ISSN: 2581-4419 Volume 1 Issue 1

EMOTION DETECTION

Durgesh Kolhe[1], Omkar Mandavkar[2], Sameer Metkar[3], Shubham More[4], Prof. Amarja Adgaonkar[5]

[1-4] Student
[5] Assistant Professor
Department of Information Technology
K.C. College of Engineering and Management Studies and Research, Thane

Mumbai, India

Abstract*—Pre-processing, feature extraction, dimensionality reduction, and classification are all part of the traditional speech emotion identification pipeline. Professional feature engineering and the classifier are crucial to recognition performance, which will be more difficult in the case of massive data. Many emotion researchers have recently shifted their focus toward automatic emotion recognition from raw signals, with the motivation that a neural network can learn representation and obtain the final result automatically. This research focuses on classification and provides an update on current developments on end-to-end speech emotion recognition issues. The network model's requirements, process methods, and existing accomplishments are discussed in this survey. We also look at some of the challenges that could arise in the future when it comes to recognising spoken emotion. Deep learning techniques are now widely used in a variety of industries, including computer vision. Indeed, a CNN model can be trained to evaluate photos and recognise facial emotion. In this project, we develop a system that can detect peoples' emotions based on their facial expressions. Face detection using Haar Cascades, normalisation, and emotion recognition using CNN on the FER 2013 database with seven types of expressions are the three steps of our approach. The obtained results suggest that facial emotion detection is possible in education, and as a result, it can assist teachers in tailoring their presentations to the emotions of their students.

I. INTRODUCTION

This is an era of change and innovation; we have seen many technologies and applications of them that few decades ago people may not have even imagined. Amidst all this comes Data Science, the gathering of Data, processing it to give outcomes that are beneficial to humankind. Today Machines are being made intelligent and are used at many places to simply work, but what separates humans from machines is Emotion .And among the multiple techniques, algorithms, applications and techniques is one such application of Emotion Recognition. To go through with this, we decided to use the Convolutional Neural Networks model or recognition algorithm. To detect Facial Expressions and emotions of the person associated with it We intend to use CNN on a Dataset, train it and test its accuracy and analyse it. Facial expression recognition has brought much attention in the past years due to its impact in clinical practice, sociable robotics and education.

Facial emotion recognition

FER typically has four steps. The first is to detect a face in an image and draw a rectangle around it and the next step is to detect landmarks in this face region. The third step is extracting spatial and temporal features from the facial components. The final step is to use a Feature Extraction (FE) classifier and produce the recognition results using the extracted features. Figure 1.1 shows the FER procedure for an input image where a face region and facial landmarks are detected. Facial landmarks are visually salient points such as the end of a nose, and the ends of eyebrows and the mouth as shown in Figure 1.2. The pairwise positions of two landmark points or the local texture of a landmark are used as features. Table 1.1 gives the definitions of 64 primary and secondary landmarks [8]. The spatial and temporal features are extracted from the face and the expression is determined based on one of the facial categories using pattern classifiers.

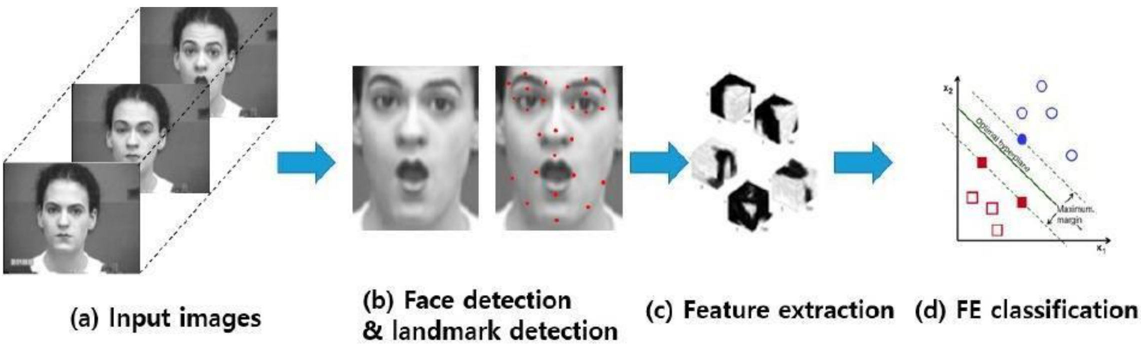


Figure 1.1 FER procedure for an image [9].

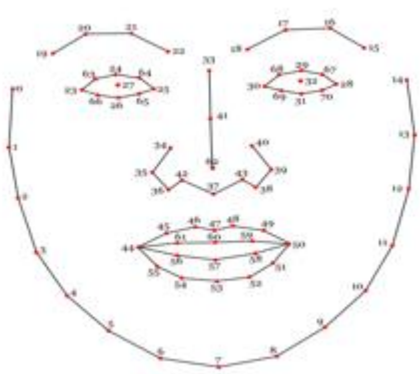


Figure 1.2 Facial landmarks to be extracted from a face.

Primary landmarks		Secondary landmarks	
Number	Definition	Number	Definition
16	Left eyebrow outer corner	1	Left temple
19	Left eyebrow inner corner	8	Chin tip
22	Right eyebrow inner corner	2-7,9-14	Cheek contours
25	Right eyebrow outer corner	15	Right temple
28	Left eye outer corner	16-19	Left eyebrow contours
30	Left eye inner corner	22-25	Right eyebrow corners
32	Right eye inner corner	29,33	Upper eyelid centers
34	Right eye outer corner	31,35	Lower eyelid centers
41	Nose tip	36,37	Nose saddles
46	Left mouth corner	40,42	Nose peaks (nostrils)
52	Right mouth corner	38-40,42-45	Nose contours
63,64	Eye centers	47-51,53-62	Mouth contours

II. LITERATURE SURVEY

Emotion Detection can be and is in real time used in many fields such as finding perpetrators and suspects in a designated area by judging their emotions or even in human computer interaction. If trained properly one of the best use of Emotion Detection is in Biometric Security such as if in a Face Lock/Unlock in a Device, the machine can judge based on the emotions that whether the device is being unlocked forcefully or not. It can be used in preventive medical treatments.

This is an era of change and innovation; we have seen many technologies and applications of them that few decades ago people may not have even imagined. Amidst all this comes Data Science, the gathering of Data, processing it to give outcomes that are beneficial to humankind. Today Machines are being made intelligent and are used at many places to simply work, but what separates humans from machines is Emotion. And among the multiple techniques, algorithms, applications and techniques is one such application of Emotion Recognition.

To go through with this, we decided to use the Convolutional Neural Networks model or recognition algorithm. To detect Facial Expressions and emotions of the person associated with it We intend to use CNN on a Dataset, train it and test its accuracy and analyse it.

Facial expression recognition has brought much attention in the past years due to its impact in clinical practice, sociable robotics and education. According to diverse research, emotion plays an important role in education. Currently, a teacher use exams, questionnaires and observations as sources of feedback but these classical methods often come with low efficiency. Using facial expression of students the teacher can adjust their strategy and their instructional materials to help foster learning of students. The purpose of choosing this Topic is because it has various applications in multiple fields ranging from preventive healthcare to cyber forensics etc. We will see about the algorithm and the topic in-general further.

III. DESIGN SYSTEM

Proposed Algorithm and System Flow Chart

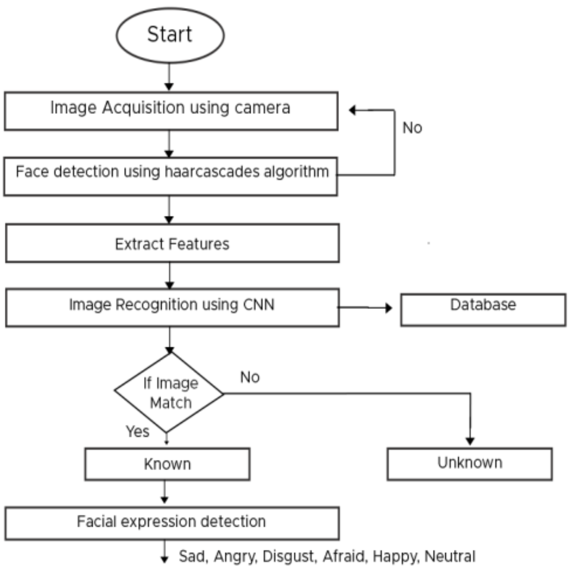


Figure 3.1 Flow Chart

Humans express themselves through emotions, and sometimes emotions show what one can’t speak. Emotion Detection can be used in many such applications such as detecting emotions on face of a baby or animals that can’t express themselves verbally or even generally we can read their emotions and interpret them. In this Project we aim to solve the basic problem of General Emotion Detection such as it can detect emotions of happy, sad , angry, fear , disgust and surprise. Once this problem has been overcome, we can work on changing its use case and expanding its applications.

Mini Xception used in Training Model

Here comes the exciting architecture which is comparatively small and achieves almost state-of-art performance of classifying emotion on this data-set.

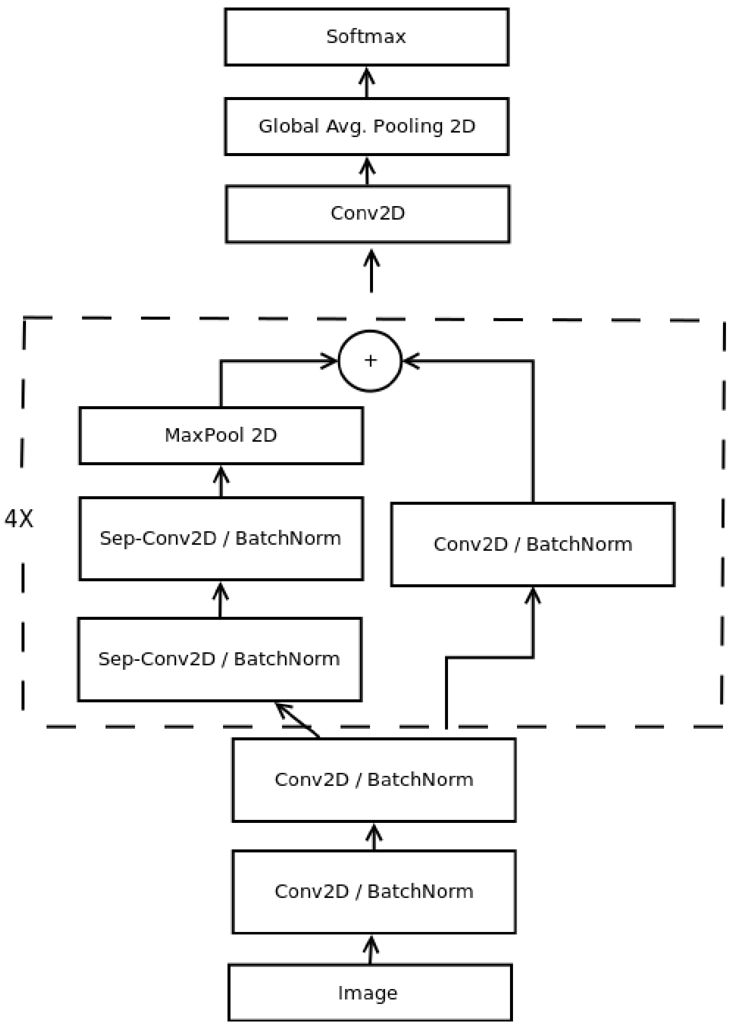


Figure 3.2 Proposed Mini_Xception architecture for emotion classification

One can notice that the center block is repeated 4 times in the design. This architecture is different from the most common CNN architecture like one used in the blog-post here. Common architectures uses fully connected layers at the end where most of parameters resides. Also, they use standard convolutions. Modern CNN architectures such as Xception leverage from the combination of two of the most successful experimental assumptions in CNNs: the use of residual modules and depth-wise separable convolutions.

IV. METHODOLOGY AND IMPLEMENTATION

A CNN is a DL algorithm which takes an input image, assigns importance (learnable weights and biases) to various aspects/objects in the image and is able to differentiate between images. The pre-processing required in a CNN is much lower than other classification algorithms. Figure 3.4 shows the CNN operations. The architecture of a CNN is analogous to that of the connectivity pattern of neurons in the human brain and was inspired by the organization of the visual cortex [32]. One role of a CNN is to reduce images into a form which is easier to process without losing features that are critical for good prediction. This is important when designing an architecture which is not only good at learning features but also is scalable to massive datasets. The main CNN operations are convolution, pooling, batch normalization and dropout which are described below.

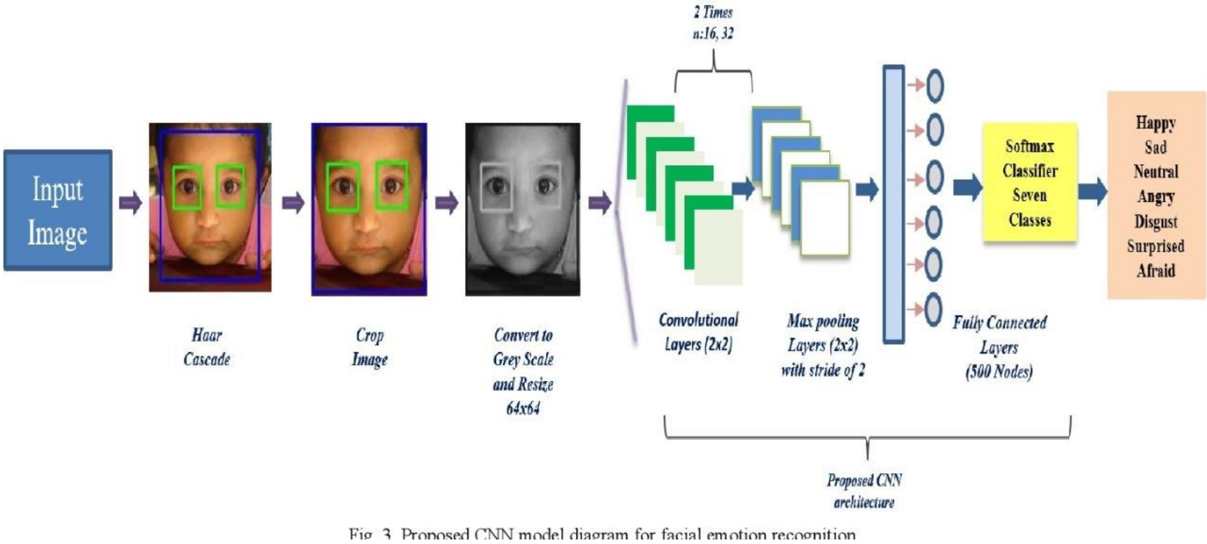


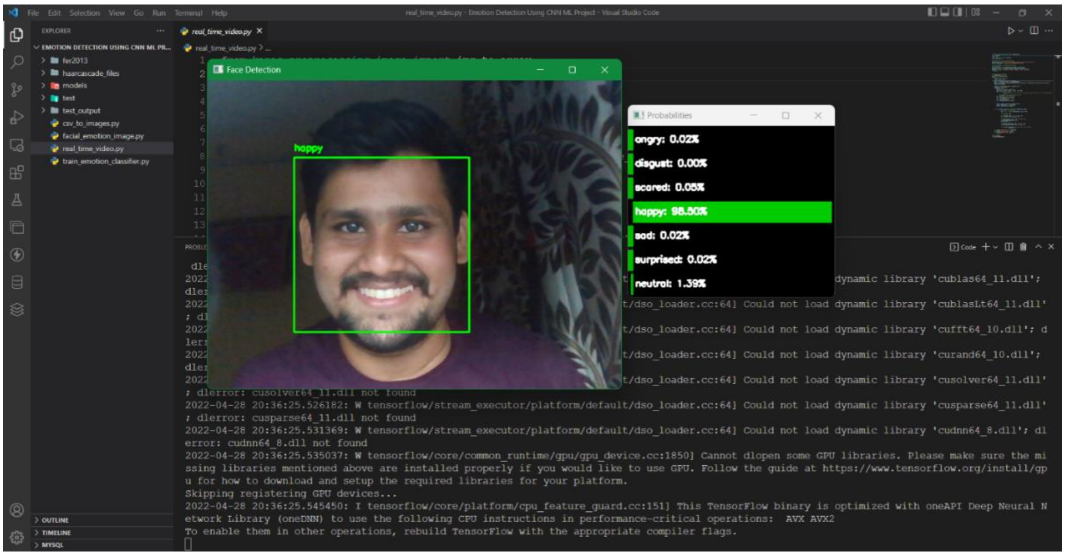
Figure 4.1 The CNN operations [33].

Key Capabilities:

- The main advantage of CNN compared to its predecessors is that it automatically detects the important features without any human supervision. For example, given many pictures of cats and dogs, it can learn the key features for each class by itself.

- Convolutional neural network is composed of multiple building blocks, such as convolution layers, pooling layers, and fully connected layers, and is designed to automatically and adaptively learn spatial hierarchies of features through a backpropagation algorithm.

<u>Tested Output:</u>



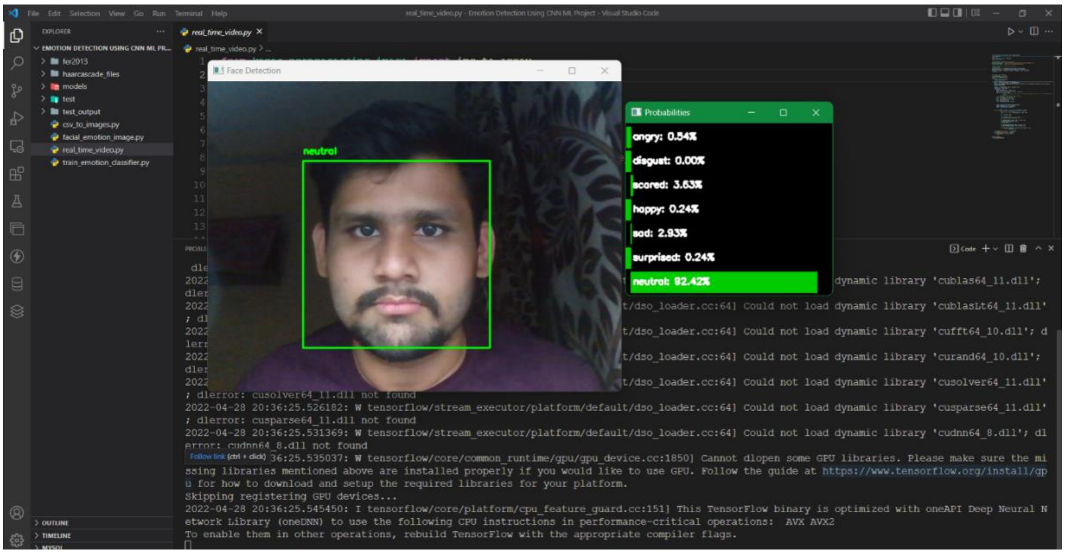


Figure 4.2

V. **COMPONENTS LIST AND SPECIFICATION**.

Hardware & Software Requirement:

A. Hardware Requirements System: Laptop Ram: 8 GB. (Minimum) Minimum i5 7 th gen Operating system: Windows 10.

B. Software Requirements

 Coding Language: Python 3.9.7 IDE: Visual Studio Code

VI. RESULT

In this chapter, the metrics used to evaluate model performance are defined. Then the best parameter values for each model are determined from the training results. These values are used to evaluate the accuracy and loss for CNN models 1 and 2. The results for these models are then compared and discussed.

6.1 Evaluation metrics

Accuracy, loss, precision, recall and F-score are the metrics used to measure model performance. These metrics are defined below.

Accuracy: Accuracy is given by

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}}$$

Loss: Categorical cross-entropy is used as the loss function and is given by

$$\text{Loss} = -\sum_{c=1}^{m} (y_{o,c} \log (p_{o,c}))$$

where y is a binary indicator (0 or 1), p is the predicted probability and m is the number of classes (happy, sad, neutral, fear, angry)

6.2 Confusion matrix:

The confusion matrix provides values for the four combinations of true and predicted values, True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN). Precision, recall and F-score are calculated using TP, FP, TN, FN. TP is the correct prediction of an emotion, FP is the incorrect prediction of an emotion, TN is the correct prediction of an incorrect emotion and FN is the incorrect prediction of an incorrect emotion. Consider an image from the happy class. The confusion matrix for this example is shown in below Figure. The red section has the TP value as the happy image is predicted to be happy. The blue section has FP values as the image is predicted to be sad, angry, neutral or fear. The yellow section has TN values as the image is not sad, angry, neutral or fear but the model predicted this. The green section has FN values as the image is not happy but was predicted to be happy.

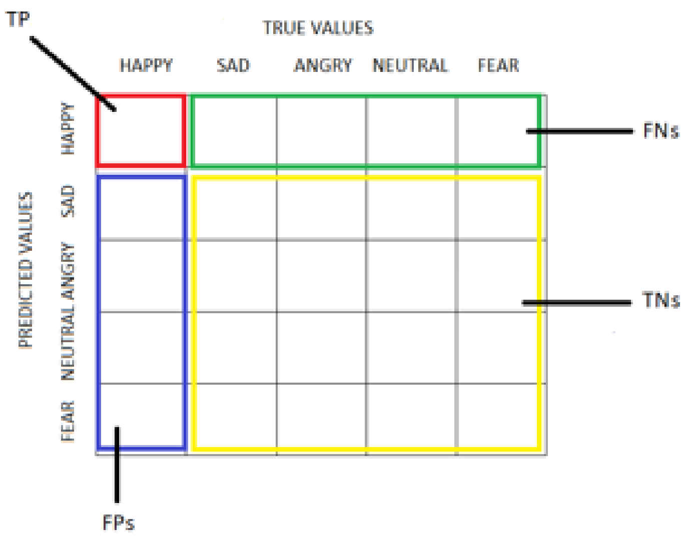


Figure 6.1

VII. CONCLUSION

So to conclude we can say that this project surely does the basic general task of Detecting Emotions if a face is given .It has been proved that self learning algorithm Convolution neural networks produces good results for naturalistic databases, also best fitted to reduce data over fitting and data imbalance. Along with that it finds various application areas like healthcare, virtual reality, robotics etc.

VIII. FUTURE SCOPE

This project has many further applications such as the machine should be able to recognize deeper emotions and recognize them even if a little bit shaky image is given. Also The project if properly maintained upgraded and if linked with proper hardware and other software devices can be used in various situations such as detecting if a person is drunk driving or not or even if someone is having suicidal thoughts or in some places when someone is being taken somewhere if they are nervous and are forcefully being taken etc. This can also be used in Biometric security and can find out if someone is being forced to unlock their device or even if someone is looking scared and we can set up prompts and alerts and if the user verify or any other verifying factors are found we can report it to the authorities. This can be used in case of Domestic Violence etc.

REFERENCES

1) https://ieeexplore.ieee.org/
2) https://www.researchgate.net
3) Exploiting multi-CNN features in CNN-RNN based Dimensional Emotion Recognition on the OMG in-the-wild Dataset Dimitrios Kollias and Stefanos Zafeiriou.
4) Real-time emotion recognition on mobile devices Denis Sokolov Mikhail Patkin
5) A Survey on Automatic Emotion Recognition Using Audio Big Data and Deep Learning Architectures Huijuan Zhao Ning Ye Ruchuan Wang