# STUDY ON HUMAN EMOTION BASED ON FACIAL EXPRESSIONS THROUGH HUMAN MACHINE COMMUNICATION

By

Tahmin Aysha Murshed
14301055

Prantar Dastider
14101052

A thesis submitted to the Department of Computer Science and Engineering in partial fulfillment of the requirements for the degree of B.Sc. in Computer Science

Department of Computer Science and Engineering
BRAC University
April 2020

© [2020]. BRAC University
All rights reserved.

## Declaration

It is hereby declared that

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

**Student's Full Name & Signature**

| **Tahmin Aysha Murshed** | **Prantar Dastider** |
|---|---|
| 14301055 | 14101052 |

# Approval

The thesis/project titled "STUDY ON HUMAN EMOTION BASED ON FACIAL EXPRESSIONS THROUGH HUMAN MACHINE COMMUNICATION" submitted by

1. Tahmin Aysha Murshed (14301055)
2. Prantar Dastider (14101052)

of [Semester], [Year] has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on April 7, 2020.

**Examining Committee:**

Supervisor:
(Member)

Amitabha Chakrabarty, PhD
Associate Professor
Department of Computer Science and Engineering

BRAC University

Program Coordinator:
(Member)

DR.Md.Golam Rabiul Alam

Associate Professor
Department of Computer Science and Engineering

Departmental Head:
(Chair)

Prof. Mahbub Majumdar
Chairperson
Dept. of Computer Science & Engineering
Brac University

DR.Mahbubul Alam Majumdar

Professor
Department of Computer Science and Engineering

BRAC University

## Ethics Statement

We, hereby declare that this thesis is based on the results we obtained from our work. Due acknowledgement has been made in the text to all other material used. This thesis, neither in whole nor in part, has been previously submitted by anyone to any other university or institute for the award of any degree.

## Abstract

Facial expressions are reactions of internal emotions that express states of human minds. In busy life people are ignorant about the emotions of others. Sometimes it's hard for a person to interact with another individual by understanding their mental condition. As a result, accidents, violent events, misunderstandings and even rates of divorce are raised across the world. But computers can be a source of relief in anger management, depression assessment or simply sharing joy. Using the projections of feelings from facial surface by mainly observing the unusual movements of nose, lips, eyes, cheeks etc. through a camera for detection of proper emotion to begin a communication between a human and a machine mainly to help the end user to get out of distress or to amplify inner positivity. Our study tends to use the methodologies of machine learning, data science, image processing and several other applied processes of these to perceive and analyze the responses of facial aspects to meet the expected outcome. Several studies have taken place considering this particular physical fatigue but none has expressed an approach to help people psychologically neither at work places or at places where they accommodate where our study is a step towards tomorrow for a more realistic interaction between machines and humans.

**Keywords:** Facial expression recognition; emotion detection; CNN algorithm; machine learning.

## Acknowledgement

Firstly, and foremost, we would like to thank our Almighty for enabling us to conduct our research, give our best e orts and complete it. Secondly, we would like to thank our supervisor Amitabha Chakrabarty sir for his feedback, support, guidance and contribution in conducting the research and preparation of the report. He encouraged us to conduct the research, gave guidance to us and always was present there for any help we could ask for. We are grateful to him for his excellent supervision to successfully conduct our research. We are also grateful to Md. Golam Rabiul Alam for his valuable guidance. In addition, we like to extend our gratitude to our family and friends, who guided us with kindness and gave inspiration and with their suggestions. Last but not the least, we thank BRAC University for providing us the opportunity of conducting this research and for giving us the chance to complete our Bachelor's degree.

# Table of Contents

## List of Tables

# List of Figures

## List of Acronyms

| | |
|---|---|
| CNN | Convolutional Neural Network |
| ROI | Region of Interest |

# Chapter1

## Introduction

In a society we are bound to meet one another at random events and whenever there's a meeting or simply a gathering the first thing we notice is face mostly. We tend to study the faces of others to understand their state of mind. Of course, this is all done unintentionally as this thought process is kind of installed into our minds. So, then we ask about the whereabouts and try to help our friend whom we are having that conversation with. But if this whole scenario takes place between a human and a machine, we can start a study to work with. The process involves a camera, a desktop and a system that will analyze human faces to understand emotions and then by reading data from the memory this whole machine will communicate on its own to help the person in distress.

In the last couple of years this idea has been worked on several times. In most cases successfully by detecting the proper facial expressions focusing on aspects of facial fatigue [1]. These individual researches have followed different paths to meet the outcome which gave us a field full of different directions all reaching similar points. According to one particular study machine learning algorithms are the best options than traditional pattern analysis and craft rules [2] as the accuracy level differs a lot. Techniques like Principal Component Analysis (PCA) have lower accuracy rates than Fisher's Linear Discriminant and 2-Dimensional Component Analysis [3]. Our work is to apply our very own along with the help of one of these several methods to detect the emotion and to start a process in the very first stages of this research.

Our methodologies include image classifiers that match expressions and output will generate textual representation of the emotion. Camera works in phases for better image detection again following recrimination of the network.

As the research is in process, a draft of the assembled data set has been presented. The data contains images of faces that are reported properly oriented almost in the same manner. The two different type levels of Convolution Neural Network have different procedures.

Through convolution pooling phase shortens the parameters to make the training time shorter as it also restrains overlap. In the classifier part programming takes over for import to read files that include libraries.

More than 5 libraries need to be associated to acquire maximum output. We incorporated sequential(), model.add(Conv2D()), model.add(Dropout()) etc. as all the libraries that hold the process. Layering stacks from input to output for convolutional layers where the trigger function consists of methods generated by all the processes.

The model compiles cross entropy categories as it is a loss function but the dame optimizer is put in also. For validating the measurement, we have used precession. With a fixed size batch, times and testing data that had been generated earlier we needed to divide the training data. In the end pattern was saved.

After detection of impression the machine is able to give outputs and tell the type of emotion it has detected. In this phase of the process we have used real time video by webcams. We have tried to build communication between human and machine as it was never the final point only to reach the detection level of emotions being portrayed in front of the camera.

## 1.1 Thought Behind Working on emotion detection

Evidence reveals that more than 90% of our interactions can be non-verbal, but technology has had challenges in preserving it. Also, our intonations and thoughts are usually not understood by standard language. Yet cognitive acknowledgement — also named affective computing — is now available to more users. Situational emotion awareness has common

social and business implications. In the public sector, the capacity of government agencies to spot feelings such as shame, anxiety and confusion may be abused. The auto-scanning TSA travelers for potential threats is not impossible to envision and makes the country a safer place. Companies also used personal appreciation to improve company success. On iPhone X, Apple also launched a novel emoji feature, dubbed Animoji, which allows to replicate facial expressions with a device.

It is not so unlikely to imagine that in other projects they will shortly use these skills. That's the basic target that pushes us to choose this topic of work.

## 1.2 Problem Statement

There were lots of studies done with emotion detection by facial expression. Though all of the studies mentioned in the paper are somewhat the same in generating outputs but almost none are actually used to support the end users psychologically exactly where our study enters. Imagine a driver who is depressed for some reason but still needs to continue driving. The situation is as complicated as it sounds but does not have to be like that. If our system is installed in the car, he can reach someone to talk to even if it's a machine rather it is an advantage because sometimes, we don't completely open our minds to our friends and family where talking with a machine is more helpful in this case.

At this early stage of research, we are still studying more published papers of institutes and side by side preparing a main and better flow chart of our own methods to build a better than before structure of this study. We are very much hopeful of reaching our goals to help the people who really need the emotional support in their life in different complicated phases of their lives.

## 1.3 Research Contribution

We recommended a blueprint for developing a future-oriented device that can tackle the emerging complexities of e client connectivity in the machine-oriented markets. In order to satisfy the purpose, we make the device more responsive by laying music by understanding the mood of the user. Therefore, by reducing error and increasing performance, the new system would improve the science of Artificial intelligence.

## 1.4 Research Objectives

The goal of the whole research is, the development of the device, that is communicative and more interactive with humans by understanding their emotion. Several of the model goals suggested are therefore listed below:

- To develop human machine interaction.
- To decrease the error of the previous works.
- To understand the emotion of humans more accurately.
- To make the machine work more like a person.

# Chapter 2

## Related Work

### 2.1 Literature Review:

A number of researches have been done to interact between human and machine by detecting emotions through facial expressions. Busso et al. [1] conducted the experiments by using a database based on one female speaker. They used three systems to train to recognize her emotion through voice and facial expressions. They have also resolved some pairs of emotions that are usually misclassified, like- sadness and neutral.

With 12 feature descriptors and 6 machine learning algorithms, JCT.Kwong et al. [2] had generated 72 models that classify 7 human emotions in their study. They have got the best result in classifying joy and the weakest suit with being anger. This study comes out with a conclusion with 25 out of 72 models having 0.90 or higher accuracies that can be used as a framework for emotion recognition through facial expression.

Hariprashad Mal & Dr Swarnalatha P. [3] had introduced a 2-dimensional sensor perceptual color framework that has the highest recognition rate and has the highest performance. This study found that the trend of alleles of genes identical to the geometric factors of facial expressions.

In the field of Facial expression recognition process as many techniques have already taken place authors Husam Salih & Lalit Kulkarni [4] actually talk about the various advantages and disadvantages of methodologies like LGBP-Top, CLM, Patch processing etc. and propose a

comparison of these applied functions and algorithms. The study reveals the accuracy of these methods under challenging scenarios.

Using the hexagon diamond shape to divide the edges of a human face to detect expressions has been elaborated by Murtaza et al. [5] and the other authors as the study proposes the methodologies of Six Facial Expressions Hexagon (SFEH) model on the outer boundary of face to recognize the six basic emotions. The technique appears to be less complicated than others comparatively as the number of devices used is much less and slight.

By using methods of Viola Jones Algorithm, ROI localization, Grey Level Difference following analysis of data to recognize stress, anger, lack of enthusiasm that is the distressed emotion of the human mind. With this paper Prianka Nair & Subha V [6] approach to help in challenging scenarios to find out someone's capability to help in different situations.

Using Active Appearance Model to design a method that studies face images saved in databases to identify variations in textures, features and shapes. According to the studies conducted by Ashwini Ann Varghese & Jacob P Cherian [7] for the first face recognizing algorithms Active Appearance Model helps and FACS parameters store the values.

According to the work of Kwang-Eun Ko & Kwee-Bo Sim [8] try to apply Active Appearance Model with DBN to structure a combined method that proposes a higher success rate than other technologies available in the field. The study proposes a ninety percent accurate recognition by using datasets of BioId.

Though different from other researches the study conducted by Saurabh Singh [9] along with his fellow researchers state a helpful method EDBL, emotion detection via body language that identifies feelings of human mind detecting physical postures. The research angleze on shoulder and other part gestures.

Using given models and other measuments to track gestures through a probabilistic inference three levels of emotional movements can be angleze. By training data advanced machine learning

techniques have been represented by Shangfei Wang [10] and team. Based on their proposal extensive experiments were conducted with satisfactory accuracy rate.

For tracking 15 facial points through particle filtering in face profile sequences Pantic [11] represented Facial Action Dynamics Recognition in his research. Achieved accuracy rate is 87 percent when in input face profile 27Aus occur without any other element or in combined forms to angleze temporal segments.

## 2.2 Emotion Detection from facial expression

The simplest form of the research is based on detecting expressions. Now as generally asked why detect expressions or why understand the changes in shapes and structures of facial fatigues, we can answer by saying to understand the human mind. As humans it is not always a problem to tell how the other person is feeling about an opinion or on a larger scale in a conversation how the other people are responding we try to understand their expressions. Because, these expressions string to the most basic feeling of human minds and they reflect the opinions which not always are expressed.

As described before it is easier for humans to understand most of the time from looking at someone's face but to make a machine be in that position and do the job is not the same thing which goes without saying. We had to give the machine the idea of a face in it is own language. Then we had to make it capture images, slowly train the changes in structures and shapes of facial fatigues. Finally, the machine was able to understand the changes and connect the expressions of the images to the texts that described the names.

The whole process of our research is based on major three words "Facial Expression Detection" in simple language. However, as stated above as simple as it sounds the process is complicated. To make it easier to understand as the process of our study we have divided the process into two specific segments.

- **Face Detection**

By detecting the faces the process begins.  Among various methods we worked with Haar feature based cascade classifiers proposed by Viola and Jones [12].

To train the classifier the algorithm works with images containing faces and images that do not contain any faces. Haar Cascade in OpenCv has an additional advantage as it has classifiers of its own [13].

- **Emotion Recognition**

When the face has been detected the expression needs to be detected. This process is conducted by deep learning. Here the process is Convolutional Neural Network, CNN. Works like the biological visual cortex that specifies areas of image [14].

# Chapter 3

## Proposed Model

## 3.1 Dataset Description

In the context of an ongoing research study, this data collection was assembled by Pierre- Carrier and Aaron Courville [16]. Here, we presented a draft version of their data for our research.

The training set has 35,888 columns. The file we used, train.csv contains gray images of faces in a 48x48 pixel size. The faces are properly reported such that the face is oriented more or less and space is roughly the same. The goal is to classify any face in one of 7 categories (0=Surprise, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise, 6=Neutral) on the basis of the mood seen in the facial expression.

Train.csv contained “emotion and pixels” in two parts. The “emotion” column includes a number tag for the emotion found in the picture, from 0 to 6. A list separated by quotes for each image is found in the pixel column.

A list enclosed by quotes for each image is found in the pixel column. The length of this list is in the maximum range a space- pixel type. Test.csv only includes the column “pixels” and the emotional column is the mission.


Fig 3.1: Part of the converted dataset that is stored as string.

The images that are shown in Fig 3.1 are originally displayed as a string in the dataset given in the Kaggle relation, which is a row 1 * 2304 of 48*48 images stored as a row vector. We have converted the strings of files.csv to images through code.

## 3.2 Deep learning (CNN)

Convolution Neural Network or CNN is a set of two different types of levels [17] –

1. The hidden layers / Feature Extraction Part
    - **Convolutions:** Convolution is a computation procedure involving two factors in order to generate a third one. Convolution for CNN, a buffer is used to create a function image for the convolution of the inputs.
    - **Pooling:** After a convolution layer, the pooling layer is applied. It constantly decreases size, as for example- it reduces the number of parameters and measurements and thereby reduces training time and prevents overlapping. One such approach is called max pooling, taking the full amount of a window that limits the map size while preserving essential details.
2. The classifier parts

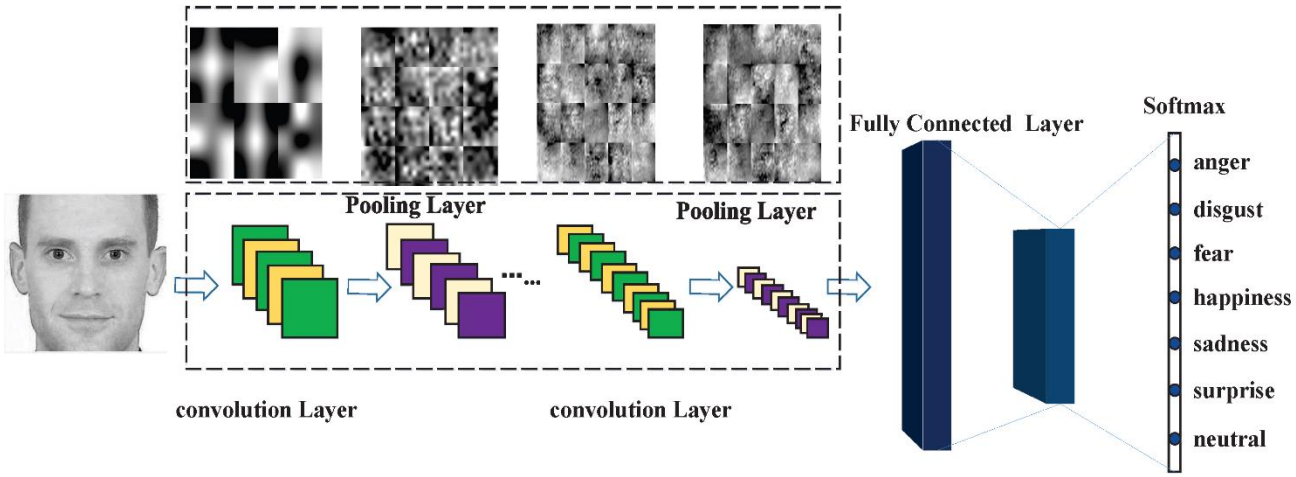



Fig 3.2: The basic design of the algorithm with convolution.

The tasks that are done in code for this research for deep learning are-

- Import and read a CSV file with libraries.
- First, we specify the variables that are required for our CNN instruction. We have a resolution of 48x 48 pixels, and thus we are as large as 48. There are also seven feelings we expect (0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4= Neutral, 5=Sad, 6= Surprise) and then we have seven names. With a batch size of 64, we filter the inputs.
- We then load the characteristics and marks in x, y and uniform x respectively, by removing the means and splitting them into standard deviations.
- Just the pixel values are used in printing the first four lines. Instead, using the train_test_split () function, we divide the data into training and testing and save the results and labels to be used later. In fact, we are splitting the validated data which will be included later in the code into another section of our training data.

## 3.3 CNN Libraries:

We are using a fusion of multiple functions to build CNN [15], which we explain individually here-

- sequential() – A sequential structure is a linear layer stack that places layers on each other as we pass from the input layer to the output layer.
- model.add(Conv2D()) – It is a 2D convolutional layer that carries out the convolution process as defined at the start of the article. In the quotation from Keras Documentation "This layer generates the convolution kernel which is paired with a level source to construct an output quaternion." This is where our trigger function consists of a 3x3 kernel and the Rectified Linear Unit (ReLU).
- model.add(BatchNormalization()): It conducts the group standardization process on the next stage inputs such that our inputs are distributed around the spot in a certain range, 0-1.
- model.add(MaxPooling2D()): This role accomplishes the data pooling process as discussed in the article. In this layout, we have a 2x2 panel window with 2x2 moves.
- model.add(Dropout()) – Dropout is a strategy that avoids picking at random neurons during the preparation. As discussed earlier All are accidentally "dropped out." The overfitting is reduced.
- model.add(Flatten())- The ND to 1D data is compressed so the huge amount scale can not be impacted.
- model.add(Dense())- Dense conducts an action according to Keras Documentation: output = activation(dot(input, kernel)where activation is a layer-based entity synchronization feature passed as activation argument, the kernel is the layer-based weights matrix. In short, it is the last nail in the receptacle that uses the technology which has been learned through layers and maps it into the mark. The ultimate label of the image being analyzed is created by this layer while testing.

- model.summary()- To get the output we will use this function.

The model is compiled with categorical_crossentropy as a loss function and the Adam optimizer [18] is used. We use precision as validation measurements. We then fit a model with a fixed size batch (64 here), times (100 here) and testing data, which we received earlier by dividing the training data. And finally, for some personalized check, we save the pattern.

## 3.4 Workflow

We use cameras for detecting the face. The camera gets introduced to the face and subtracts the other aspects of the visual. As real time records the image it focuses on letting imutils, OpenCv work on the field of the visual.

From the libraries we try to match the newly recorded images to detect the expression. With changes in shapes and structures we propose the system to identify the expression that best matches with other images in the library through training. That is why outputs occur in percentages

of all the emotions and the max percentage gets selected for display.

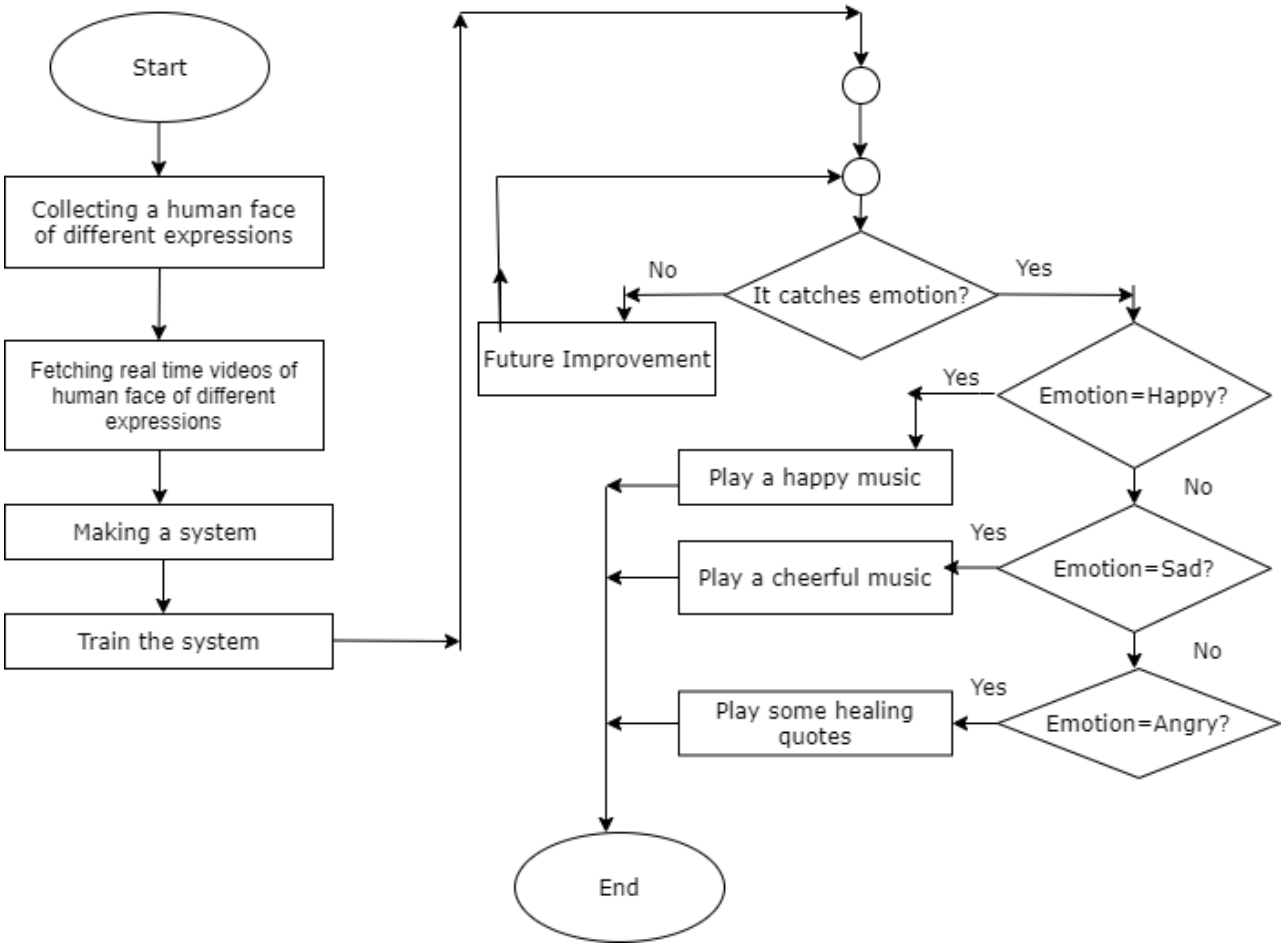



Table 3.4: Workflow in whole

When the emotion is recognized we tend to wait for further assessment. If not required or the output remains unchanged and we proceed to our next level.

In the next level our machine checks the emotion to play a song according to the emotion. For for example if somebody is angry, we play a healing song that will calm him down. The programming checks the emotions in the syntax and plays accordingly.

## 3.5 Device Diagram

We had to work with a graphics processing unit considering the fact that TensorFlow and raspberry pi are not compatible. The latest TensorFlow 2.1 was compatible with the methods of the libraries. The Nvidia graphics version also worked properly with the system.

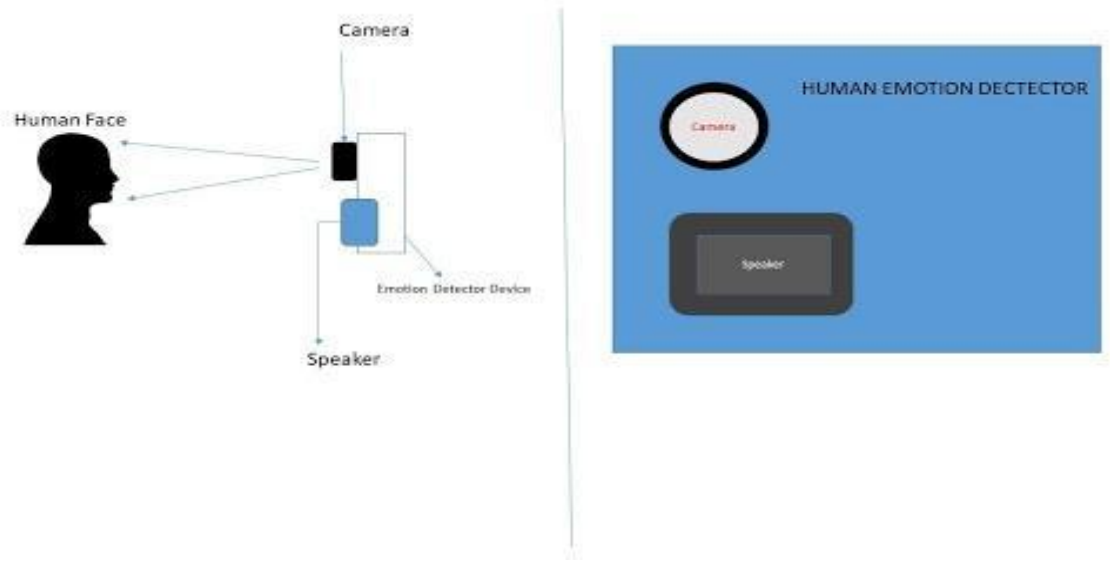



Fig 3.5: Device diagram

The whole setup can work both with internal and external camera modules. We used a ram of 16 gigabytes for the whole process to run without any lags or pause.

Initially the work started by using a raspberry pi, 8 gb ram, ver 1. While working on hardware, firstly we had chosen raspberry pi for this project. We found it impossible. Raspberry pi cannot hold the power while running the TensorFlow. So, we used a GPU and camera for this research.

## 3.6 RealTime testing:

For real time testing, we have imported the keras.processing.image library. Also, a platform named imutils and Opencv is imported. From the keras.model, we loaded the models. To detect the face, we have loaded the haar cascade file.Haar Cascade is the algorithm for object

detection[19]. It is a machine learning approach, in which a cascaded method is educated by both positive and negative images (where positive images are the object to be recognized, negative images are those not). This is then used in other images to detect objects. Fortunately, OpenCV offers pretrained, categorized haar cascade algorithms, based on the images they have been trained

on (faces, eyes, etc.). We prepare emotion classifiers and set an array with elements of angry, disgust, scared, happy, sad, surprised, neutral.

While starting the video streaming, there is a namedWindow opened, the camera is on to read the frame. This frame resizes it to a fixed 48*48px and gray scale. To ignore the unwanted face while reading face, we extract the ROI [21] of the face from the grayscale image.

```
def Speech_to_text ():
    if emotion == 'sad':


        tx = repr('You Look Sad Here is an Exciting music for you')
    elif (emotion == 'happy '):
        tx = repr('You Look Happy so  Here is happy music for you')
    elif (emotion == 'angry'):
        tx  = repr('You Look Angry  so  Please take a deep breath and heal')

    import os
    import platform

    syst = platform.system()
    if syst == 'Linux' and platform.linux_distribution()[0] == "Ubuntu":
        os.system('spd-say %s' % tx)
    elif syst == 'Windows':
        os.system('PowerShell -Command "Add-Type -AssemblyName System.Speech; (New-Object System.Speech.Synthesis.SpeechSynt
    elif syst == 'Darwin':
        os.system('say %s' % tx)
    else:
        raise RuntimeError("Operating System '%s' is not supported" % syst)
```

To predict the emotion, we search for the match of ROI using emotion.classifier. From here we labeled the emotions that are found as angry/ disgust/ scared/ happy/ sad/ surprised/ neutral. A rectangular frame appears to detect the face and calculation done to detect which emotion dominates from the facial expression. Thus it shows the probabilities of the emotion in canvas. The maximum probability is accepted to show the emotion of the facial expression of the user.

With Speech_to_text class we declare the detected emotion with speech. We made the system suitable for Linux, windows and also other operating systems.

## 3.7 Playing music according to emotion of user:

By defining play_music() operation, we import music from the playsound library[20]. The code below shows the work here-

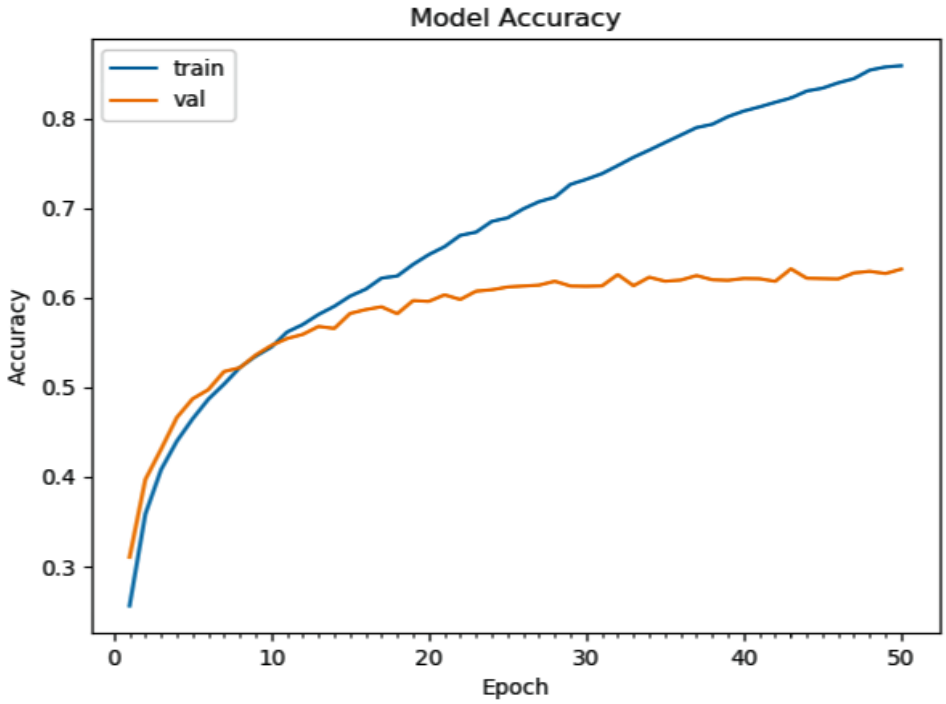
```python
from playsound import playsound
def play_music ():
    if emotion == 'happy':
        playsound('happy.mp3')

    if emotion == 'sad':
        playsound('HealingSong.mp3')

    elif (emotion == 'angry'):
         playsound('HealingSound.mp3')
```

PlaySound plays the audio defined by the file name, resource, or device case. From the preselected music, the system plays music or sounds according to the emotion of the user. We have worked with three emotions to play music for our device. So, when the machine finds the user happy, it will play ‘happy songs’. As well as when it found the user sad or angry- it will play ‘cheerful song’ and ‘healing quotes’.

# Chapter 4

## Result & Analysis

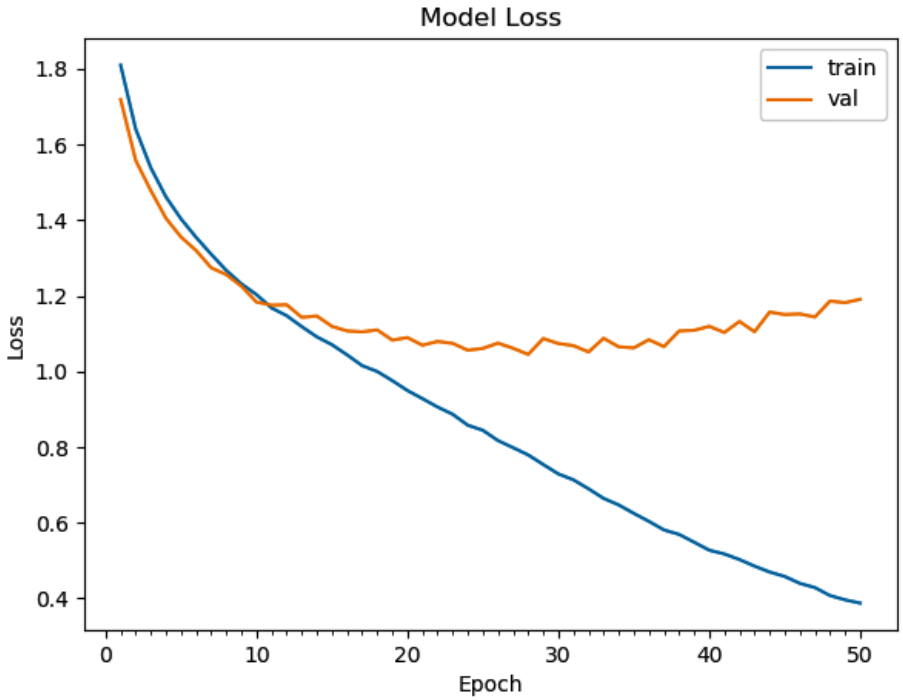
In our trained model we have found accuracy around 66% on average. The Model Accuracy graph demonstrates that our training accuracy was decently high compared to epoch.




*Model Accuracy*

Table 4.1 : Epoch vs Accuracy

Our validation accuracy becomes consistent above 60 percent that we can see from table 4.1.




*Model Loss*

Table 4.2 : Epoch vs Loss

On the other hand, the Model Loss graph shown in Table 4.2 illustrates that our loss was decreasing by the progression of epoch and it becomes consistent in 50 epochs.

In the testing we have found the “happy” emotion and “sad” emotion creates a significant difference in the test image of a person which makes our model more accurate in terms of recognizing the emotion. Whereas, the “angry” and “Surprised” emotions are more unlikely to be distinguished compared to “happy”, and “Sad”.

Our real time emotion detection works well as it is classifying by loading pertained weights of models. For working on a low powered device, it would take more time. The result in the monitor shows the emotion in real time video.

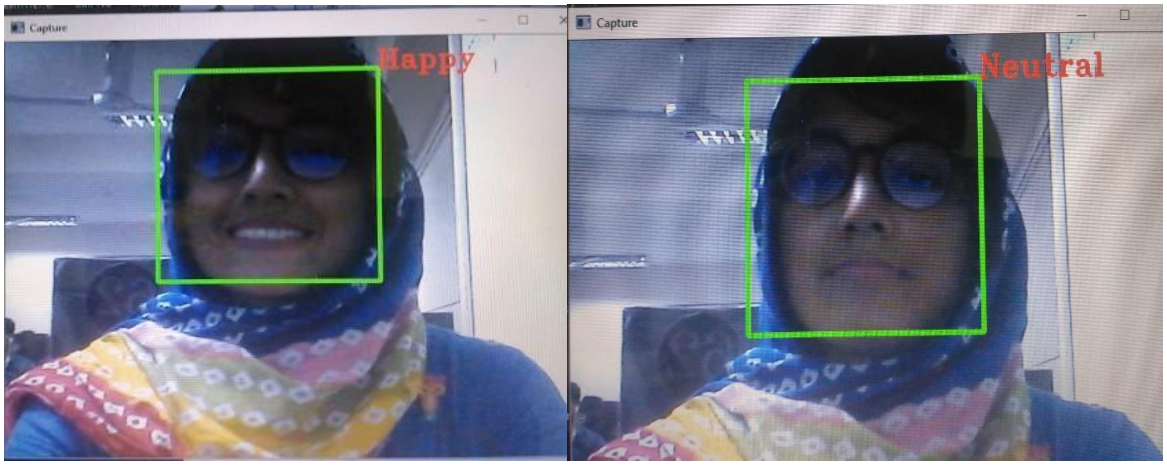



Fig 4.1: The device can grab the emotion accurately and plays music

# CHAPTER 5

## Conclusion and Future Work

### 5.1 Conclusions:

If we consider a common scenario of our country where a driver is driving reckless and he is angry then we can approach the necessity of our study. While driving if someone is using the machine the detection of the anger can help him to calm him down and drive more carefully. The approach has always been to help people psychologically because machines and technologies are more widespread than ever. We have an accuracy rate of approximately seventy percent. We do consider the fact that this is lower than previous research done in the field but our research includes hardware and is applied through devices. The use of OpenCv, CNN and all the libraries has helped to know a lot about data science and we are eager more to apply data science in our research as data programming, machine learning is something we believe we can help the country with. We still have some obstacles to cross over as wearing glasses sometimes is a problem while detecting expressions and detecting two or more faces together is also not entirely done. But with the results being satisfactory we do want to carry forward. We do not want to stop here as the field has more depth to continue such work. We hope to conduct more experiments on other physical fatigue and care to take the country forward. We are greatly thankful to our institution and respected faculties for all the knowledge and help.

### 5.2 Future Work

For the advanced work we thought to add a number of features in our project. These are-

- Make an emotion detection robot, that can understand emotion from facial expression. We consider this option as a great scope as robotics can directly help us by creating a position

in the world map. By building such robots we can actually start a revolution in the fields of psychology as well as robotics.

- Build a communication system between machine and human.
- We want to establish meaningful conversations between machine and human. Similar works have been generated but in the fields of data science a machine will be able to understand the emotions and will be able to answer according to that.
- For image processing along with object detection we can apply architectures like Resnet, Vgg-16. We would like to run various models in distributed architectures.
- For more accurate results, we will detect emotion with facial expression, voice and gesture at a time. The pitch and energy of voice and gesture will help to get the exact emotion for conversation.
- Explore ways to detect mixed emotions (happily crying does not mean that a person is sad). As our study is giving results with an accuracy of more than sixty percent to take the rate higher we are willing to merge other physical aspects for detecting feelings.
- For processing language we would like to apply methods on Hierarchical Neural Networks with that including bidirectional LSTM [22].
- Add bangla as the textual language.
- We would like to apply the whole method to work in domestic help to factory work.

# Bibliography

[1] Carlos Busso, et al., Analysis of Emotion Recognition using Facial Expressions, Speech and Multimodal Information, *ICMI '04 Proceedings of the 6th International Conference on Multimodal Interfaces*, pp. 205-211(2004).

[2] JCT. Kwong, et al., Emotion Recognition via Facial Expression: Utilization of Numerous Feature Descriptors in Different Machine Learning Algorithms, *TENCON 2018 - 2018 IEEE Region 10 Conference*, (2018).

[3] Hariprashad Mal & Dr Swarnalatha P., *2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS)*,(2017).

[4] Husam Salih & Lalit Kulkarni, Study of Video based Facial Expression and Emotions Recognition Methods, *2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*,(2017).

[5] M Murtaza, et al., Facial expression detection using Six Facial Expressions Hexagon (SFEH) model, *IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)*, (2019).

[6] Priyanka Nair & Subha V, Facial Expression Analysis for Distress Detection, *Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*,(2018).

[7] Ashwini Ann Varghese & Jacob P Cherian, Overview on Emotion Recognition System, *2015 International Conference on Soft-Computing and Network Security (ICSNS -2015), Feb. 25 – 27, 2015, Coimbatore, INDIA*

[8] Kwang-Eun Ko & Kwee-Bo Sim, Development of a Facial Emotion Recognition Method based on combining AAM with DBN, *2010 International Conference on Cyberworlds*

[9] Saurabh Singh & Ravi Bhall, EDBL - Algorithm for Detection and Analysis of Emotion Using Body Language, *2015 1st International Conference on Next Generation Computing Technologies (NGCT-2015) Dehradun, India, 4-5 September 2015*

[10] Yongqiang Li & Shangfei Wang, Simultaneous Facial Feature Tracking and Facial Expression Recognition, *IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 22, NO. 7, JULY 2013*

[11] Maja Pantic, Dynamics of Facial Expression: Recognition of Facial Actions and Their Temporal Segments From Face Profile Image Sequences, *IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART B: CYBERNETICS, VOL. 36, NO. 2, APRIL 2006*

[12] Paul Viola & Michael Jones, Rapid Object Detection using a Boosted Cascade of Simple Features, *ACCEPTED CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION 2001*

[13]Face Detection using Haar Cascades — OpenCV-Python Tutorials 1 documentation, Opencv-python-tutroals.readthedocs.io
https://opencv-python-tutroals.readthedocs.io/en/latest/py_tutorials/py_objdetect/py_face_detection/py_face_detection.html

[14]Convolutional Neural Networks in Python,
DataCamp Community
https://www.datacamp.com/community/tutorials/convolutional-neural-networks-python?utm_source=adwords_ppc&utm_campaignid=1455363063&utm_adgroupid=65083631748&utm_device=c&utm_keyword=&utm_matchtype=b&utm_network=g&utm_adpostion=&utm_creative=332602034364&utm_targetid=dsa-473406581915&utm_loc_interest_ms=&utm_loc_physical_ms=9069450&gclid=CjwKCAjwg6b0BRBMEiwANd1_SFVOyzRjx3fZhglaXZ-6XNHN5FS3_nptFjlceRnrAiOzlewDPwfOKRoCp5MQAvD_BwE

[16] Challenges in Representation Learning: Facial Expression Recognition Challenge

https://www.kaggle.com/c/challenges-in-representation-learning-facial-expression-recognition-challenge/data

[17] Yamashita, R., Nishio, M., Do, R.K.G. *et al.* Convolutional neural networks: an overview and application in radiology. *Insights Imaging* **9,** 611–629 (2018). https://doi.org/10.1007/s13244-018-0639-9

[18] https://keras.io/optimizers/

[19] R. Yustiawati et al., "Analyzing Of Different Features Using Haar Cascade Classifier," 2018 *International Conference on Electrical Engineering and Computer Science (ICECOS), Pangkal Pinang, 2018*, pp. 129-134.

[20] https://pypi.org/project/playsound/

[21] NAIR, R. A. N. J. I. T. A. M. A. N. I. Y. A. R. J. A. I. M. I. N. (2019). *Image Processing With Python: embedding image processing to boost the computer vision and... machine vision*. S.l.: PACKT PUBLISHING LIMITED.

[22] Brownlee, J. (2020, January 7). How to Develop a Bidirectional LSTM For Sequence Classification in Python with Keras. Retrieved from https://machinelearningmastery.com/develop-bidirectional-lstm-sequence-classification-python-keras/