Artificial Intelligence*.
41. Knyazev, G.G., Slobodskoj-Plusnin, J.Y. and Bocharov, A.V., 2010. Gender differences in implicit and explicit processing of emotional facial expressions as revealed by event-related theta synchronization. *Emotion*, *10*(5), p.678.
42. Zheng, W.L., Zhu, J.Y. and Lu, B.L., 2017. Identifying stable patterns over time for emotion recognition from EEG. *IEEE Transactions on Affective Computing*.
43. Atkinson, J. and Campos, D., 2016. Improving BCI-based emotion recognition by combining EEG feature selection and kernel classifiers. *Expert Systems with Applications*, *47*, pp.35-41.
44. Mert, A. and Akan, A., 2018. Emotion recognition from EEG signals by using multivariate empirical mode decomposition. *Pattern Analysis and Applications*, *21*(1), pp.81-89.
45. Katsigiannis, S. and Ramzan, N., 2017. DREAMER: A database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-shelf devices. *IEEE Journal of biomedical and health informatics*, *22*(1), pp.98-107.
46. He, L., Hu, D., Wan, M., Wen, Y., Von Deneen, K.M. and Zhou, M., 2015. Common Bayesian network for classification of EEG-based multiclass motor imagery BCI. *IEEE Transactions on Systems, Man and Cybernetics: Systems, 46*(6), pp.843-854.
47. Velchev, Y., Radeva, S., Sokolov, S. and Radev, D., 2016, July. Automated estimation of human emotion from EEG using statistical features and SVM. In *2016 Digital Media Industry & Academic Forum (DMIAF)* (pp. 40-42). IEEE.
48. Chen, M., Han, J., Guo, L., Wang, J., and Patras, I., 2015, September. Identifying valence and arousal levels via connectivity between EEG channels. In *2015*

International Conference on Affective Computing and Intelligent Interaction (ACII) (pp. 63-69). IEEE.

49. Speier, W., Arnold, C., Lu, J., Deshpande, A. and Pouratian, N., 2014. Integrating language information with a hidden Markov model to improve communication rate in the P300 Speller. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *22*(3), pp.678-684.
50. Zheng, W.L. and Lu, B.L., 2016, July. Personalizing EEG-based effective models with transfer learning. In *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence* (pp. 2732-2738). AAAI Press.
51. Murugappan, M., Rizon, M., Nagarajan, R., Yaacob, S., Zunaidi, I. and Hazry, D., 2007. EEG feature extraction for classifying emotions using FCM and FKM. *International Journal of Computers and Communications*, *1*(2), pp.21-25.
52. Khosrowabadi, R., Quek, H.C., Wahab, A. and Ang, K.K., 2010, August. EEG-based emotion recognition using self-organizing map for boundary detection. In *2010 20th International Conference on Pattern Recognition* (pp. 4242-4245). IEEE.
53. Längkvist, M., Karlsson, L. and Loutfi, A., 2012. Sleep stage classification using unsupervised feature learning. *Advances in Artificial Neural Systems*, *2012*, p.5.
54. Li, K., Li, X., Zhang, Y. and Zhang, A., 2013, December. Affective state recognition from EEG with deep belief networks. In *2013 IEEE International Conference on Bioinformatics and Biomedicine* (pp. 305-310). IEEE.
55. Martinez, H.P., Bengio, Y. and Yannakakis, G.N., 2013. Learning deep physiological models of affect. *IEEE Computational intelligence magazine*, *8*(2), pp.20-33.
56. Jirayucharoensak, S., Pan-Ngum, S. and Israsena, P., 2014. EEG-based emotion recognition using deep learning network with principal component-based covariate shift adaptation. *The Scientific World Journal*, *2014*.
57. Tripathi, S., Acharya, S., Sharma, R.D., Mittal, S. and Bhattacharya, S., 2017, February. Using Deep and Convolutional Neural Networks for Accurate Emotion Classification on DEAP Dataset. In *Twenty-Ninth IAAI Conference*.
58. Zheng, W., 2016. Multichannel EEG-based emotion recognition via group sparse canonical correlation analysis. *IEEE Transactions on Cognitive and Developmental Systems*, *9*(3), pp.281-290.
59. Yang, Y., Wu, Q.J., Zheng, W.L. and Lu, B.L., 2017. EEG-based emotion recognition using hierarchical network with subnetwork nodes. *IEEE Transactions on Cognitive and Developmental Systems*, *10*(2), pp.408-419.

60. Wen, Z., Xu, R. and Du, J., 2017, December. A novel convolutional neural network for emotion recognition based on EEG signal. In *2017 International Conference on Security, Pattern Analysis and Cybernetics (SPAC)* (pp. 672-677). IEEE.
61. Mei, H. and Xu, X., 2017, December. EEG-based emotion classification using a convolutional neural network. In *2017 International Conference on Security, Pattern Analysis and Cybernetics (SPAC)* (pp. 130-135). IEEE.
62. Russell, J.A., 1980. A circumplex model of effect. *Journal of personality and social psychology*, *39*(6), p.1161.
63. S. Koelstra, et al. (2012), DEAP dataset. Available at http://www.eecs.qmul.ac.uk/mmv/datasets/deap/ accessed on 15th February 2018.
64. Lang, P.J., Bradley, M.M., & Cuthbert, B.N. (2008). International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-8. University of Florida, Gainesville, FL.
65. Bradley, M. M. & Lang, P. J. (2007). The International Affective Digitized Sounds (2nd Edition; IADS-2): Affective ratings of sounds and instruction manual. Technical report B-3. University of Florida, Gainesville, Fl.
66. Lang et al. (2008) IAPS dataset available at: https://csea.phhp.ufl.edu/Media.html#topmedia
67. Cao, H., Cooper, D.G., Keutmann, M.K., Gur, R.C., Nenkova, A. and Verma, R., 2014. CREMA-D: Crowd-sourced emotional multimodal actors dataset. *IEEE transactions on affective computing*, *5*(4), pp.377-390.
68. Campbell, N., 2001. Building a corpus of natural speech-and tools for the processing of expressive speech. In the *Seventh European Conference on Speech Communication and Technology*.
69. Soleymani, M., Lichtenauer, J., Pun, T. and Pantic, M., 2011. A multimodal database for affect recognition and implicit tagging. *IEEE transactions on affective computing*, *3*(1), pp.42-55.
70. Stibbard, R., 2000. Automated extraction of ToBI annotation data from the Reading/Leeds emotional speech corpus. In *ISCA Tutorial and Research Workshop (ITRW) on Speech and Emotion*.
71. Sneddon, I., McRorie, M., McKeown, G. and Hanratty, J., 2011. The belfast induced natural emotion database. *IEEE Transactions on Affective Computing*, *3*(1), pp.32-41.
72. Alpaydin, E., 2004. Introduction to machine learning (adaptive computation and machine learning series). *The MIT Press Cambridge*.

73. Reed, R. and MarksII, R.J., 1999. *Neural Smithing: supervised learning in feedforward artificial neural networks*. MIT Press.
74. Naser, D.S. and Saha, G., 2013, March. Recognition of emotions induced by music videos using DT-CWPT. In *2013 Indian Conference on Medical Informatics and Telemedicine (ICMIT)* (pp. 53-57). IEEE.
75. Bahari, F. and Janghorbani, A., 2013, December. Eeg-based emotion recognition using recurrence plot analysis and k nearest neighbor classifier. In *2013 20th Iranian Conference on Biomedical Engineering (ICBME)* (pp. 228-233). IEEE.
76. Zhuang, X., Rozgić, V. and Crystal, M., 2014, June. Compact unsupervised EEG response representation for emotion recognition. In *IEEE-EMBS international conference on Biomedical and Health Informatics (BHI)* (pp. 736-739). IEEE.