Emotion Recognition using Machine Learning Model from ECG and GSR

Author
AMNA RAHIM
Reg. Number
00000277185

Supervisor
DR SAJID GUL KHUWAJA
Co-Supervisor
Dr. MUHAMMAD USMAN AKRAM

DEPARTMENT OF COMPUTER & SOFTWARE ENGINEERING
COLLEGE OF ELECTRICAL & MECHANICAL ENGINEERING
NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY
ISLAMABAD

July 2021

Emotion Recognition using Machine Learning Model from ECG and GSR

Author

AMNA RAHIM

Reg. Number

00000277185

A thesis submitted in partial fulfillment of the requirements for the degree of

MS Computer Engineering

Supervisor:

DR SAJID GUL KHUWAJA

Co-Supervisor

Dr. MUHAMMAD USMAN AKRAM

Thesis Supervisor's Signature:

____

DEPARTMENT OF COMPUTER & SOFTWARE ENGINEERING

COLLEGE OF ELECTRICAL & MECHANICAL ENGINEERING

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY,

ISLAMABAD

July 2021

DECLARATION

I certify that this research work titled “Emotion Recognition using Machine Learning Model from ECG and GSR” is my work. The work has not been presented elsewhere for assessment. The material that has been used from other sources has been properly acknowledged/referred to.

Signature of Student

AMNA RAHIM

MS – 18 – CE

PLAGIARISM REPORT (TURNITIN REPORT)

This thesis has been checked for Plagiarism. Turnitin report endorsed by Supervisor is attached.

Signature of Student

AMNA RAHIM

00000277185

Signature of Supervisor

COPYRIGHT STATEMENT

ACKNOWLEDGEMENT

I am thankful to my Creator Allah Subhana-Watala to have guided me throughout this work at every step and for every new thought indeed. Whosoever helped me throughout my thesis, whether my parents or any other individual was your will, so indeed none be worthy of praise but you.

I would also like to express special thanks to my supervisor and my lovely sister Aqsa Rahim for their help throughout my thesis. Finally, I would like to express my gratitude to all the individuals who have rendered valuable assistance to my thesis.

ABSTRACT

Human Emotion Recognition has enabled state of art applications in medical healthcare, surveillance, digital entertainment and various other sectors. Therefore, their prediction remained an interesting aspect in the field of research. It helps in making decisions and in effective communication. Emotions are reflected through numerous ways which includes speech processing, behavioral changes, facial expression and physiological signals. Most of these methods are biased. Limitations may appear in emotion recognition model because of limited number of facial expression, fake emotions or among people with disease where they cannot communicate their emotions. Physiological signals give a better insight of emotion classification. This research presents a self-collection of physiological signals database for emotion classification into 7 classes for 27 subjects using heart and skin signals. We designed a generalized deep learning classification system for emotion detection using ECG and GSR signals to divide signals into 7 different classes. The classification model has been built around extracting several features and employing three different CNN architectures individually: Alexnet, Resnet and Inception. Inception performs best among all other architectures. The training dataset containing 5000 samples each has been used to train all three classifiers. The proposed methodology performs reasonably well for most of the classes achieving around an accuracy of 80% and 79.2% for ECG and GSR signals respectively. This setup is useful for the health monitoring system and also for the investigation purpose where one can easily predict the victim’s emotion.

Key Words: Emotion Classification, Physiological Signals, Deep Learning, Convolutional Neural Network

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1. INTRODUCTION

Emotion is defined as mixture of feelings, thinking patterns, behavioral response and psychophysiological signals. It has an important role in making decisions, effective communication and in extracting useful information from messages in both speech and images. Negative emotions adversely affect an individual's life leaving an impact on quality of life as their long term accumulation leads to clinical depression. It causes transition in different organs of human body.[1]. It is identified by various means which includes speech processing, face expression, behavioral analysis and physiological signals. Most of these emotion recognition methods are biased, such as an individual can easily hide their true emotional state. . Limitations may appear in emotion recognition model because of limited number of facial expression, fake emotions or among people with disease where they cannot communicate their emotions. Physiological signals give a better insight of emotion classification. [2]. As they cannot be faked nor easily affected by the environment and can be observed in real time as well. Moreover, physiological signals offer more information related to emotional state. Therefore, we considered two physiological signals of a subject which includes ECG and GSR from our self-collected database and developed an emotion recognition system.

Emotion is defined as state of mind that arises quickly and comes with changes in physical and physiological state of an individual which are reflected by variations in different organs of body e-g heart, brain, skin etc. Various databases are available online for emotion classification using different techniques. This research proposed an online self-collected database for emotion classification into 7 classes based on ECG and GSR signals and an automated classification model for 7 different types of emotions.

1.1 MOTIVATION

Due to advancement in technology many applications are becoming operational around the globe. This has led to several new discoveries which have helped to provide solutions for multiple problems and enhanced the quality of life. Understanding of emotions enables humans to perform their tasks in a better way and excel in their respective jobs. In addition, recent advances in computation have allowed us to manipulate data on human emotion with ease [3]. The collected dataset and computational expertise can lead to better, fast and accurate solution.

Accurate and swift classification of emotions can significantly cut back on the manual effort required by researchers and can really help the diverse scientific community to achieve their goal.

Emotion Recognition is one of fundamental components of affective computing [4]. It can help physiologists to keep record history of an individual and can be helpful in their speedy recovery. It also has great significance in health monitoring systems especially those in which patients cannot communicate conveniently such as paralytic, stroke, alexithymia and autism patients. An exact human mind and heart reading might still be a dream but the pursuit of understanding and detection of human emotions is the motivation for taking on this research.

1.2 PROBLEM STATEMENT

Swift and accurate prediction of human emotion using physiological signals can contribute in development of many applications enhancing the quality of life and adding ease in the lives of people. Precise and timely emotion recognition is still limited to our knowledge because of the complexity of mutual intersection of physiological and psychological patterns in emotional states.

We intend to develop model using deep learning techniques that precisely classify objects into their respective classes and develop a database for emotion determination.

1.3 AIMS AND OBJECTIVES

Major objectives of the research are as follow:

- To develop a local dataset for emotion classification
- To attain direct access of ECG and GSR signals
- To process signals from sensors using AI and Machine Learning
- Devise an algorithm for classification of GSR and ECG signals into seven classes of emotion
- To facilitate people so they can make their true emotional assessment this can be further utilized in variety of ways.

1.4 STRUCTURE OF THESIS

Main contents of thesis are listed below:

Chapter 2 covers the emotion classification models basics and specifically discusses emotion classification using physiological signals.

Chapter 3 gives review of the literature and the significant work done by researchers in past few years for classification of emotions using Machine learning techniques.

Chapter 4 consists of the proposed methodology in detail.

Chapter 5 includes all the experimental results accompanied by relevant figures.

Chapter 6 concludes the thesis and reveals future scope of this research

CHAPTER 2. EMOTION RECOGNITION SYSTEM

Human Recognition system has numerous applications in multiple fields. It creates an impact on making decisions, predictions in business, health care and academics. Emotion is described as intellectual condition of someone that is generated quickly instead of aware attempt and is accompanied with changes in body. Emotions are extensively divided into two classes: constructive and destructive [5]. Constructive emotions improves quality of life and work performance, at the same time destructive emotions may cause health issues which can lead to depression, anxiety and other mental and emotional disorder. Emotion is immediately related to excessive depth and excessive hedonic content material called and mind. Emotions are biological states variously linked with thinking process, emotions, response of behavior and a level of satisfaction or dislikeness [6].

2.1 Emotion Classification Models

Emotion Classification models are used to divide emotions into different categories using different methods. Emotions are broadly classified using following two approaches [7].

1. Discrete models
2. Dimensional models

2.1.1 Discrete Models:

It considers primary emotions which are identified by all people. These emotions are called discrete because they are individually identifiable using individual's features and biological markers. Paul Ekman described emotions a basic level in six ways which are anger, disgust, worry, happiness, unhappiness and wonder as shown in Figure 1.

Figure 1 Basic types of emotion [8]

2.1.2 Dimensional Models:

Dimensional Framework of emotion advise that a not unusual and interconnected neurophysiological network is in charge for all affective states. They try to conceptualize human feelings with the aid of 3 dimensions contrasting theories of primary emotion, which suggest that distinctive feelings get up from separate neural structures. Numerous dimensional architectures of emotion are build, even though there are only some that stay as the dominant fashions currently well-known via maximum. Plutchik's emotion classification scheme is shown in Figure 2.

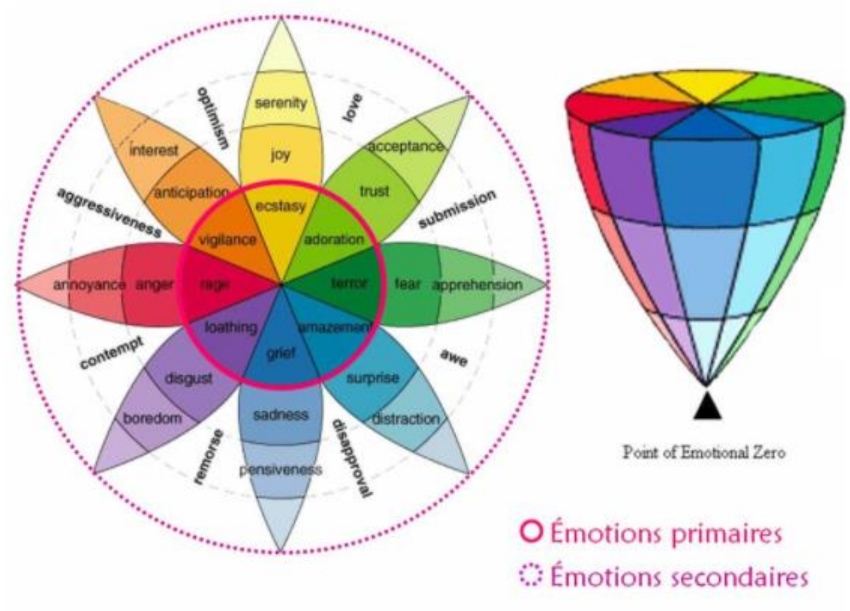


Figure 2 Plutchik Model [9]

Dimensional models encompass three independent and bipolar dimensions which can be valence, arousal and dominance respectively as shown in figure three. Valence is high quality or terrible affectivity, while arousal measures how calming or thrilling the records is. Arousal comes from our reptilian mind. It inspires a fight-or-flight response that aids our survival. Dominance refers to a person's feel of getting a capability to have an effect on the environment.

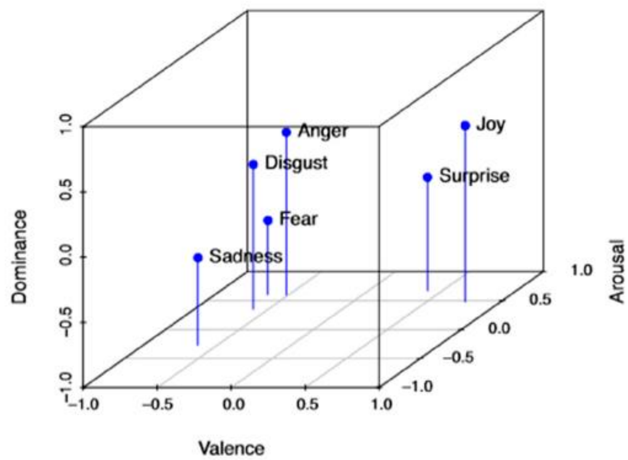


Figure 3 3 Dimensional Emotion Model

Most of time arousal and valence are considered ignoring dominance. Consequently, commonplace framework for dealing with emotional encounter is characterized in a two-dimensional area as shown in Figure 4. Valence ranges from fairly terrible to enormously wonderful this means that pleasure to displeasure, and arousal tiers from calming to thrilling.

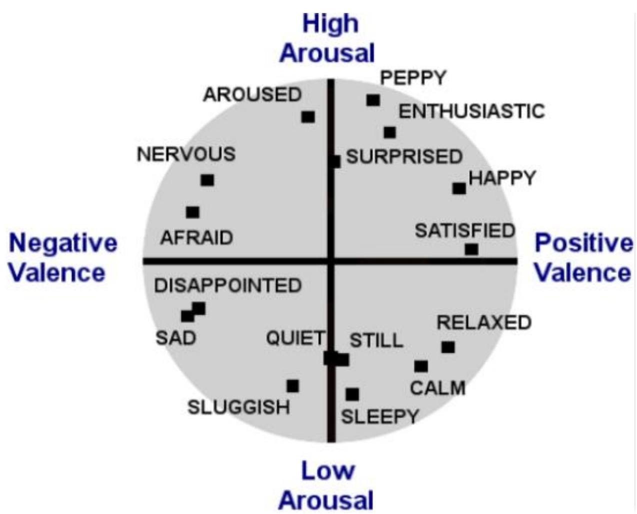


Figure 4 2Dimensional Model of emotion [10]

2.2 Significance of ECM

Emotion identification is currently a relevant area of look at for lots regions along with health care in particular intellectual health monitoring and social safety and so forth. The goal of emotion popularity is to perceive a human’s temporal emotional kingdom automatically. Unavoidably, feelings play a considerable role in making members of the family with others. Affective Computing has affected many components of cutting-edge existence. Any such developments is Emotional intelligence (EI), in which artificial intelligence (AI) techniques are used for emotion identification to serve several beneficial programs, along with patient monitoring, mental remedy, criminal detection, disabled assistance, safety offerings, robotics and conversation tools.

CHAPTER 3. LITERATURE REVIEW

Emotion identification using of physiological alerts is a critical task to enhance the interaction among human and machine. It is been correctly applied to advertising and marketing, healthcare, training, recommender systems, video games, social media and plenty of different applications [13]. Physiological signals had been proved to provide high accuracy in detecting feelings because they reflect spontaneous affect-related statistics that's fresh and do no longer require an additional manipulate or interpretation. maximum of the previous works recognition on less correct sensor-loose approach to restriction intrusiveness and decrease costs or are based on visible assets of data, which include facial expressions or eye-tracking. Such sources have the advantage that they generalize well between customers and allow the construction of inter-problem models. However sensor-based totally structures tend to be extra intrusive and highly-priced after which their applicability are reduced in actual-international state of affairs. Data has always been a counted of difficulty for every subject. So the primary purpose of records mining or deep learning is to boom the efficiency of the information so we can make better choices on its basis. Each dataset has its worth in its respective discipline. Numerous strategies and proposed methodologies are used on exclusive varieties of datasets to improve their accuracy. We talk physiological sign related dataset which could be very crucial because of its relation with human life. By way of growing the performance of the dataset the selection-making of the viewer or physician may be higher. And this higher expertise can enhance the first-rate of life for lots human beings. The aim of our look at is likewise to boom the performance of the dataset so the decision making primarily based on data can be better to get the most advantage. The literature overview is mentioned beneath about the different supervised learning strategies and on CNN models and the effectiveness of the implemented version at the same time as enhancing the accuracy of the dataset. As we recognize CNN models are very powerful inside the category of the photos and additionally in the course of the extraction of the important capabilities from the input images, so our motive is to apply this capability of the CNN model for the improvement of the accuracy of the supervised datasets. It will proved a critical asset for the complex datasets having much less accuracy whilst classification.

3.1 TECHNOLOGIES FOR HUMAN EMOTION RECOGNITION

Emotion is a complicated condition that mixes emotions, mind, and behavior and is human being's psychophysiological reactions to inner or outside stimuli. It plays a critical role in people's decision-making, belief and conversation. Affective computing has an extensive variety of programs. It is observed via certain physiological adjustments, together with growth in heartbeat, respiratory, temperature, pores and skin conductance and muscular anxiety in the frame. Emotion reputation is an interdisciplinary place, and it achieves an enormous attention of the researchers within the past few years [14]. Human emotions may be identified by way of exclusive manner such as facial features, speech, conduct, or physiological alerts. However, the primary three strategies of emotion reputation are by some means biased. Moreover, difficulties and barriers may also arise in standard emotion recognition software due to limited variety of facial features triggers, dissembling of emotions or among human beings with alexithymia.

3.1.1 Facial Expression

Facial expressions are vital in facilitating human conversation and interactions. Also, they are used as a vital component in behavioral research and in scientific rehabilitation. Facial picture primarily based totally temper detection strategies may also offer a quick and realistic method for non-invasive temper detection. Ekman and Friesen advanced the facial motion coding system (FACS) to degree the facial behavior. The FACS codes distinct facial actions into Action Units (AU) primarily based totally at the underlying muscular interest that produces short-term adjustments with inside the facial expression. An expression is in addition identified through effectively figuring out the motion unit or mixture of motion units associated with a specific expression.

3.1.2 Speech Signals

Speech is one of the most common modes of communications among people. Humans have specific emotions via written or spoken languages. Speech emotion has an important role to describe feeling of an individual. If one desires to recognize which means of an utterance, it must be in right emotion in any other case semantically utterance will cross inside the incorrect route and give incorrect results. Whilst people can efficaciously perform this assignment as a natural a part of speech verbal exchange, the potential to behavior it routinely using programmable modules remains an ongoing issue of research. At present, speech emotion classification is an emerging crossing discipline of synthetic intelligence and artificial psychology; except, it became a warm research subject matter of sign processing and sample popularity. The research was extensively applied in HI interaction, interactive teaching, entertainment, safety fields, and so on.

3.1.3 Physiological Signals

Emotion recognition is a promising area of studies in a discipline of human-computer interplay. It is viable to understand feelings using facial expression, audio alerts, frame poses, gestures and so on. However physiological signals are very useful on this field because they're spontaneous and not controllable [15]. Conventional research, which has made massive achievements, is based totally at the recording statistical analysis of physiological alerts from ANS. A few researchers were doing their exceptional to develop wearable devices.

To develop a sensible reputation algorithm, the physiological signals which we will choose are very crucial in addition to very limited at the same time. Despite the fact that electroencephalogram (EEG), facial electromyogram and blood stress can be useful for the research, the attachment of electrodes to the scalp or face seems no longer to be tolerable for practical use, as a result they are ignored of consideration. Other physiological signals encompass: electrocardiogram (ECG), galvanic pores and skin response (GSR), pores and skin temperature (SKT), skin conductance (SC) and coronary heart fee (HR) [16]. ECG is a critical and effective indicator of emotion states. Galvanic skin response (GSR), skin conductance (SC), every now and then additionally called the Electrodermal hobby (EDA), is every other crucial signal to symbolize the interest of the ANS. It characterizes modifications within the electrical residences of the pores and skin because of the interest of sweat glands and is physically interpreted as conductance. ECG is dually controlled by way of the sympathetic and parasympathetic branch of the ANS which could act independently [17]. in addition, time-domain features, inclusive of imply and widespread deviation of the HRV (coronary heart rate variability) time series have also been taken into consideration to be sizable for the exploration of autonomic anxious machine in lots of preceding studies for cardiac function evaluation and psychophysiological research, they had been regularly used as capabilities [18]. Considering the fact that ECG signal can be acquired notably without problems and both time-area functions of

HRV and HR can be computed from it [19], ECG signal seems to be very crucial on this take a look at. Further, breathing could be very important in emotion studies. Particular emotion expressions, including crying, laughing, or shouting, have precise respiration signatures. A detailed quantification of volume, timing and shape parameters in the breathing pattern waveform can map into special emotional states along the size of calm-pleasure, rest-tenseness, and active vs. passive coping. Preliminary proof shows that breathing parameters additionally map into the 'affective area' dimensions of valence and arousal [20].

3.2 RELATED WORK

Harsh Dabas et al. [21] proposed a three-D version for classifying feelings of a person by usage of EEG alerts. The emotional states are categorized on the idea of numerous parameters which include arousal; valence, dominance, and liking. Emotions are studied and analyzed the usage of Database for Emotion Analysis the usage of Physiological Signals (DEAP) Dataset. Proposed method is carried out on Machine Learning Algorithms inclusive of Naive Bayes and Support Vector Machine (SVM) giving an accuracy of 78.06% and 58% respectively.

Yuling Luo et al. [22] proposed an answer that acknowledges emotional modifications and arousal valence, dominance and liking stages running on Electroencephalograph (EEG), the usage of spiking neural networks (SNNs).For DEAP dataset the emotion states of arousal, valence, dominance and liking may be categorized with accuracies of 74%, 78%, 80% and 86.27% the usage of SNN with variance information processing approach and an average accuracy of 96.67% is accomplished for the SEED dataset.

Nagarajan Ganapathy et al. [23] proposed Convolutional neural network (CNN) which makes usage of Electrodermal Activity (EDA) alerts and time-frequency capabilities. Thirty 8 time-frequency capabilities are extracted from phasic factor of EDA. Representative and key capabilities are discovered the usage of CNN to differentiate numerous feelings. Achieved F-degree is 79.30% and international accuracy alongside arousal and valence dimensions is 71.41% for classifying exclusive emotional states.

Lim Jia Zheng et al. [24] develop a completely unique technique for emotion reputation the usage of eye-tracker, scholar diameter as simplest function for classifying feelings into 4 classes. The stimuli used on this test are 360° movies supplied the usage of a Virtual Reality (VR) headset. Three classifiers had been used for the emotion class that are Support Vector Machine

(SVM), k-Nearest Neighbor (KNN), and Random Forest (RF). SVM gave great overall performance for class mission at a median accuracy of 57.05%, that's extra than two times the accuracy of a random classifier.

SeungJun Oh et al. [25] proposed a most reliable emotion reputation approach that makes usage of self-series of Respiration (RSP) and coronary heart fee variability (HRV) alongside their derived parameters. A newly designed deep-getting to know version primarily based totally on a convolutional neural network (CNN) became followed for detecting the identity accuracy of person feelings. Additionally, the sign aggregate of the received parameters became proposed to acquire excessive class accuracy of 94%. Furthermore, a dominant component influencing the accuracy became observed through evaluating the relatedness of the parameters, imparting a foundation for assisting the consequences of emotion class.

Ayan Seal et al. [26] supplied an Electroencephalogram (EEG) database and brought discrete wavelet transform (DWT) for Emotion classification into 4 classes with the usage of EEG alerts. The ELM set of rules is used for channel choice observed through sub band choice with 94.72% accuracy.

Yang Liu et al. [27] proposed a brand new physiological dataset PAFEW for emotion class into 7 classes of emotion the usage of Electrodermal activity (EDA).The dataset is gathered through E4 wristband and Eye Tribe with Acted Facial Expressions with inside the Wild (AFEW) used as stimulus.6 capabilities had been extracted from every information collection similar to every video because the center of the community. Leave-one-observer-out became hired on this class mission.2 networks had been defined. One became a 3 layer primary community which achieves an accuracy of 42.08%. It additionally educated classifiers to categories the information through arousal and valence, ensuing in an accuracy of 68.66% and 72.72% respectively.

Tongshuai Song et al. [28] proposed a model based totally on Convolution Neural Network (CNN) to system EEG and peripheral alerts respectively for emotion class into 3 classes. Taking the extraction of conventional capabilities into consideration, the primary version is primarily based totally on a 2 -dimensional Convolutional Neural Network (2D-CNN) with the usage of authentic EEG information, wherein its kernel is one-dimensional to extract identical varieties of capabilities for each channel. In the second one version, it carried out one-dimensional Convolution Neural Network (1D-CNN) to each channel of peripheral alerts after which concatenates consequences for class. MAHNOB-HCI database is used for assessment of each fashion. The class accuracies in arousal and valence size of the 2 fashions the usage of CNN are 61.5%, 58.01% and 58%, 56.28% respectively.

Deger Ayata et al. [29] completed emotion reputation from galvanic alerts the usage of time area and wavelet primarily based totally capabilities. Feature extraction has been completed with numerous function set attributes. Valence and arousal were labeled and courting among physiological alerts and arousal and valence has been studied with the usage of Random Forest getting to know set of rules accomplishing 71.53% and 71.04% accuracy for arousal and valence respectively.

Maryam Memar et al. [30] supplied a unique approach of strain stage class with the usage of physiological alerts at some point of the real-global riding mission. Proposed framework has a look at evaluates the feasibility and effectiveness of the evaluation of variance (ANOVA) classifier learner at the Galvanic Skin Response (GSR) signal. Data gathered through writer and his co-people is to be had with inside the PHYSIONET database. Three stages of strain had been taken into consideration and impartial capabilities inclusive of growing time and amplitude had been extracted. These capabilities are extracted from foot and hand GSR alerts in 3 exclusive situations for the sake of schooling. The end result suggests that the foot amplitude function of the GSR sign simplest is a dependable supply of strain class with an accuracy of 95.83% through making use of the ANOVA technique.

Sali Issa et al. [31] worked on an user-impartial emotion class approach that classifies 4 feelings with the usage of Electroencephalogram (EEG) alerts with the usage of public DEAP and MAHNOB-HCI databases. Just one EEG electrode channel is chosen for the function extraction system. Continuous wavelet transform (CWT) is then applied to extract the proposed gray-scale photograph (GSI) function which describes the EEG mind activation in each time and frequency domains. Finally, the brand new BLS is built for the emotion class system. The test consequences display that the proposed framwwork produces a sturdy device with excessive accuracy of about 93.1% for the DEAP database and a median accuracy of about 94.four% for MAHNOB-HCI database.

Jinxiang Liao et al. [32] proposed a convolutional recurrent neural community (RCNN) primarily based totally approach for multi-modal physiological signal emotion reputation mission. The approach used convolutional neural community to analyze the spatial representations of multi-channel EEG alerts and the Long Short-time period Memory community to analyze the temporal representations of peripheral physiological alerts which include EOG, EMG, GSR, RSP, BVP and TMP. The representations are blended for emotion reputation and class. In the 2 emotion dimensions of Arousal and Valence, experiments carried out at the open supply dataset DEAP suggests that, this approach attain 89.68% and 89.19% common accuracy with inside the EEG emotion class, 63.06% and 62.41% common accuracy with inside the peripheral physiological sign emotion class,93.06% and 91.95% common accuracy with inside the blended function emotion class.

R. Chinmayi et al. [33] supplied a version framework the usage of each facial and vocal capability for emotion class into six classes which had been surprise, worry, disgust, anger, happiness and disappointment. For pictures database FER2013, 25838 samples had been used for schooling and ninety samples of Amrita Emote database (ADB) for trying out. The speech database includes 4 exclusive datasets, with a complete of 20,000 examples. 3/4 of this data is used for training purpose and 1/four of is used is for testing. The function Extraction approach used for speech information is Mel Frequency Cepstral Coefficients (MFCC).The output of the function extraction acquired is function coefficient vectors of a cepstrum with inside the Mel scale and those coefficient vectors are then used for class. By dividing the information accordingly, the device is educated after which examined for each the information sorts the usage of RNN classifier for speech and CNN for the images. The train accuracy and train accuracy acquired became 91.57% and 78.12% respectively for speech. The acquired train accuracy and test accuracy became 97.22% and 59.06% respectively for photograph.

Arturo Martínez Rodrigo et al. [34] brought a wearable hardware and software program primarily based totally device for emotion detection into seven classes which incorporates neutral, tenderness, entertainment, anger, disgust, worry and disappointment from self-gathered information on 39 topics the usage of a movie temper induction procedure. Seventeen capabilities are calculated on pores and skin conductance reaction and coronary heart variability information, grouped into 5 statistical, 4 temporal and 8 morphological capabilities. Then those capabilities are used to run emotion class version thinking about aid vector machines, selection trees, and quadratic discriminant evaluation. The consequences provide an international accuracy of 82% in bad emotion class. For superb feelings accuracy for entertainment and tenderness is 92% and 66% respectively. A summary of all the research work discussed above is presented in Table 1.

Table 1 Comparison of already exiting work

Year	**Author**	**Feature**	**Technique**	**Emotions**	**Accuracy**
2018	Harsh Dabas et al. [21]	EEG	Naive Bayes SVM	4	78.06% 58.90%
2020	Yuling Luo et al. [22]	EEG	SNN	4	96.67%
2020	Nagarajan Ganapathy et al. [23]	EDA	CNN	2	71.41%
2020	Lim Jia Zheng et al. [24]	Pupil diameter	SVM	4	57.05%
2020	SeungJun Oh et al. [25]	RSP,HRV	CNN	6	94%
2020	Ayan Seal et al. [26]	EEG	ELM	4	94.72%
2020	Yang Liu et al. [27]	EDA	CNN	7	42.08%
2020	Tongshuai Song et al. [28]	EEG,PERI	CNN	3	61.5% 58.01%
2020	Deger Ayata et al. [29]	GSR	RF	2	71.53% 71.04%
2021	Maryam Memar et al. [30]	GSR	ANOVA	3	95.83%
2020	Sali Issa et al. [31]	EEG	BLS	4	94.4%
2020	Jinxiang Liao et al. [32]	EEG,PERI	RNN	2	93.06% 91.95%
2019	R. Chinmayi et al. [33]	Facial and Vocal	RNN(speech) CNN(image)	6	78.2% 59.06%
2019	Arturo Martínez Rodrigo et al. [34]	SCR,HRV	SVM	7	82%

CHAPTER 4. METHODOLOGY

This thesis presents a method for automatic classification of physiological signals to identify different emotional states observed by an individual. The proposed methodology consists of two main phases i.e. dataset collection and emotion classification. Figure 7 shows the constituent steps of the proposed methodology which presents the flow of the steps which we have followed. First of all, we collected the data to validate our proposed methodology. Data has more importance because it depends on what we are trying to do, in which field and about what problem. It depends on the aim of the study about what we are going to solve. After the dataset collection, there is some preprocessing which is compulsory to do to bring the data into an understandable form. Because data is gathered from different resources and can be in different formats and could have many anomalies and blunders. So motivation behind preprocessing data is to purify it from these anomalies and blunders. After the pre-processing, we need to collect those features which are important and contribute towards the output label. Feature importance depends on its contribution to the output label. So we collected those features which have a high correlation to the output label. After that, the last step of our flow is to classify the extracted features using different supervised algorithms, and finally, we will compare their accuracies.

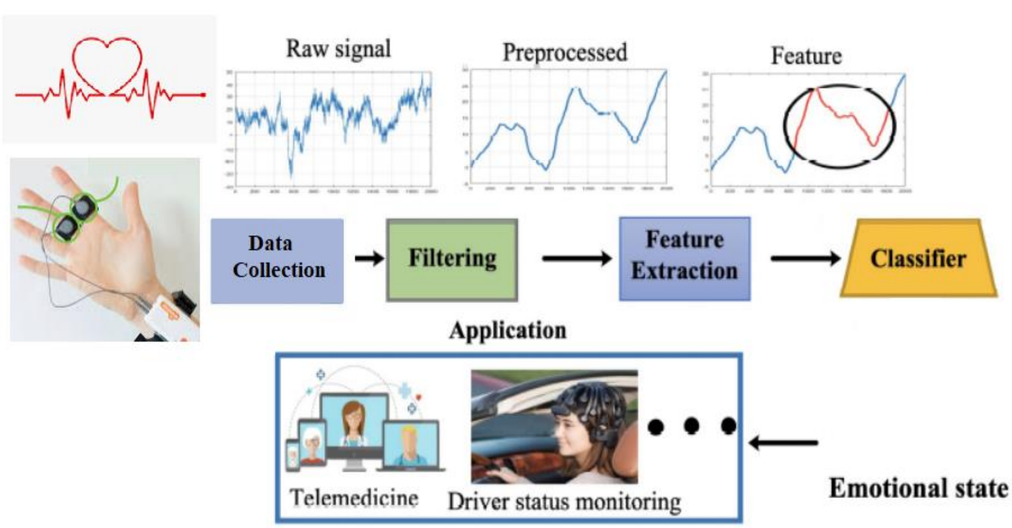


Figure 5 Proposed Methodology

4.1 DATASET COLLECTION

The first and the main step of the methodology is the selection of data on which we are going to implement our proposed methodology. The selection of the dataset depends on which field you are interested to do work or which specific area needs improvement. So our dataset collection aims to improve the accuracy of the Emotion Classification Model (ECM) which is helpful in enhancing quality of life for individuals having emotional impairments. A large number of people are affected by emotion disorders in the whole world due to different causes and it negatively impacts their mental health and lives. According to an estimation round 1 billion people suffer from depression worldwide. The population with a mental health is shown in Figure 8. Despite being vital to average well-being and bodily fitness, diagnoses and furthermore, remedy or support, stay a great deal decrease than this estimate. The real incidence of intellectual fitness problems globally stays poorly understood. Diagnosis facts might no longer carry us near the real figure; mental health is generally underreported, and under-diagnosed. If counting on intellectual fitness diagnoses on own, incidence figures might be probably to mirror healthcare spending in place of giving a consultant attitude on variations among nations; high-profits nations might probably display substantially better incidence due to greater diagnoses. Mental fitness can comprise a selection of various however every now and then related problems including melancholy, tension, bipolar, ingesting problems, schizophrenia and alcohol and drug use problems. Many recognized organizations including World Health Organization (WHO) frequently do now no longer degree worldwide incidence throughout character intellectual fitness problems past melancholy and substance use. Around 1-in-7 humans globally (11-18 percentage) have one or greater intellectual problems. Globally, this indicates round 1000000000 humans in 2017 skilled one. The biggest range of humans had a tension disorder, envisioned at round four percentage of the population.

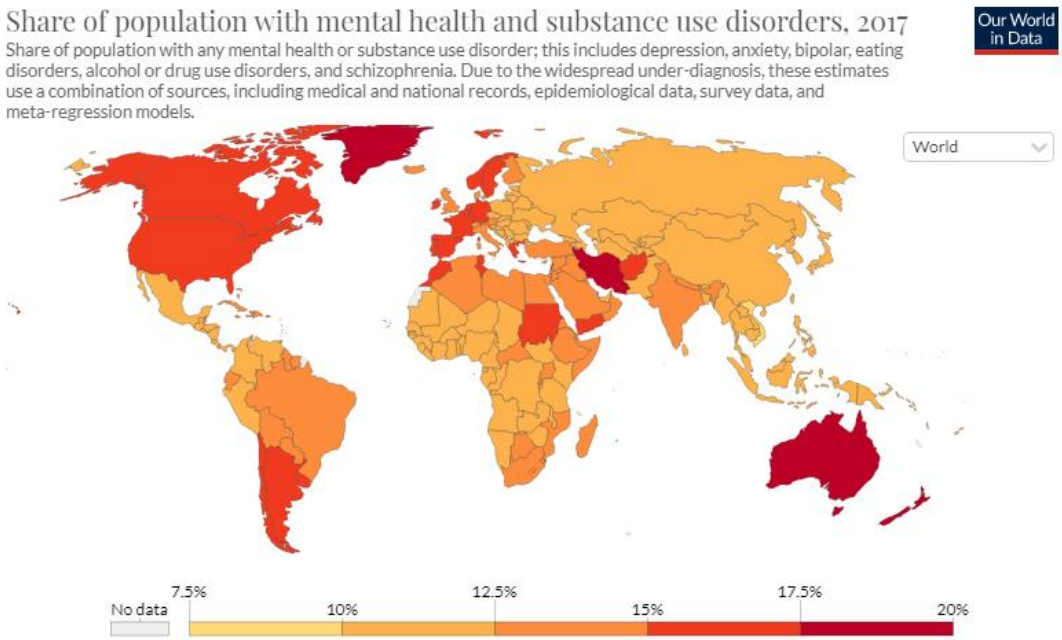


Figure 6 Population with mental health issues [35]

This thesis research focuses on development of physiological signals based dataset in which emotions are labelled by subjects both explicitly and implicitly. Values for valence, arousal and dominance are simultaneously reported. Meanwhile, two good qualities physiological signals of 24 bits ADC for GSR and ECG signals are collected over total 27 subjects, which include 16 male and 11 female by watching videos for respective emotions. The videos used for inducing an emotion are chosen based on literature. Moreover, dataset is also important for the validation of the proposed methodology. Our proposed methodology is used to increase the efficiency of the dataset. So the improvement in the dataset can help its stakeholders and it is also important in its field. The improvement in the efficiency of that dataset can help in accurate emotion classification. This dataset has eighteen attributes including output labels.

So the Aim of our study to improve the accuracy of the supervised dataset and it will work for almost every supervised dataset with slight changes. We tried our best to improve the efficiency of the dataset so it would be helping in improving quality of life.

4.1.1 SUBJECTS

Twenty-seven volunteers, who encompass 16 males, range of age 8–25 years and eleven ladies, with age between eleven–25 years from one-of-a-kind cultural backgrounds participated inside the data series test. Those participants had been recruited from an organization-extensive name for volunteers. Upon registering for the test, an e-mail containing well known statistics and commands for the test become sent to the participants. On this e-mail, they were asked to put on loose apparel to facilitate the placement of sensors. Furthermore statistics at the gender, age-organization, scientific history and many others of the contributors is to be had in the metadata to the dataset.

4.1.2 STIMULANT

Within the major experiment, the aim was to detect elicit happy, sad, neutral, surprise, disgust, anger and fear emotional states through video-stimuli. For this purpose, 42 videos few of them were previously used by different research have been shortlisted. The emotional content material of these motion pictures turned into then tested, where 10 contributors without a overlap with the contributors of this look at viewed and rated these motion pictures remotely with the usage of a web-primarily based interface. based totally on the effects of this pre-take a look at and in addition inner evaluate, twenty-one videos have been decided on for the principle test, such that there have been three films for the emotional condition that we wanted to elicit.

4.1.3 EXPERIMENTAL DESIGN

Experiments have been carried out to categories seven emotions and evaluate the classifiers. All individuals watched and annotated the one-of-a-kind video-stimuli used for the test. It became an open label experiment finished from November to December 2020 at college of electrical and Mechanical Engineering (CEME), national college of technology and era (NUST), Islamabad, Pakistan. The experiment turned into conducted in a controlled surroundings wherein a screen able to displaying video clips with seven awesome feelings and sensors capable of acquiring physiological alerts have been set up. To stimulate the members, movies expressing emotional states of satisfied, unhappy, impartial, surprise, disgust, anger and worry had been played for 2 minute in that order.

Earlier than the start of the test, contributors were sufficiently briefed approximately the experiment and its facet outcomes. While seated in the front of the display, the participant placed a Shimmer GSR sensor on hands and a Shimmer ECG sensor at the chest. Whilst the experiment commenced and it became determined that physiological balance changed into ensured, we measured the physiological sign in the neutral nation for 10 seconds for use as a different emotion for the manipulate institution. The player watched the happiness video for two mins and maintained a secure condition for 2 minutes thereafter. This two-minute period also allowed the contributors to relaxation in-between annotating the motion pictures. Whilst the cozy country ended, the participant watched the unhappiness video and the final videos within the order indicated in Figure 9. Similarly, markers had been displayed on the signal at the begining of every video to facilitate information processing.

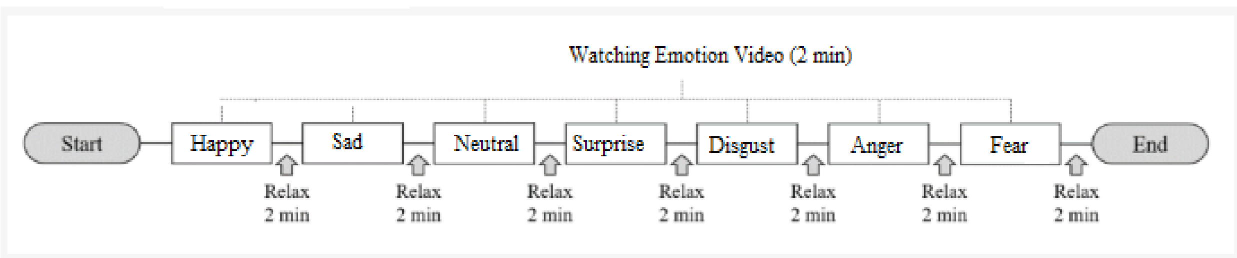


Figure 7 Video sequence for data collection

Shimmer3 ECG and Shimmer3 GSR units are used to capture ECG and GSR signals respectively. Each beat of coronary heart is caused by means of an electrical impulse commonly generated from special cells in the top chamber of the heart. An electrocardiogram notes the timing and energy of these signals as they tour through the coronary heart. Sampling frequency can be changed in keeping with the requirements. ECG signals are sampled at a frequency of two fifty six Hertz and GSR signals are sampled at a frequency of one twenty eight Hertz and overall 5000 samples are taken into consideration for every signal. Duration for saving records is one hundred fifty sec. We have got used five lead shimmer3 ECG unit. The electrodes are commonly made of Ag/AgCl. There is no need to apply gel before placing electrodes on the chest. Length of wire can affect the quality of signal. It is recommended to use short wires so that it can minimize the interface experienced in the signal. The placement of ECG electrodes over the chest is important as shown in Figure 10.

Figure 8 ECG and GSR electrodes placement

Participant is kept in a neutral condition for 5 seconds before and after watching the video to remove any biases in his emotion. A five second offset is added before and after showing video to participant so neutral state can be maintained. Therefore, baseline is added in all signals.

The Shimmer MATLABTM Instrument Driver allows for real-time streaming of data from a Shimmer device into MATLABTM. Shimmer ECG and GSR sensor units are paired and connected with the PC over the Bluetooth. Each sensor has its own Com Port number through which their signals are shared over the Bluetooth to PC and saved as mat files. Noise was avoided at time of signal acquisition by removing mobile and all other electronic devices around the subject and subject was asked to stay in the static position without involving unnecessary movements.

4.1.4 EXPERIMENTAL PROTOCOL

Data Collection for emotion classification task is held in two steps: GSR and ECG signals acquisition and self-assessment. The goal is to obtain a standardized dataset. ECG and GSR signals are captured congenitally, they have got repeating sequence styles and a few classification methodologies make it possible to categories the indicators in actual-time. Then three-D arousal, valence and dominance version changed into furnished and any doubts about the same have been clarified. Distinct information turned into then furnished on the annotation system. Following this, physiological sensors have been connected and the player changed into seated dealing with a 50” tv display screen as shown in Figure 11. The signals have been captured using Shimmer3 ECG unit and Shimmer3 GSR unit from subjects with no discernible disability.

Figure 9 Subject watching video during signal acquisition

After watching a video subject is asked to fill a questionnaire\self-assessment form in which emotional state is assessed with respect to 3 dimensional (arousal, valence and dominance) model and seven basic emotions model. Arousal means how calming or exciting the information is, valence is positive or negative affectivity and dominance is action defined on an emotion. Subject was asked to make proper selection by developing ratings for overlapping emotional states. Participant assessment is carried out in a Google sheet self-assessment form percentagewise after watching each video for all 7 videos on targeted emotions. The Self-Assessment Manikin (SAM) is used for 3 dimensional emotional assessments of subjects in questionnaires. It is a non-verbal pictorial assessment methodology that is quick, inexpensive and easy and quantifies emotional state accordingly. These manikins were mapped with numbers. Ten levels from 0 to 9 are defined for each dimension which represents increase in intensity as shown in the questionnaire attached in Appendix. Similarly for emotion wise categorical assessment percentage is defined against each emotion. Subject’s age, gender and health condition are also noted. Each video is 3 minutes long.

4.2 FEATURES

Feature selection contributes a vital role in improving performance. In feature selection, those features are chosen which are highly correlated to the output label. In our proposed methodology we are working on ECG and GSR signals which are explained below in detail.

4.2.1 ECG

ECG represents the electrical activity of heart. It can be used for several purpose as it is one of important key performance index for human health. ECG indicators consist of 3 most important waves as shown in Figure12.

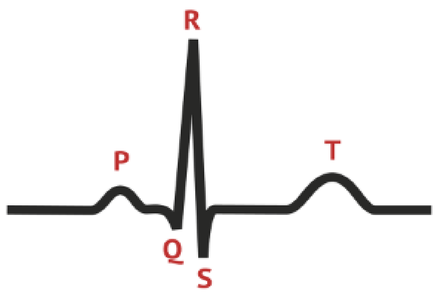


Figure 10 ECG waveform

4.2.1.1 ECG Signal Acquisition\ECG Equipment

Shimmer3 ECG Unit shown in Figure 13 is used to acquire ECG alerts from the subject thru using electrodes placed at the challenge's frame. Shimmer's sensing generation permits for simple signal capture and transmission of biophysical and kinematic records in real-time which may be logged or streamed thru Bluetooth to every other Bluetooth enabled device like a laptop or cell tool. Each beat of heart is triggered by means of an electrical impulse commonly generated from unique cells inside the upper chamber of coronary heart referred to as pacemaker cells.ECG signals are recorded at 256 Hz sampling rate and total 5000 samples are considered for each signal with duration of capturing data is 150 seconds. In acquiring ECG signals five electrodes were used and are placed at fixed positions on subject's chest as shown in Figure 14.

Figure 11 Shimmer ECG unit [35]

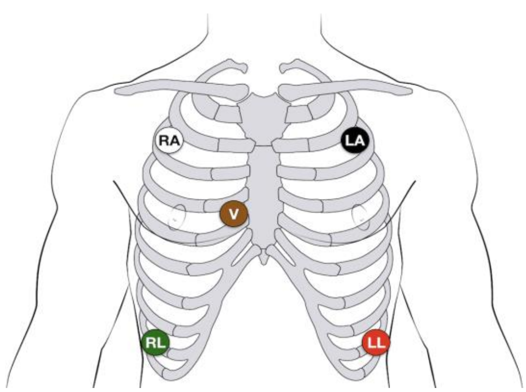


Figure 12 Placement of ECG electrodes [37]

4.2.2 GSR

Conductance values from the GSR sensor will rely upon the quantity that a person sweats; the extra the individual sweats, the better the conductance. It is in general getting used to classify anger and worry. The GSR sign is also known as Electro-dermal Activity (EDA), Electrodermograph (EDG) and Electrodermal response (EDR) consists of both tonic and phasic additives. It use electrodes for measuring conductance levels in the skin. One electrode is placed on index finger and other on the ear using an ear clip on the respective side. As a result, the electric levels of the skin alternate. GSR triumph over 3 limitations together with the difficulty of acquiring continuous measurement, subjects' incapacity to appropriately report their emotions. One major disadvantage of GSR is that it could be laid low with outside factors together with temperature. a way to counter this is thru the encouraged size and calibrations.

4.2.2.1 GSR Signal Acquisition\GSR Equipment

As a precaution it is important to observe that the GSR electrodes aren't to be applied to the situation's body while unit is in a USB dock or multi-charger. The choice of sampling frequency is completely as much as the user wants, GSR signals are sampled at a frequency of one twenty eight Hertz and total 5000 samples are considered for every signal. The length for recording signal is hundred and fifty seconds. One electrode has to be placed at the palmar floor of the medial phalange and the alternative on the palmar floor of the distal phalange and ear clip on that aspect. Alternatively, the electrodes may be placed as shown in figure 15. It's far encouraged to use snap connector Ag/AgCl electrodes. The floor area of the electrodes must be stored to a minimum; 1 cm2 are perfect.

Figure 13 Placement of Shimmer GSR unit on finger [38]

4.3 FEATURE PRE-PROCESSING

All signals acquired from the experiment are being recorded using real time period and showed visually. We selected the statistics appropriate for the examining and splitting them. Noisy indicators had been discarded due to signal instability or other motives. The signal to be analyzed was selected and changed into stored as time-series information. Subsequently, we split the records in line with the markers, as recorded in the test and extracted the information similar to the seven emotions. Records processing turned into completed with the aid of labeling every of the seven extracted alerts with their corresponding emotion.*The overall procedure of data processing is represented in Figure 14.

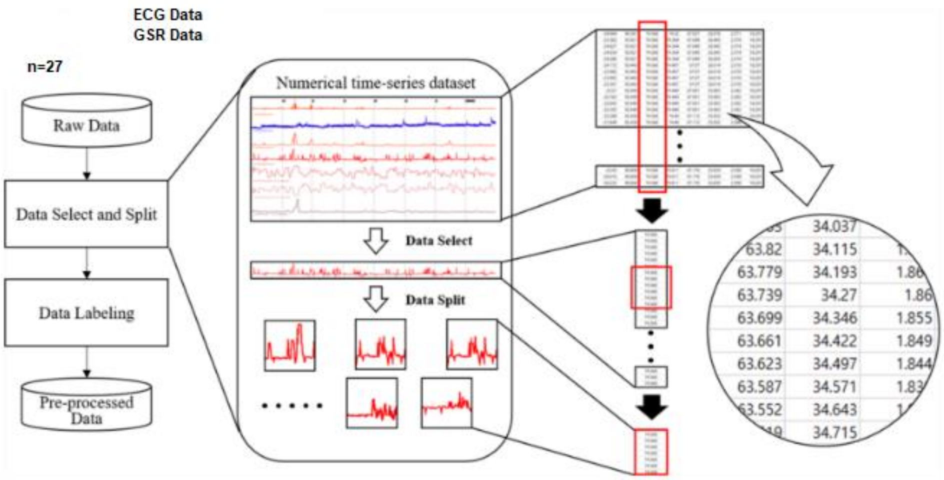


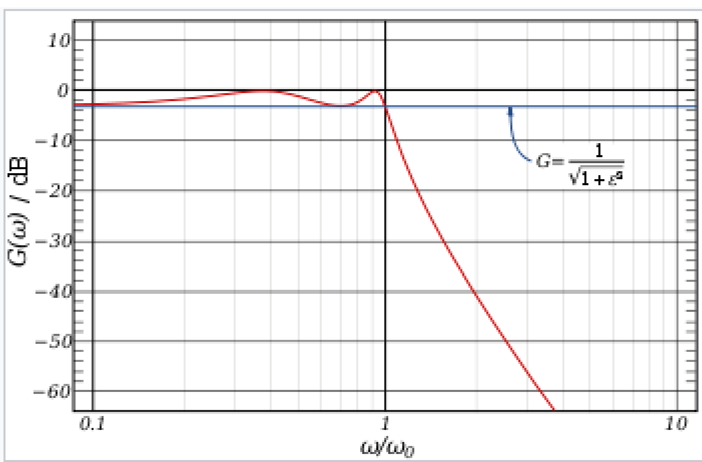
Figure 14 Data preprocessing procedure

The ECG data is pre-filtered with a second order Chebyshev low pass filter (LPF) having a corner frequency slightly smaller as compared to the NYQuist rate and a second order Chebyshev high pass filter (HPF) having a corner frequency of 0.5Hz as shown in Figure 19.

Chebyshev Filter:

The pass band ripple of Chebyshev is greater than 1.

$$G_n(\omega) = |H_n(j\omega)| = \frac{1}{\sqrt{1+\varepsilon^2 T_n^2\left(\frac{\omega}{\omega_0}\right)}}$$



The frequency response of a fourth-order type I Chebyshev low-pass filter with $\varepsilon = 1$

Figure 14 Chebyshev filter

4.4 SENSORS

A sensor is a device which is used to detect an event or change in its environment and send that acquired data to any computer processor or to a cloud system. Sensors can also send data to processors through the Bluetooth connectivity. We have used Shimmer sensors for data acquisition. Shimmer is a Dublin based wearable brand, wearable wireless sensors which can be used to display health, athletic overall performance and biophysical responses. Shimmer's sensing technology lets in for easy seize and transmission of biophysical and kinematic records in actual-time which can be logged or streamed via Bluetooth to any other Bluetooth enabled tool like a computer or cell device. We have used Shimmer ECG and GSR unit for data collection purpose.

4.4.2 Shimmer3 ECG unit

Shimmer ECG unit records electrical activity of heart at rest or at moving positions using five electrodes as shown in Figure 20. The collected data can be used for numerous purposes.

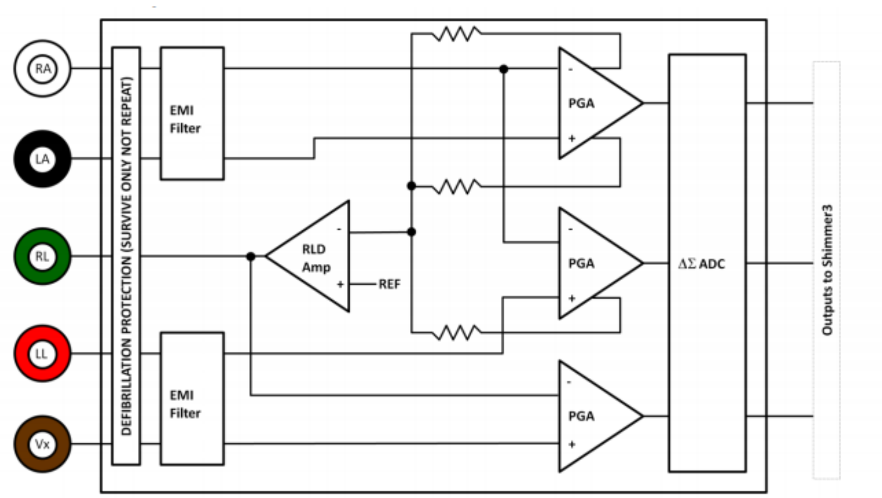


Figure 15 5 ECG Electrodes

4.4.2 Shimmer3 GSR+ unit

.4.4.3 SHIMMER MATLAB TOOL KIT

The Shimmer MATLABTM Instrument Driver allows for real-time streaming of data from a Shimmer device into MATLABTM. It is designed to work with MATLABTM 2008a.

Pre-Requisites

The pre-requisites for Shimmer MATLABTM Instrument Driver are listed below:

1. The most recent version of the Shimmer User Manual, which holds information on programming firmware onto Shimmer and information on how to pair a Shimmer. The Shimmer User Manual is available for download from the documentation section of the Shimmer website.

2. MATLABTM, 2008a (version 7.6) or later installed on PC.

3. A Shimmer2/2r with the latest BtStream firmware or a Shimmer3 programmed with the latest version of LogAndStream firmware.

4. The Shimmer needs to be paired with the PC over Bluetooth.

5. Use the latest version of the Instrument Driver and accompanying User guide.

6. Ensure that Real-term is installed on your PC.

Real term

Download and install an application called Realterm Serial Terminal, which acts as the communication link between the Shimmer device and MATLABTM. Realterm version 2.0.0.57 was used in the development of this driver.

Bluetooth Connectivity:

In order to use a Shimmer3 ECG and GSR devices, the PC should be connected with sensors via Bluetooth. Once Shimmer3 ECG and GSR devices have been paired with computer, they will be automatically connected every time. Each sensor has its own Com Port number through which their signals are shared via Bluetooth to the PC. We can use Bluetooth connectivity to connect multiple shimmers with the same computer simultaneously. For this purpose we need to get comport of the shimmers as shown in Figure 21. Go to the Bluetooth settings to get comport of the connected shimmers. There will be two directions for every port i-e Incoming and outgoing. We will select comport with outdoing direction for the connectivity

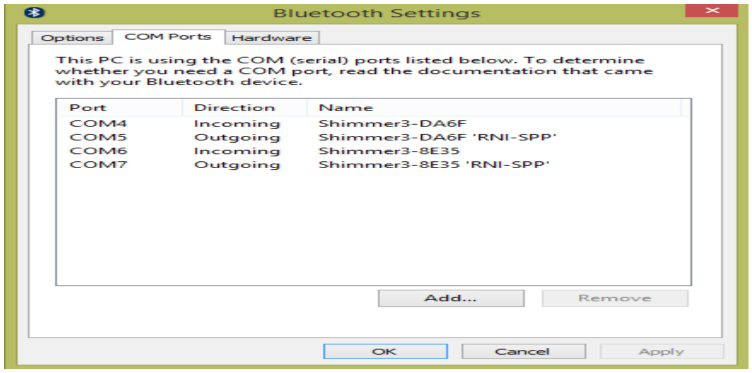


Figure 16 Shimmer comports in PC

Using the Instrument Driver

Start MATLABTM and set the current working directory to the Shimmer MATLABTM Instrument Driver folder.

Shimmer State Machine

The Shimmer essentially behaves as a state machine. Figure 22 illustrates the behavior of the state machine.

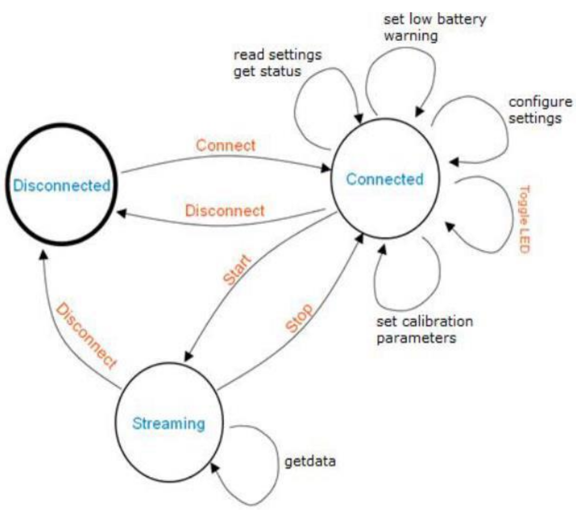


Figure 17 Shimmer State Machine

4.5 DATA ACQUISITION

The collection of real time signals from the user is called signal acquisition. Experimental setup for data collection is shown in Figure 23. Videos for different emotions are shown on a large TV screen. The physiological data was acquired over Bluetooth using the Shimmer sensors. The collected data is divided into three files which includes: 2 files for physiological data and 1 file for self –assessment by participant. Same procedure is repeated for all other subjects.

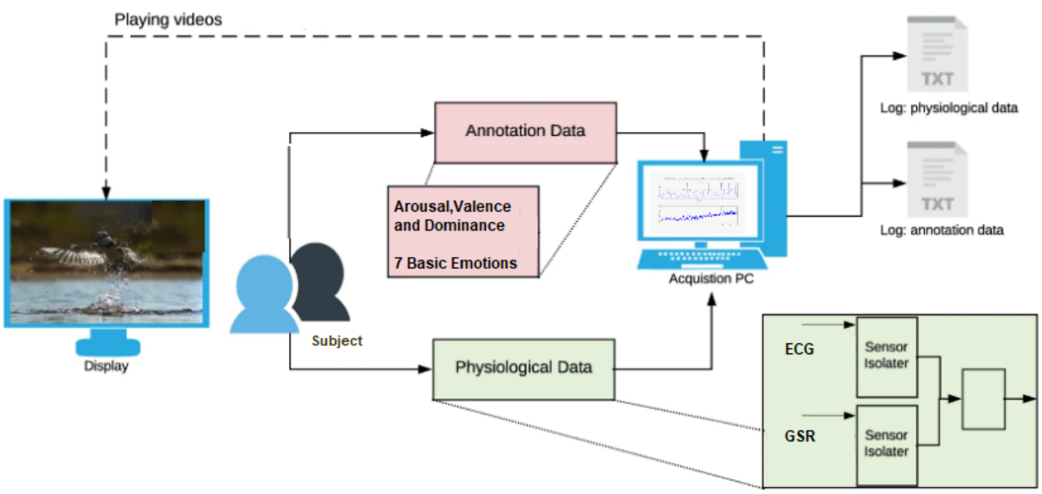


Figure 18 Data collection overview

ECG signal Acquisition:

Every beat of heart is prompted by an electrical impulse generally generated from unique cells in the higher proper chamber of heart (pacemaker cells). An electrocardiogram captures the timing and electricity of those indicators as they travel through the coronary heart.

Safety Information

ECG lead is applied to participant's body by removing the sensor from power supply.

Skin Preparation

Skin preparation can help to improve the quality of the signal by removing unnecessary hair and cleansing the oil.

Wires

Length of the wire can change the quality of the signal. It is recommended to use the short wires so that it can minimize the interface experienced in the signal.

GSR signal Acquisition:

Safety Information

Skin preparation can help to improve the quality of the signal by removing unnecessary hair and cleansing the oil.

Sampling Frequency:

Sampling frequency can be adjusted by user. 128 Hz sampling frequency is set and total 5000 samples are captured per signal for 120 seconds.

Electrode Positioning:

One electrode is placed is placed on index finger and other is placed on ear of subject.

4.6 CONVOLUTIONAL NEURAL NETWORK

Deep convolutional Neural Network is a deep learning algorithm in which the image is taken as input and different weights and biases are assigned to them which is used for the classification and differentiates one image from the other. Convolutional Neural Network consists of a sequence of layers as shown on Figure 24. Four main layers are used in the convolutional neural network architecture i.e. Input Layer, Rectified Linear Units layer, Convolutional Layer, Fully-Connected Layer, and a Pooling Layer. These layers are arranged in the form of a stack that forms a full convolutional neural network. The number of layers can be increased as we go deep in the algorithm. Now we will discuss every single layer of Convolutional Neural Network in detail.

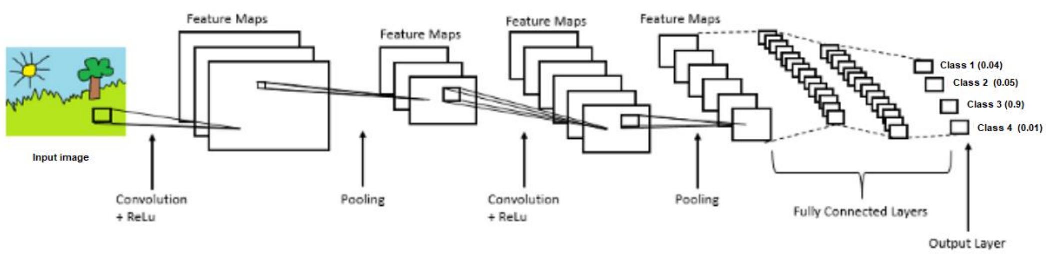


Figure 19 CNN architecture [39]

Input Layer

The input images are encoded into the color channels one of the most common is the red, green, blue (RGB) channel (W x H x D) as shown in Figure 25. When we have to transform the image in the pixel values than based on the intensity of these channels, we form three matrixes. Each matrix represents the intensity of the color at each point in the image. These three matrixes together form a tensor.

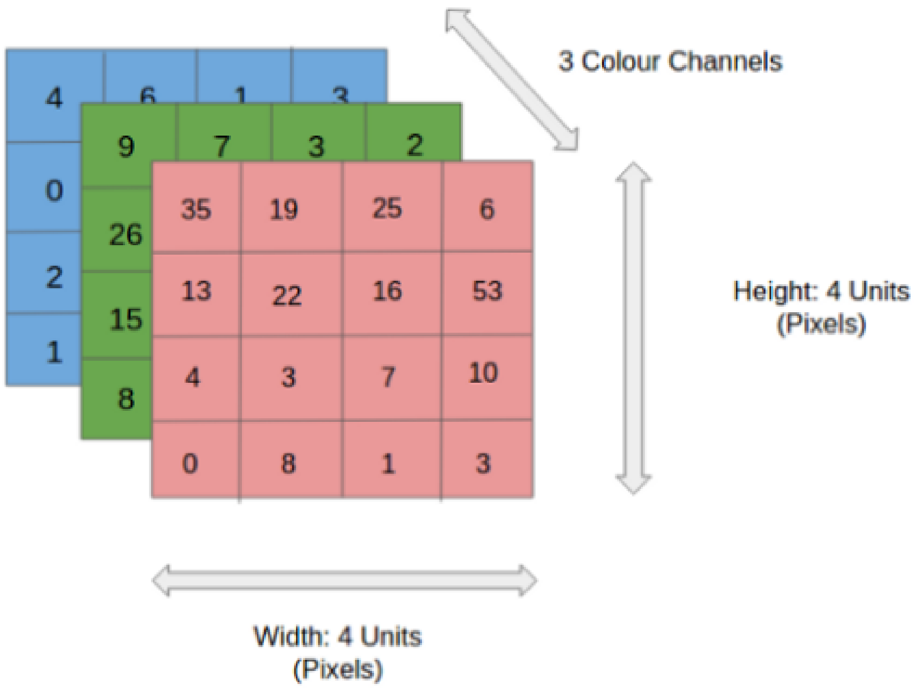


Figure 20 25 4x4 RGB image

Convolution Layer- The Kernel

Convolutional Layer is that layer where the features are extracted from the image, so it is also known as the feature extractor layer. The features are extracted by using the kernel convolution. It involves passing the kernel or filter through the image and in this way, image is transformed based on the values of the filter. The final output of this convolution is calculated using equation 1.

$$G[m,n]=(f*h)[m,n]=\sum\sum h[j,k]f[m-j,n-k] \quad (1)$$

Here, f shows input image, h shows the filter, and m, n represents dimensions of the final output matrix. If the image dimension is (n, n) and the filter dimension is (f, f) then the output dimension can be calculated using equation 2.

$$\text{Dimension of Output}=((n-f+1),(n-f+1)) \quad (2)$$

The output of the convolutional layer is sent to the Rectified Linear Units layer (ReLU) which applies the activation function on the output. The images contain the non-linear features so this function is applied to decrease the non-linearity in the image. The function removes all the negative values from the output and keeps the positive values only.

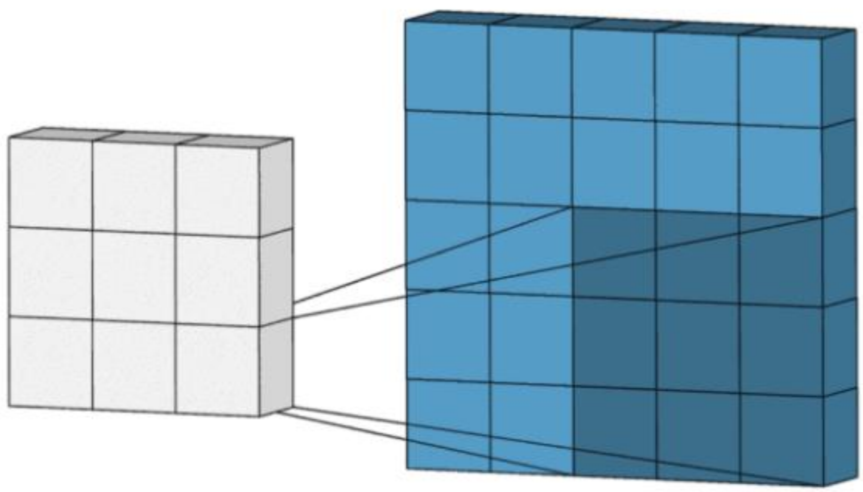


Figure 21 illustration of convolution operation

Pooling Layer

The size of the image is quite and by running our image through different filters the size of the image keeps on increasing which makes our algorithm inefficient. To overcome this problem, we will stack the pooling layer after every convolutional layer. It benefits in reducing the image size drastically by replacing the pixel with min, max, or average of the pixel. The main idea is that we will take a pixel and replace it with soma some function (min, max, average). The most common is the max function that is also called max-pooling. In the max-pooling layer, the size of the image is reduced drastically as we move the filter by multiple pixels in such a way that each pixel is seen by exactly one filter position. It uses two hyperparameters, first is the stride S, and second is their spatial extent. But, the most famous technique is max pooling, which reviews the maximum output from the neighborhood as shown in Figure 27.

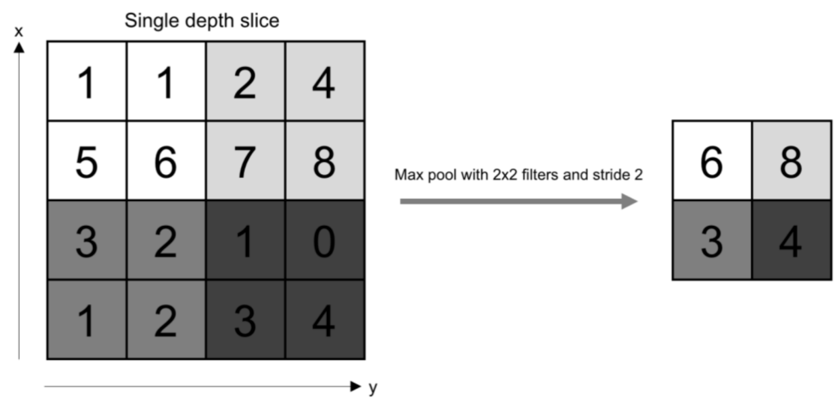


Figure 22 maximum pooling

Classification - Fully Connected Layer

Adding a totally-connected layer is mostly a reasonably-priced manner of gaining knowledge of non-linear combinations of the high-stage features as represented by way of the output of the convolutional layer. The absolutely-related layer is getting to know a possibly non-linear characteristic in that area. Now that we've got transformed our input image into a suitable shape for our Multi-degree Perceptron, we will flatten the image into a column vector. So the output from the pooling layer is transformed right into a vector after which given as an input to the completely related layer. The flattened output is fed to a feed-forward neural community and backpropagation is applied every iteration for each training session. Over a chain of epochs, the version is ready to distinguish among dominating and certain low-stage features in images. One of kind sorts of activation features including sigmoid, hyperbolic tangent, softmax, rectified linear unit (ReLU), exponential linear gadgets (ELUs) capabilities are utilized in absolutely related Layer relying on the application as shown in Figure 28.

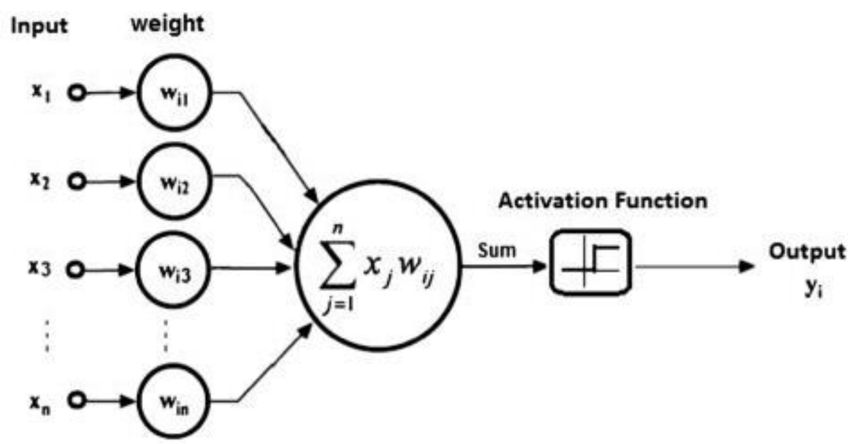


Figure 23 Activation function

After passing from the fully connected layer the very last layer used the softmax activation function which classifies input in the given classes primarily based on the rate of probability. A whole diagram representing all of main components of a Convolutional Neural community is represented in Figure 29.

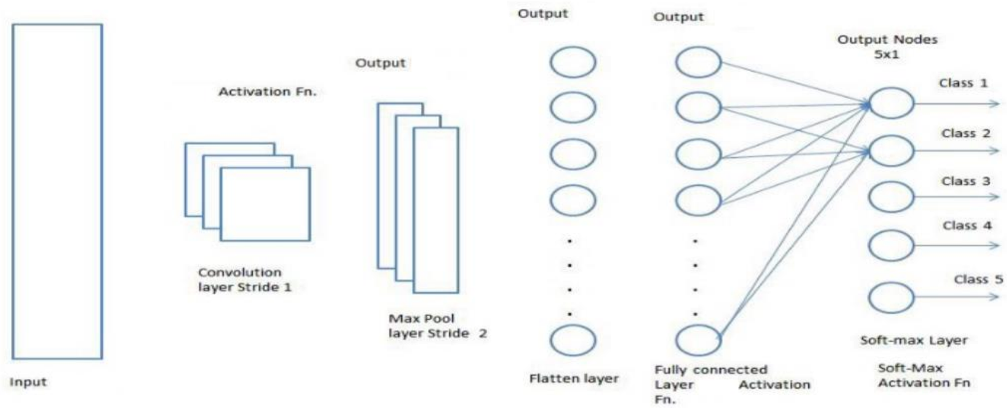


Figure 24 Convolutional Neural Network

There are various architectures of CNNs available which have played a key role in building algorithms which power Artificial Intelligence as a whole in the foreseeable future. Some of them are LeNet, AlexNet, VGGNet, GoogleNet, ResNet etc. In this research thesis we have worked on Alexnet, Resnet and Inception. They are explained below in detail.

4.6.1 ALEXNET

AlexNet is a deep neural network that's 8 layers deep, advanced by way of Alex Krizhevsky and others in 2012. It became designed to categories pictures for the ImageNet LSVRC-2010 opposition, in which it performed nation of the art effects. It also worked with multiple GPUs. AlexNet has been trained on over 1,000,000 images and can classify images into one thousand item classes. The network has found out rich characteristic representations for a wide variety of images. The community takes an image as input and outputs a label for the item within the pictures collectively with the possibilities for every of the object classes. It became a lot larger than previous CNNs used for computer imaginative and prescient duties with 60 million parameters and 650,000 neurons. As an end result, the community has learned wealthy feature representations for a huge variety of pictures. The community has a photo input size of 227-by-227.today there are a lot greater complicated CNNs that may run on faster GPUs very correctly even on very big datasets.

AlexNet includes five Convolutional Layers and 3 completely related Layers as shown in figure 30. Kernels extract interesting capabilities from a picture. In a convolutional layer, there are generally many kernels of the equal size. As an example, the first Conv Layer of AlexNet includes 96 kernels of length 11x11x3. The width and height of the kernel are normally the same and the depth is the same as the quantity of channels. The primary Convolutional layers are accompanied through the Overlapping Max Pooling layers. The 0.33, fourth and 5th convolutional layers are linked immediately. The 5th convolutional layer is followed by an Overlapping Max Pooling layer, the output of which goes into a sequence of two absolutely linked layers. The second one completely related layer feeds right into a softmax classifier with repective labels. ReLU nonlinearity is implemented after the entire convolution and fully linked layers. The ReLU nonlinearity of the first and 2nd convolution layers are observed by way of a neighborhood normalization step earlier than doing pooling.

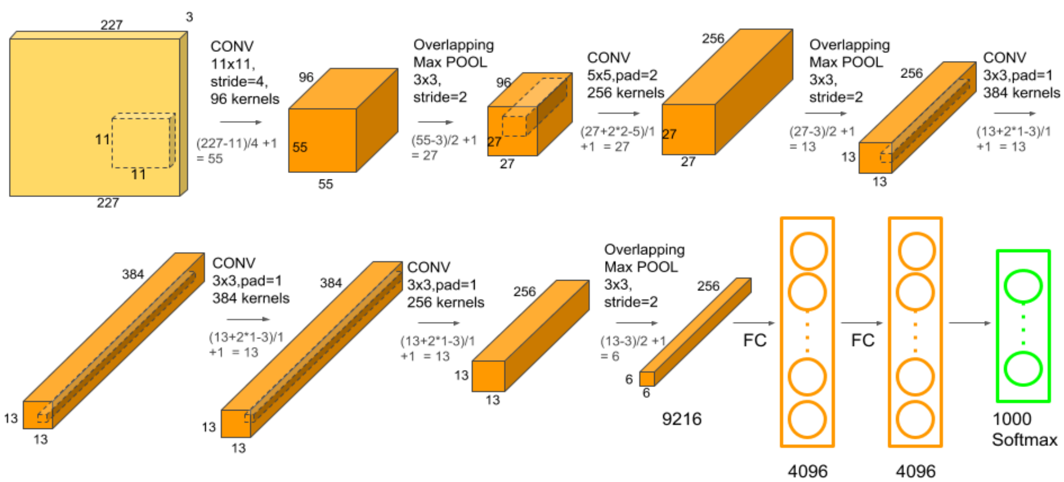


Figure 25 Alexnet Architecture

4.6.2 RESNET

With the increasing network depth, saturation of accuracy occurs and it then rapidly degrades. Over fitting does not cause this degradation and training error becomes higher as more layers are added to a deep model. A deep residual learning framework is proposed and introduced by Microsoft. They explicitly allow stacked layers to fit in a residual underlying. Feed forward neural networks with shortcut connections form F(x) + x as shown in Figure 31.

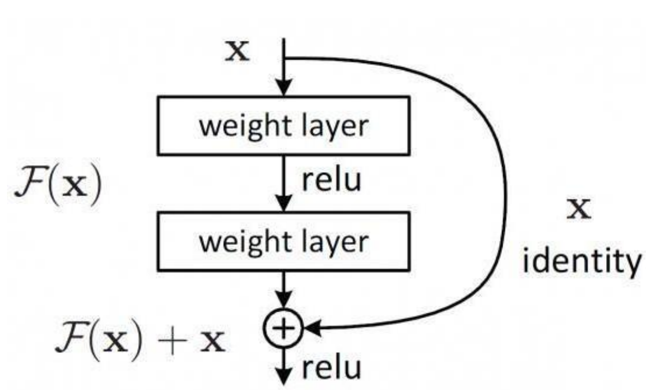


Figure 26 Feed forward Neural Network

For the network into its counterpart residual version a shortcut connection is added. ResNet block is either three layers (ResNet 50, 101, 152) or two layers deep (ResNet 18, 34).

4.6.3 SHUFFLENET

ShuffleNet is a convolutional neural network designed specifically for mobile devices with very limited computing strength. The architecture makes use of two new operations, pointwise institution convolution and channel shuffle, to reduce computation value at the same time as keeping accuracy. Its architecture is shown in Figure 32.

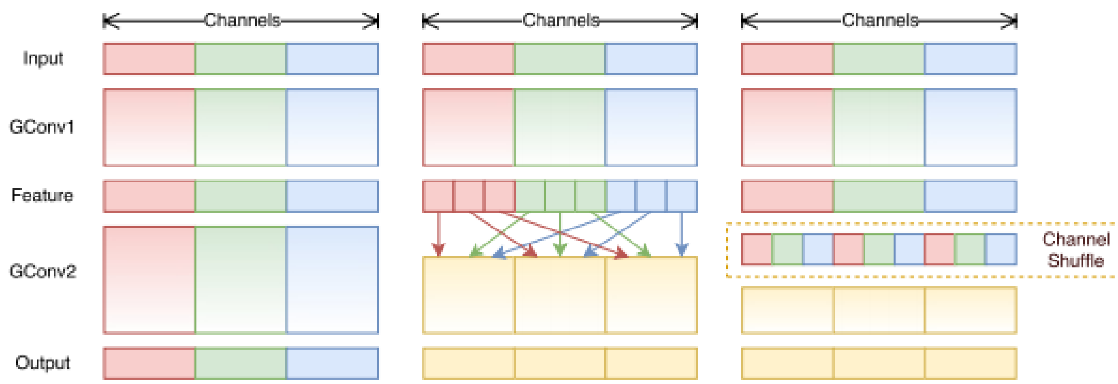


Figure 27Shufflenet architecture

This network is constituted of different ShuffleNet units which are combined in three different stages as shown in Figure 33.

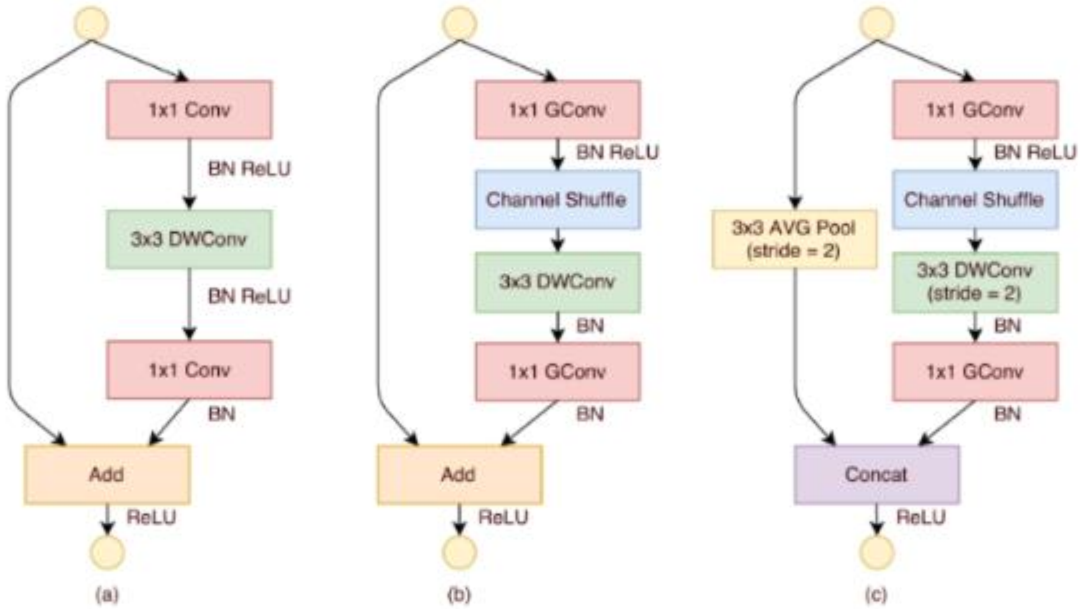


Figure 28Shuffle Unit

4.6.4 INCEPTION

An Inception Module is an image model block that aims to approximate a top-rated nearby sparse shape in a CNN. It may be used to carry out specific obligations as demonstrated in Figure 34. It lets in using a couple of styles of filter length, instead of being constrained to a clear out size, in a single block of image, which then concatenate and bypass onto the following layer. It became an essential milestone inside the development of CNN classifiers. It is far a 48 layers deep convolutional neural network. It may classify images into one thousand categories. As a result, the network has learned rich feature representations for a wide variety of images. It has an image input length of 299-through-299.Inception Modules are utilized in Convolutional Neural Networks to allow for greater efficient computation and deeper Networks through a dimensionality reduction with stacked 1×1 convolutions. The modules had been designed to clear up the hassle of computational price, in addition to overfitting, among different troubles. The answer, in quick, is to take multiple kernel filter out sizes within the CNN, and in place of stacking them sequentially, ordering them to perform on the identical stage.

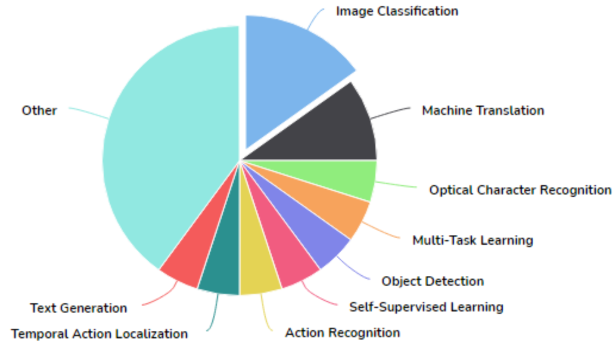


Figure 29 Applications of Inception

Architecture for inception is shown in figure 35.

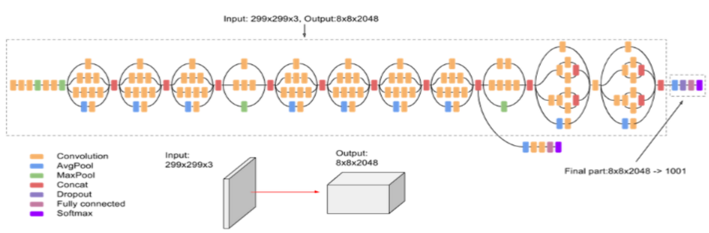


Figure 30 Architecture of inception

4.7 SUMMARY

A detailed introduction of our proposed methodology is explained above. The main focus of this study is to improve the performance of self-collected supervised data via CNN model. For this purpose we used different classifiers for the classification of the ECG and GSR features. We utilized different classifiers on the collected dataset and found out their outcomes discussing classifier giving good results. Classifiers used are Alexnet, Resnet, Inception and others. By following the whole process we achieve good results using Inception.

CHAPTER 5. RESULTS AND DISCUSSION

It explains analysis and results for the collected data. It comprises of two main parts. Firstly, it explains the self-collected dataset. Secondly, it discusses transfer learning techniques which were used for emotion classification. We will discuss the results that we achieved using our proposed methodology. The discussion will be about the techniques which we used and also about the dataset. We will also discuss the result we achieved using the proposed methodology. We will discuss the results a little bit about the proposed methodology and dataset. The sequence of the discussion is as dataset, dataset division and signal conversion from 1D to 2D, attributes assignment, performance measures, transfer learning and classification results. This is the flow of our discussion in the portion of results and discussion.

5.1 DATASET

Dataset is important for the validation of the proposed methodology. Our proposed methodology is used to increase the efficiency of the dataset. So the improvement in the dataset can help its stakeholders. So it is also important in this field. In this case we deal with the medical-related dataset which is very crucial. The dataset that we collected is an emotion classification dataset. This is a medical related dataset. Emotion classification model finds its applications in multiple areas with emotional impairment where an individual cannot properly understand or communicate their true emotional states for example Alzheimer, Parkinson's disease, stroke related aphasia etc. Emotional impairment is a serious condition that adversely affects the quality of life. There are several cause factors due to which an individual cannot clearly understand their emotional state. The improvement in the efficiency of that dataset can help in figuring out the level or stage of the disease of the patient and timely medication of patient can help in improving quality of life. A lot of research has been carried out on this. This self-collected dataset has eighteen attributes including output labels. Shimmer3 ECG and Shimmer3 GSR units are used to capture ECG and GSR signals as shown in Figure 36 and Figure 37 respectively. ECG signals are sampled at a frequency of 256 Hz. Similarly the GSR signals are sampled at frequency of 128 Hz and total 5000 samples are considered for each signal. Signal is recorded for a duration of 150 seconds. We have used 5 lead shimmer3 ECG unit. The Shimmer MATLABTM Instrument

Driver allows for real-time streaming of data from a Shimmer device into MATLABTM. Shimmer ECG and GSR sensor units are paired and connected with the PC over the Bluetooth. Each sensor has its own Com Port number through which their signals are shared over the Bluetooth to PC and saved as mat files. Noise was avoided at time of signal acquisition by removing mobile and all other electronic devices around the subject and asking the subject to stay in the same position without involving unnecessary movements. The ECG data is pre-filtered with a second order Chebyshev low pass filter (LPF) with a corner frequency slightly smaller than the Nyquist frequency and a second order Chebyshev high pass filter (HPF) with a corner frequency of 0.5 Hz for reducing the distortion caused by muscular movement and noise in the environment.

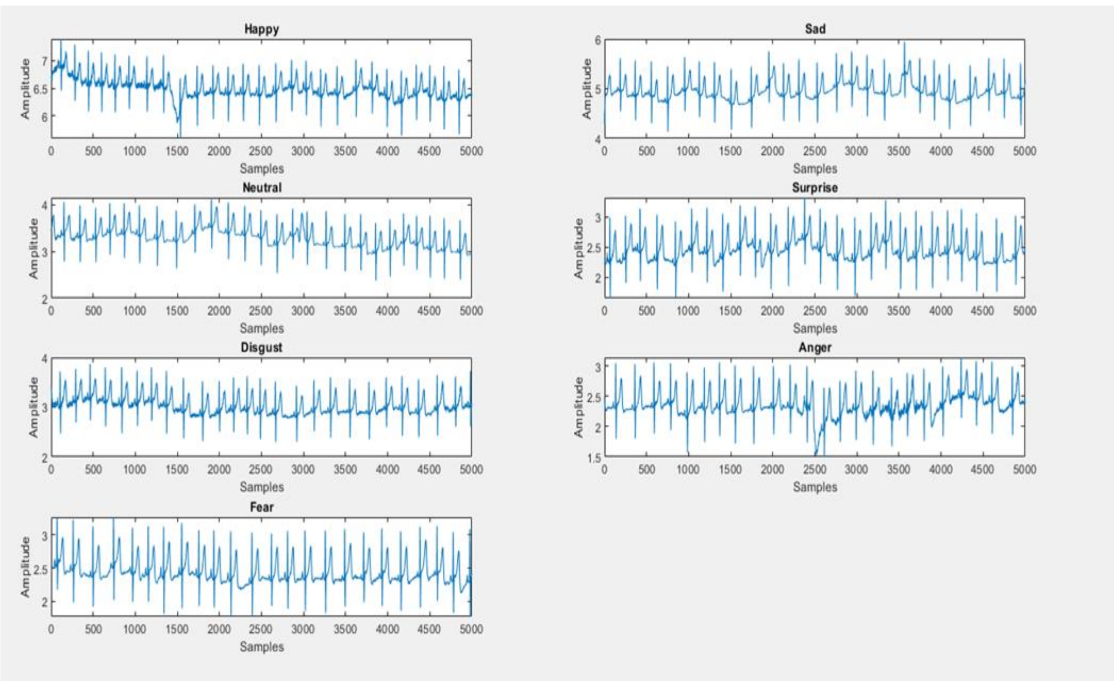


Figure 31 ECG waveform for seven emotions

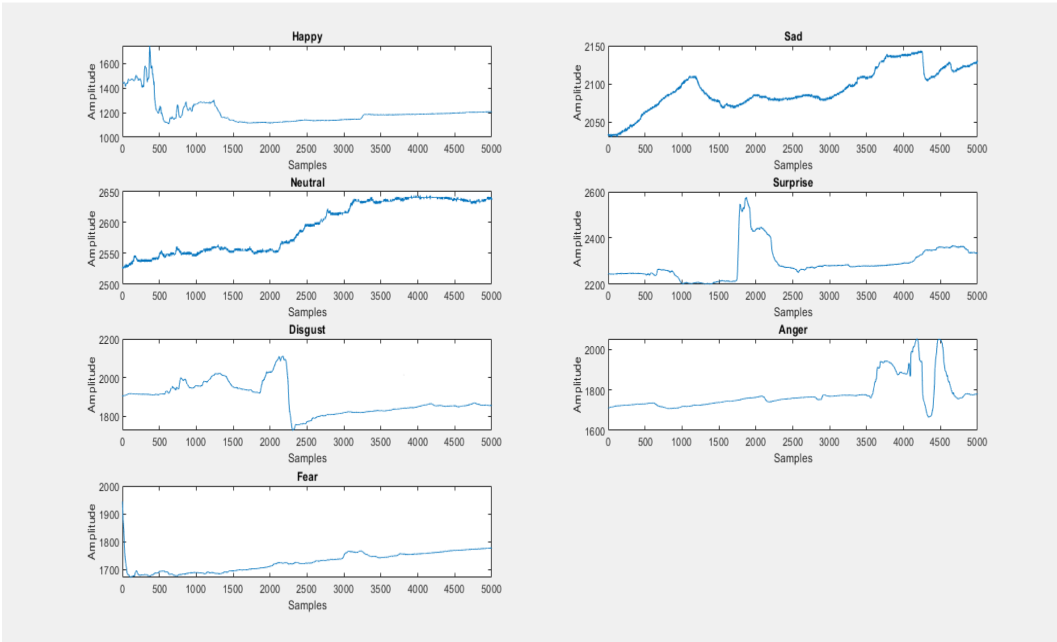


Figure 32 GSR waveform for seven emotions

5.2 DATASET DIVISION AND SIGNAL CONVERSION FROM 1D TO 2D

We divide the dataset into different mat files converted into scalograms for both ECG and GSR signals according to the number of output labels. We have seven labels to identify the images into seven different classes. So we divide it into seven files for each subject. The division of the dataset aims to obtain the separate data of each label as shown in Figure 38.

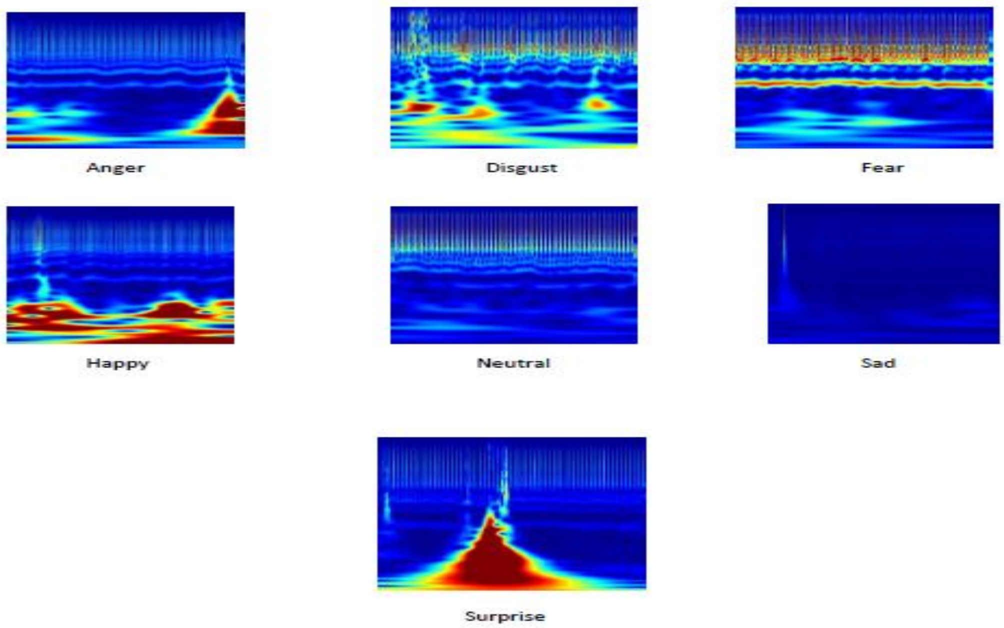


Figure 33 Spectogram for seven emotions using ECG signal

5.3 ATTRIBUTES ASSIGNMENT

We have seventeen attributes in our dataset excluding output label. So our assignments of the attributes for emotion classification model are as follows.

x1 => Participant Identification Number

x2 => Session Identification Number

x3 => Video Identification Number

x4 => Participant Name

x5 => Participant Age

x6 => Participant Gender

x7 => Participant Medical History

x8 => Arousal Level

x9 => Valence Level

x10 => Dominance Level

x11 => Percentage for Happy emotional state

x12 => Percentage for Sad emotional state x13 => Percentage for Neutral emotional state

x14 => Percentage for Surprise emotional state

x14 => Percentage for Anger emotional state

x16 => Percentage for Fear emotional state

x17 => Video familiarity score

5.4 PERFORMANCE MEASURES

For examining the efficiency of different detection algorithms, key performance index can be obtained using the true classification and misclassifications in a confusion matrix, as described in Table 2, as an example for a simple binary classification. The elements in diagonal show the correct classification for a specific class, class A in this case and are termed as a True Positive and when the remaining classes which are not of interest are also correctly identified as class B. The two terms namely False Positive and False Negative indicate 'false alarm' and 'missed identification, respectively.

Table 2 Performance metrics

True/Predicted	A	B
A	True Positive	False Negative
B	False Positive	True Negative

5.5 TRANSFER LEARNING

Transfer learning is known as a popular method in system gaining knowledge because it lets in us to build correct models in a timesaving manner. With switch studying, instead of beginning the learning technique from scratch, you begin from styles which have been learned while solving a different problem. This way you leverage previous learnings and keep away from starting from scratch. It is usually expressed via using pre-trained frameworks. A pre-trained model is a model that became trained on a big benchmark dataset to resolve a problem just like the only that we need to clear up. As a result, due to the computational cost of training such models, it's far not unusual exercise to import and use architectures from published literature. Pre-trained networks have special traits that depend while selecting a community to apply to your hassle as described in Figure 39. The maximum crucial characteristics are network accuracy, pace and length. Selecting a network is commonly a tradeoff among those characteristics. We are able to take a pre-trained network that has already learned to extract effective and useful informative capabilities from natural images and use it as a start line to learn a brand new problem.

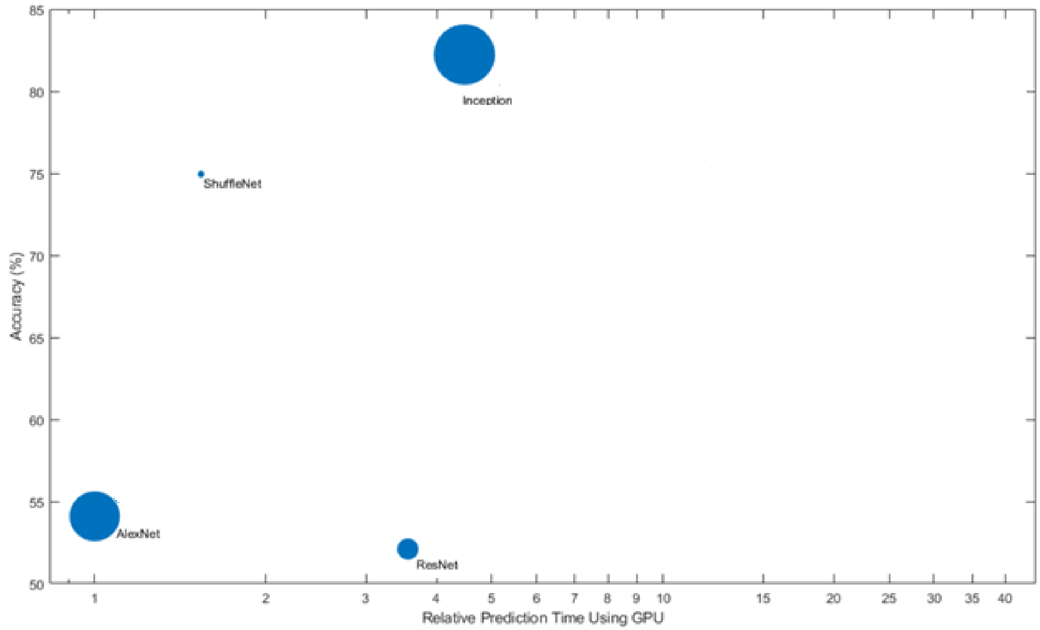


Figure 34 Comparison of different pre-trained models

You can fine tune deeper layers inside the network with the aid of network training to a new data set with the pre-trained network as a place to begin. Fine-tuning a network with transfer learning is frequently quicker and less difficult than constructing and building a brand new network. The network has already found out a wealthy set of training images, however whilst you fine tune the network it can analyze capabilities precise for your new images. When you have a totally large data, then transfer learning may not be quicker than training from scratch. Fine tuning a network is slower and calls for extra attempt than simple feature extraction. It generally works better than function extraction as long as the new dataset isn't very small, due to the fact then the network has information to study new functions from. It is represented in Figure 40.

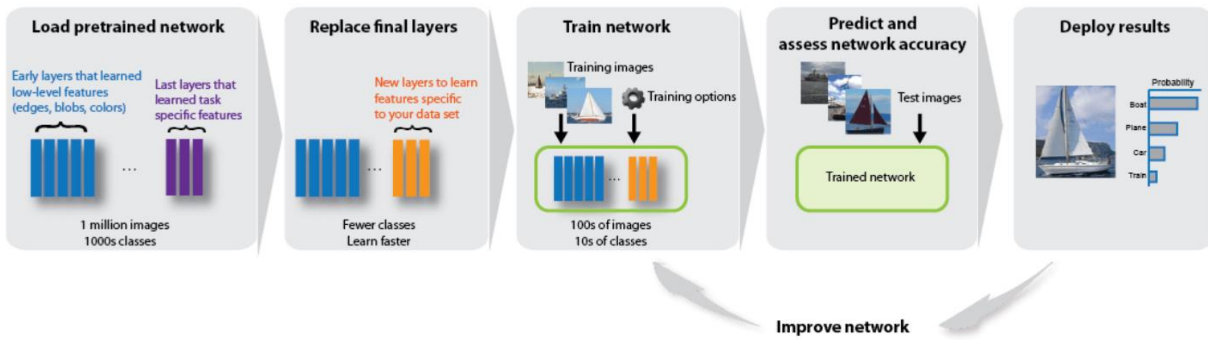


Figure 35 Transfer Learning process

5.6 CLASSIFICATION RESULTS

This section presents classification results obtained using different methods. For validation of our proposed methodology, we collect the accuracy results using four different CNN architectures for self-collected dataset. These four different architectures are Alexnet, Resnet, Shufflenet and Inception. This section discusses the results of our proposed technique. The dataset is collected by showing short term videos to 27 subjects in three different sessions. It consists of ECG and GSR signals of subjects which allow the study of their personality, mood and stimuli duration. Participant's emotions have been both self-annotated and external annotated. 21 short videos were shown to subjects. This dataset was further preprocessed. Labels self-assessment contains 12 columns.1st column represents arousal (integer between 0 and 9), 2nd column shows valence (integer between 0 and 9),3rd column shows dominance(integer between 0 and 9),4th column shows familiarity(integer between 1 and 9),5th column shows medical history. Remaining columns shows basic emotions i.e. happiness, sadness, neutral, surprise, disgust, anger and fear are shown in 6th,7th,8th,9th,10th,11th,12th columns respectively which are taken percentagewise. For data classification, we utilized the pre-trained Inception architecture and transfer learned it on scalograms representing seven different emotions for seven different classes. The scalograms show the signal spectrogram in different colors. These colors represent different movements and speeds and are utilized for training the deep learning network for classification of emotions. Furthermore, we directly train Inception network and compared it with the transfer learned AlexNet, Resnet and Shufflenet architectures and demonstrate the effectiveness of our proposed transfer learned model. The multi-classification performance of 7 emotional states is illustrated using confusion matrices as shown in Figures 41 (a) and (b). The proposed technique provides a classification accuracy of 88.3% for both ECG and GSR signals as shown in Figure 45.In comparison; Alexnet provides a classification accuracy of 59.7% and 65.7% for ECG and GSR signals respectively. Whereas Resnet provides a classification accuracy of 10.4% and 39% for ECG and GSR signals respectively. Furthermore, the proposed model achieves high classification accuracy for Anger, Happy and Sad emotion detection using ECG signals and Fear detection using GSR signals. So both Alexnet and Resnet provide relatively lower classification accuracy as compared to Inception.

The result obtained using Alexnet on the self-collected dataset is shown in Figure 41. The results are obtained on the original dataset which we collected and after applying the preprocessing. We achieved an accuracy score of 63.9 % for ECG and GSR signals.

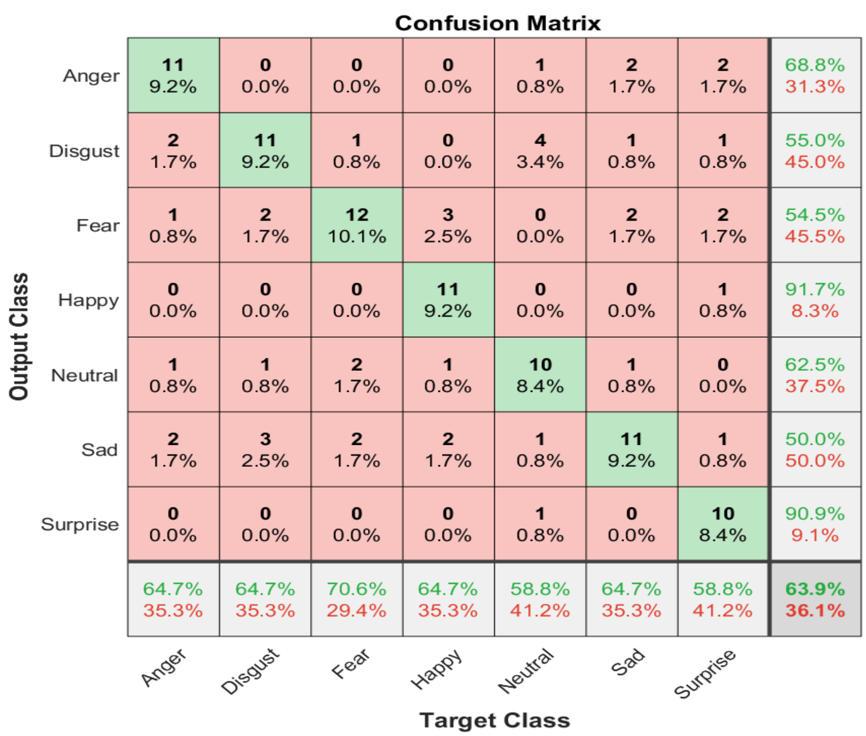


Figure 36 Confusion Matrix for Alexnet using ECG & GSR signals

Confusion matrices for Resnet are shown in Figure 42 for both ECG and GSR signals and we achieved an accuracy score of 39%. There is not too much good progress in achieving good results using Resnet architecture. The obtained efficiency is very low for decision making. For proper decision making and to obtain good results we need to classify the data with high accuracy which would be reliable.

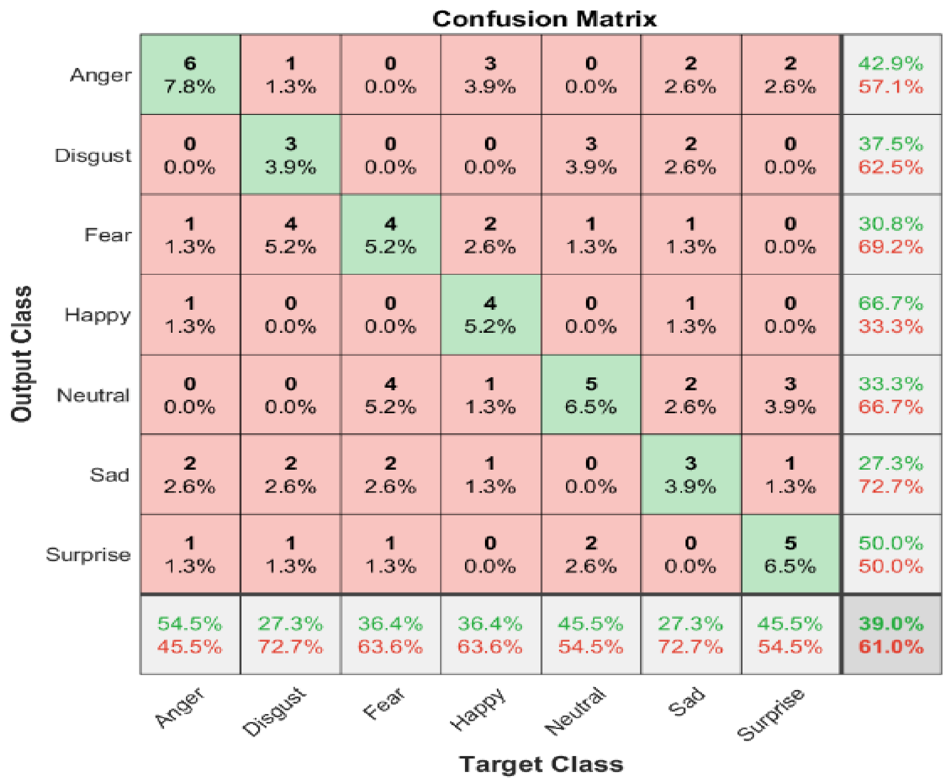


Figure 37 Confusion Matrix for Resnet using ECG & GSR signals

The accuracy score of the Shufflenet is shown in Figure 43 for both ECG and GSR signals which 76.6% accuracy. The confusion matrix of Inception using both ECG and GSR signals is shown in Figure 44 which achieves an accuracy of 88.3% that is a good number of accuracy for decision making. The accuracy score which we obtained is very excellent for the decision making on the basis of the collected dataset. This accuracy score is very high from the accuracy score that we obtained using Alexnet, Resnet and other architectures. This is how our proposed methodology is performing in increasing the accuracy of the collected dataset which is very complex and has very little accuracy using the other architectures. We are quite good at proving out that our proposed methodology is right. This is how our proposed methodology improves the accuracy of the dataset from 50% to 88%. This is a high level of margin to achieve.

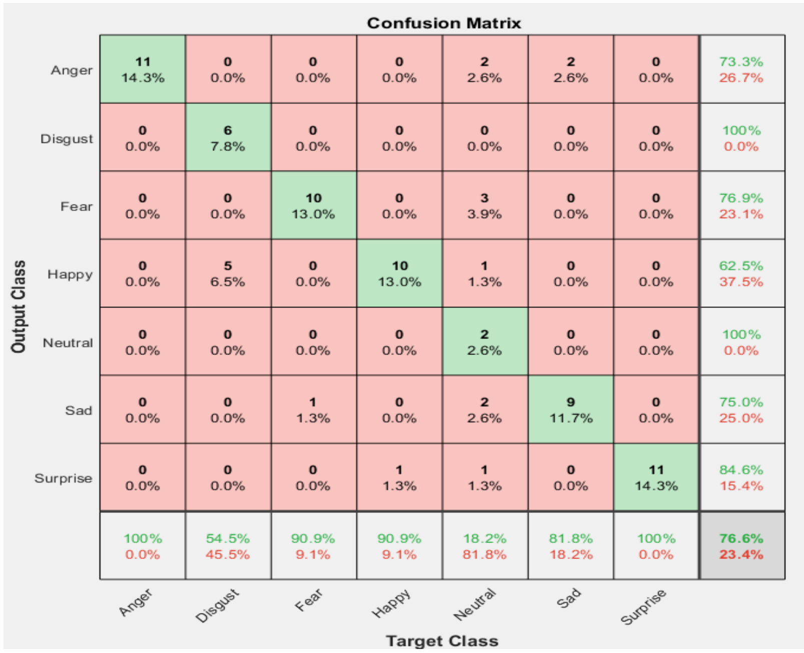


Figure 38 Confusion matrix for Shufflenet using ECG and GSR

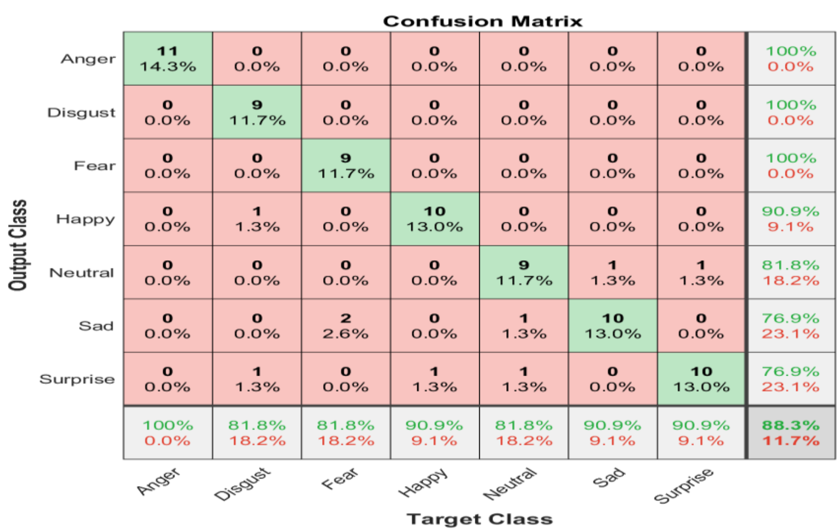


Figure 39 Confusion Matrix for Inception using ECG & GSR signals

The comparison among all the architectures that we have worked on is shown in Table 3.

Table 3 Comparison of different architectures

Algorithm	**Layers**	**Input Size**	**Computational Time (hours)**	**Accuracy**
ECG				
Alexnet	25	227x227	1	59.7%
Resnet	177	224x224	3.5	10.4%
Inception	315	299x299	4	80.5%
GSR				
Alexnet	25	227x227	1	65.7%
Resnet	177	224x224	3.5	39%
Inception	315	299x299	4	79.2%
ECG and GSR				
Shufflenet	44	224x224	1.5	76.6%
Inception	315	299x299	4	88%

The algorithm assigns weight to each class label and class having maximum weight among all the classes is identified as target class as shown in Figure 46. For example for first video of participant 1 in first session the class for sad gets maximum score among all other classes so the predicted class label for P1S1V1 is sad.

	A	B	C	D	E	F	G	H
1		Anger	Disgust	Fear	Happy	Neutral	Sad	Surprise
2	P1S1V1	0.131394	0.000536	0.003192	0.042975	0.212842	0.609016	4.61E-05
3	P1S1V2	0.131394	0.000536	0.003192	0.042975	0.212842	0.609016	4.61E-05
4	P1S1V3	0.131394	0.000536	0.003192	0.042975	0.212842	0.609016	4.61E-05
5	P1S1V4	0.568794	0.000882	0.020522	0.041685	0.28015	0.085948	0.002018
6	P1S1V5	0.568794	0.000882	0.020522	0.041685	0.28015	0.085948	0.002018
7	P1S1V6	0.514322	5.44E-05	0.000735	0.006767	0.056874	0.419555	0.001693
8	P1S1V7	0.00036	2.18E-05	6.14E-05	0.023274	0.005217	0.969176	0.00189

Figure 45 Determination of Predicted class

Table 4 provides the comparison of state-of-the-art techniques with our proposed system and model. Our proposed model utilizes a deeper CNN model as compared to the proposed CNN models. We achieve emotion classification accuracy of 88.3% using both ECG and GSR signals respectively. Presently, a massive majority of emotion classification research make use of Electroencephalograms (EEG) and facial features to carry out emotion classification. Fewer research were conducted the use of the ECG and GSR to this give up. It was observed that there are constrained studies concerning emotion classification for GSR or in aggregate with ECG. Proposed technique focuses on ECG and GSR being the signal of desire EEG is certainly the most customarily used signal in emotion classification however ECG is also one of the common used signals apart from EEG and is favored due to its non-invasive nature.

Table 4 Comparison of state of art techniques

Author	**Emotions**	**Signals**	**Technique**	**Accuracy**
Tzirakis et al. [40]	2	ECG,EDA	CNN RNN	0.1% error
Muhammad Zeeshan Baig et al.[41]	1	ECG	SVM	90%
G. Kanagaraj et al. [42]	4	ECG	SVM, KNN,	75%–80%
Xu et al. [43]	2	ECG,EDA	CNN	71.4%
W. Wei et al. [44]	3	EEG,ECG,GSR	SVM	68.75%
Khairun Nisa et al. [45]	3	ECG	SVM	68%
Hany Ferdinando et al. [46]	2	ECG	KNN	66.1%
Katsigiannis et al. [47]	3	ECG,EEG	SVM	62.63%
Alam et al. [48]	6	ECG,EMG,EDA,PPG	RNN	52%

CHAPTER 6. CONCLUSION & FUTURE WORK

6.1 Conclusion

Human body response is a good indicator of emotion detection as it cannot be concealed. We proposed a new physiological dataset for emotion classification into seven states of emotions and 3d arousal, valence and dominance dimensions. Results have shown that performance of the data is improved by using the proposed approach. ECG and GSR signals against different emotions are obtained through experiments and an optimal classification model is found by experimenting with several CNN models for classifying emotions. The prediction accuracy of dataset was low with other architectures e-g Alexnet and Resnet. But our proposed methodology aimed to improve the efficiency of the dataset using the Inception architecture. After using the proposed methodology the results are different. Unlike preceding studies using many physiological indicators for emotion identification, the present have a look at attempted to achieve an best emotion category model by simply the use of ECG and GSR. The proposed emotion type version shows excellent consequences, it takes an exceedingly long time due to the various parameters and the calculation requires excessive computing energy. Therefore, in addition studies must be carried out to improve the overall performance of emotion type with fewer parameters and advanced CNN fashions. Indicators were captured for 2 mins inside the test. If the constructed model has been commercialized, this can be an alternatively long term process. Although there are obstacles in the present studies, it is possible to develop a more efficient emotion estimation technology.

6.2 Future Work

Our proposed methodology can be used for almost every supervised dataset with slight changes. It can increase the efficiency of the dataset. ECG and GSR based data was the input for the CNN model. As we know CNN has different architectures like LeNet, AlexNet, VGG, GoogLeNet, ResNet, Sequential, and some variations of connections of layers as well as fully connected and have some dimensions variations as well, such as: 1D, 2D, 3D CNN architectures. All have their own benefits. Transfer learned CNN automatically extracts the features from the images by using propagations. L2 regularization can play an important role during the determination of the accuracy for those datasets which have fewer numbers of training data. It will avoid over-fitting in the model. The number of features can be more or less by adding and subtracting numbers of layers accordingly. And then you can use those features for the classification of the data. For future investigations, ECG and GSR signal acquisition from subject can be further simplified using wearable technology. The classifier used plays a vital position in obtaining a correct result for classification; therefore a more particular research as to which classifier offers the best accuracy for seven-class emotion type using ECG and GSR signals may be essential for future work.

REFERENCES

[1]https://www.lap-publishing.com/catalog/details//store/gb/book/978-620-2-91996-8/human-emotion-recognition-using-bio-signals

[2] Shu, Lin et al. "A Review of Emotion Recognition Using Physiological Signals." *Sensors (Basel, Switzerland)* vol. 18,7 2074. 28 Jun. 2018, doi: 10.3390/s18072074

[3] Robinson, Peter, and Rana el Kaliouby. "Computation of emotions in man and machines." *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* vol. 364, 1535 (2009): 3441-7. doi:10.1098/rstb.2009.0198

[4] Ayan Seal, Puthi Prem Nivesh Reddy, Pingali Chaithanya, Arramada Meghana, Kamireddy Jahnavi, Ondrej Krejcar, Radovan Hudak, "An EEG Database and Its Initial Benchmark Emotion Classification Performance", *Computational and Mathematical Methods in Medicine*, vol. 2020, Article ID 8303465, 14 pages, 2020. https://doi.org/10.1155/2020/8303465.

[5] Lim, Nangyeon. "Cultural differences in emotion: differences in emotional arousal level between the East and the West." *Integrative medicine research* vol. 5,2 (2016): 105-109. doi:10.1016/j.imr.2016.03.004

[6]S. Issa, Q. Peng and X. You, "Emotion Classification Using EEG Brain Signals and the Broad Learning System," in IEEE Transactions on Systems, Man, and Cybernetics: Systems, doi: 10.1109/TSMC.2020.2969686.

[7] Costanzi M, Cianfanelli B, Saraulli D, Lasaponara S, Doricchi F, Cestari V, Rossi-Arnaud C. The Effect of Emotional Valence and Arousal on Visuo-Spatial Working Memory: Incidental

Emotional Learning and Memory for Object-Location. Front Psychol. 2019 Nov 19;10:2587. doi: 10.3389/fpsyg.2019.02587. PMID: 31803120; PMCID: PMC6877739.

[8] https://www.verywellmind.com/an-overview-of-the-types-of-emotions-4163976

[9] Mantelet, Fabrice, "Consideration of the emotional perception of the consumers within the design process",(2006).

[10] Ercheng Pei, Dongmei Jiang, Hichem Sahli,"An efficient model-level fusion approach for continuous affect recognition from audiovisual signals",Neurocomputing,Volume 376,Pages 42-53,ISSN 0925-2312,https://doi.org/10.1016/j.neucom.2019.09.037,2020.

[11] https://europepmc.org/articles/PMC6069143/figure/sensors-18-02074-f005/

[12]https://open.oregonstate.education/aandp/chapter/12-1-structure-and-function-of-the-nervous-system/

[13] S. Vijayakumar, R. Flynn and N. Murray, "A Comparative Study of Machine Learning Techniques for Emotion Recognition from Peripheral Physiological Signals," 2020 31st Irish Signals and Systems Conference (ISSC), 2020, pp. 1-6, doi: 10.1109/ISSC49989.2020.9180193.

[14] Mannepalli, Kasiprasad & Narahari, Panyam & Suman, Maloji. (2016). FDBN: Design and development of Fractional Deep Belief Networks for speaker emotion recognition. International Journal of Speech Technology. 19. 10.1007/s10772-016-9368-y.

[15] S. Wioleta, "Using physiological signals for emotion recognition," 2013 6th International Conference on Human System Interactions (HSI), 2013, pp. 556-561, doi: 10.1109/HSI.2013.6577880.

[16] A. Rahim, A. Sagheer, K. Nadeem, M. N. Dar, A. Rahim and U. Akram, "Emotion Charting Using Real-time Monitoring of Physiological Signals," 2019 International Conference on Robotics and Automation in Industry (ICRAI), 2019, pp. 1-5, doi: 10.1109/ICRAI47710.2019.8967398.

[17] Posada-Quintero, Hugo F et al. "Time-varying analysis of electrodermal activity during exercise." PloS one vol. 13,6 e0198328. 1 Jun. 2018, doi:10.1371/journal.pone.0198328

[18] Shaffer, Fred, and J P Ginsberg. "An Overview of Heart Rate Variability Metrics and Norms." Frontiers in public health vol. 5 258. 28 Sep. 2017, doi:10.3389/fpubh.2017.00258

[19] Pham T, Lau ZJ, Chen SHA, Makowski D. Heart Rate Variability in Psychology: A Review of HRV Indices and an Analysis Tutorial. *Sensors*. 2021; 21(12):3998. https://doi.org/10.3390/s21123998

[20] Ante Topic, Mladen Russo,Emotion recognition based on EEG feature maps through deep learning network, Engineering Science and Technology, an International Journal,2021,ISSN 2215-0986,

[21] H. Dabas, C. Sethi, C. Dua, M. Dalawat and D. Sethia, "Emotion Classification Using EEG Signals", Proc. ACM Int. Conf. Computer Science and Artificial Intelligence, pp. 380-384, 2018.

[22] Y. Luo et al., "EEG-Based Emotion Classification Using Spiking Neural Networks," in IEEE Access, vol. 8, pp. 46007-46016, 2020, doi: 10.1109/ACCESS.2020.2978163.

[23]Ganapathy N, Veeranki YR, Swaminathan," Convolutional neural network based emotion classification using Electrodermal activity signals and time-frequency features", R (2020), Expert Syst Appl 159:113571. https://doi.org/10.1016/j.eswa.2020.113571

[24] Zheng, Lim & Mountstephens, James & Teo, Jason, "Multiclass Emotion Classification Using Pupil Size in VR: Tuning Support Vector Machines to Improve Performance",Journal of Physics: Conference Series. 1529. 052062. 10.1088/1742-6596/1529/5/052062, (2020).

[25] Oh, SeungJun,Lee, Jun-Young,Kim, Dong K, "The Design of CNN Architectures for Optimal Six Basic Emotion Classification Using Multiple Physiological Signals" Sensors 20, no. 3: 866. https://doi.org/10.3390/s20030866,2020.
[26] Seal, Ayan et al, “An EEG Database and Its Initial Benchmark Emotion Classification Performance”,Computational and mathematical methods in medicine vol. 2020 8303465., doi:10.1155/2020/8303465,3 Aug. 2020.
[27] Yang Liu, Tom Gedeon, Sabrina Caldwell, Shouxu Lin, and Zi Jin,” Emotion Recognition Through Observer's Physiological Signals”,arXiv preprint arXiv:2002.08034 (2020).
[28] Song, T., Lu, G., Yan, J.,”Emotion recognition based on physiological signals using convolution neural networks”,. In: Proceedings of the 2020 12th International Conference on Machine Learning and Computing, pp. 161–165 (2020).
[29]Ayata,Deger & Yaslan, Yusuf & Kamasak, Mustafa,” Emotion Recognition from Multimodal Physiological Signals for Emotion Aware Healthcare Systems”, Journal of Medical and Biological Engineering. 40. 10.1007/s40846-019-00505-7,(2020).
[30]Memar, M., Mokaribolhassan,“A. Stress level classification using statistical analysis of skin conductance signal while driving”, SN Appl. Sci. 3, 64 (2021).
[31]S. Issa, Q. Peng and X. You, "Emotion Classification Using EEG Brain Signals and the Broad Learning System," in IEEE Transactions on Systems, Man, and Cybernetics: Systems, doi: 10.1109/TSMC.2020.2969686.
[32]Jinxiang Liao, Qinghua Zhong, Yongsheng Zhu and Dongli Cai,”Multimodal Physiological Signal Emotion Recognition Based on Convolutional Recurrent Neural Network”,in IOP Conference Series: Materials Science and Engineering, Volume 782, 2. Energy Utilization and Smart Grids,Sci. Eng. 782 032005,2020.
[33]R. Chinmayi, N. Sreeja, A. S. Nair, M. K. Jayakumar, R. Gowri and A. Jaiswal, "Emotion Classification Using Deep Learning," 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), 2020, pp. 1063-1068, doi: 10.1109/ICSSIT48917.2020.9214103.
[34]Arturo Martínez-Rodrigo, Luz Fernández-Aguilar ,Roberto Zangróniz ,José M. Latorre ,José M. Pastor, Antonio Fernández-Caballero,”Film mood induction and emotion classification using physiological signals for health and wellness promotion in older adults living alone,”in Expert Systems, 10.1111/exsy.12535, 37, 2, (2020).
[35] https://ourworldindata.org/global-mental-health
[36] https://www.researchgate.net/figure/Shimmer3-ECG-EMG-Bluetooth-device-configured-in-EMG-mode_fig5_315047201
[37] https://litfl.com/wp-content/uploads/2018/08/ECG-5-lead-electrode-ECG.png
[38] https://www.shimmersensing.com/products/shimmer3-wireless-gsr-sensor
[39]https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

[40] P.Tzirakis, G.Trigeorgis, M.A.Nicolaou, B. W. Schuller, and S.Zafeiriou,”End-to-end multimodal emotion recognition using deep neural networks”, IEEE Journal of Selected Topics in Signal Processing, 11(8):1301–1309, 2017
[41]M.Z. Baig, M. Kavakli,”A survey on psycho-physiological analysis & measurement methods in multimodal systems”,Multimodal Technol Interact, 3 (2) (2019), p. 37
[42]G. Kanagaraj, S.G. Ponnambalam, N. Jawahar,”Trends in intelligent robotics, automation, and manufacturing”,Commun Comput Inf Sci, 330 (2012), pp. 491-501

[43]Y.Xu and G.Yuan Liu,”A method of emotion recognition based on ecg signal. In Computational Intelligence and Natural Computing”,2009.CINC'09,International Conference on, vol. 1, p. 202-205,IEEE,2009
[44]W. Wei, Q. Jia, Y. Feng, G. Chen, “Emotion recognition based on weighted fusion strategy of multichannel physiological signals”, Comput Intell Neurosci, 1 (2018), p. 2018
[45]K.N. Minhad, S.H.M.D. Ali, M.B.I. Reaz,”A design framework for human emotion recognition using electrocardiogram and skin conductance response signals”,J Eng Sci Technol, 12 (11) (2017), pp. 3102-3119
[46]H. Ferdinando, T. Seppänen, E. Alasaarela,”Enhancing emotion recognition from ECG signals using supervised dimensionality reduction”,ICPRAM 2017 - Proc. 6th Int. Conf. Pattern Recognit. Appl. Methods, 2017 (2017), pp. 112-128 Janua, no. October
[47]S. Katsigiannis, N. Ramzan,”DREAMER: a database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-shelf devices”,IEEE J Biomed Heal Inf, 22 (1) (2018), pp. 98-107
[48]M.Golam,R.Alam, S.F.Abedin, S.Il.Moon, A.Talukder, A.K.Bairagi,M. S. Munir, S.Ullah, and C. S. Hong,”Affective mood mining through deep recurrent neural network”, p. 501–503, 2017.