Facial Emotion Recognition

Gireesh kumar Muppalla, Ajay Kumar Kancharla,
Vishnumonish kankanala and Mohan Reddy Tamma

Department of Computing and Engineering,
University of Missouri-Kansas City
Kansas City, Missouri, USA

ABSTRACT Over the course of two decades, a substantial research effort was undertaken to create human-computer interaction. The main objective of this project is to combine facial emotion recognition and applications. The expression of the face is the process of nonverbal communication between verbs and nonverbs which displays the human's point of view and emotional display. In this project we have worked on the following topics Introduction to the method for recognizing facial emotions, implementation, a comparison of existing schemes for recognizing face expressions, and several different levels for automated systems for facial identification. The goal of our paper is to recognize faces in any image, extract facial expressions (eyes and lips), and bifurcate them into six emotions (happy, fear, anger, disgust, neutral, and sorrow). The Training data is skillfully filtered and processed before being described by a Support Vector Machine (SVM), enhanced victimization Grid Search. The most common method for interpreting human emotion is through facial expressions, IN general the human emotions are classified into two types of emotions happy feelings and negative emotions. There are four sorts of systems that are commonly used face detection, extraction, classification, and recognition. A overview of existing Facial Emotion Recognition stages, techniques, and data sets (FER). For so many years, FER has been recognized as an important aspect in the fields of computer vision and machine learning. Automated FERs are beneficial in a wide range of applications, including healthcare, education, criminal research, human-robot interaction (HRI), and so on.

I. INTRODUCTION

Emotional aspects have a huge impact on social intelligence, including understanding, communicating, and making decisions.
Emotion is important throughout a conversation. Vocal (audible) communication, facial expression, and emotional recognition can all be expressed in a variety of ways, both verbal and nonverbal. While just 7% of the message's power is delivered verbally, face expression has a voice contribution of 38% and a speaker contribution of 55%. Facial expression that is automated and efficient is also important in the connection between humans and technology. Facial expressions may be used to identify people in a variety of settings, from human facilities to healthcare operations.

The fundamental instrument for describing human emotion is facial expression.
Human emotions fluctuate dramatically during the day, maybe due to changes in their mental or

physical health. While human beings are full of varied emotions, modern psychology defines six fundamental face expressions like pleasure, grief, surprise, fright, disgust, and rage as universal emotions.

Face movement aids in recognizing people's emotions. Basic facial features include the brow, lips, nose, and eyes.

2. Problem Statement

In this project, we must accurately anticipate a person's emotions based on a grayscale image. The accuracy of every emotion displayed by the person is the major metric used in this study (a part of well classified images). This is reinforced with a confusion matrix, which determines certain emotions are more well-known than others.

Fig 1. Emotions to classify

3. Related Work

The theme of classifying photo emotions based on facial expressions has received a lot of attention. Based on our research, we have identified this paper, "Empathy: Neural Networks that Classify Facial Expressions." Journal of Cognitive Neuroscience [1]. Then apply a raw photo and Gabor filter to take various transformations and PCA, and finally a 3-layer neural network. Later we also found several papers such as the following [2],[3],[4] which utilized convolutional and deep neural network for classifying emotions. Majority of these papers deal with identifying emotions in video footage or doing similar tasks utilizing audiovisual data, although the majority of them do not use convolutional neural networks. Neural networks may be used to extract emotions from still photographs.

4. Solutions

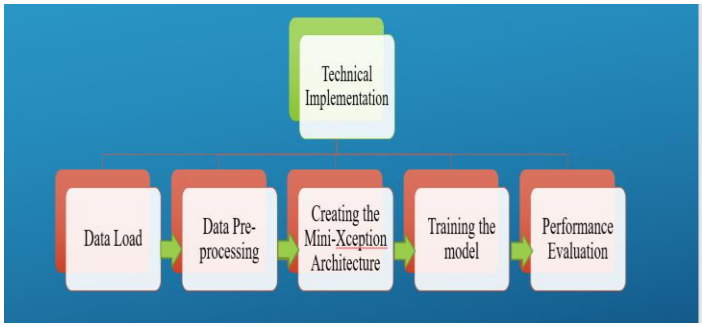

Fig 2. Project flow

We have chosen keras as our framework, Keras is an open-source software library for artificial neural networks that includes a Python interface. We have also used Flask as our Web framework along with that we have used several State of Art models as shown below out of all the models we used, we found Mini_Xception has performed better.

Model Used	Accuracy
VGG-16	63.19
ResNet-18	63.9
VGG-19	63.97
VGG-Face	64.15
Mini_Xception	**65.7**

Fig 3. Models Comparison

4.1Architecture

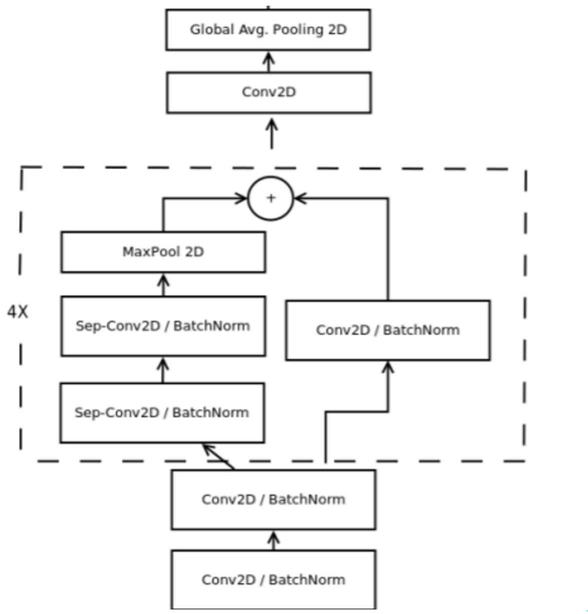

Fig 4. Architecture

Based on what we have learned in the class we have used Categorical Cross entropy function and optimizer Adam is a stochastic optimization approach that just requires first-order gradients and uses less memory.

5. Dataset Used

In this project we have used FER2013 as our dataset https://www.kaggle.com/msambare/fer2013.

6. Pre-Processing

In Pre-Processing we have converted RGB to grayscale and then we have done resizing later we used One hot encoding is used for allowing the representation of categorical data to be more accurate.

7. Model

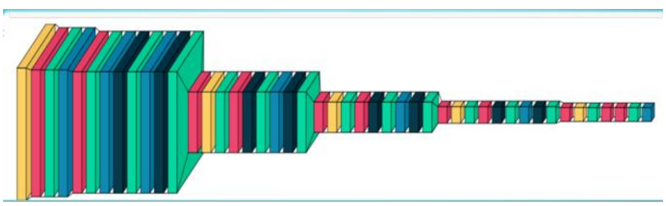

Fig6. Mini Xception model

We have used the Mini Xception State of Art model. After creating the model in keras as per the figure4 it was noticed the below no. of model parameters.

Total Params	58,423
Trainable params	56,951
Non-trainable params	1472

8. Post Processing

***Confusion Matrix**

0:'angry', 1:'disgust', 2:'scared', 3:'happy', 4:'sad', 5:'surprised', 6:'neutral.

We have trained the model on 36000 samples and the train to test ratio is 80:20, epoch as 110, with each epoch time approx. took 210 seconds on CPU.

After training process, we have evaluated the model performance with the test data above figure shows the confusion matrix for the same. The size of the test data on which we have evaluated the model performance is 7178.

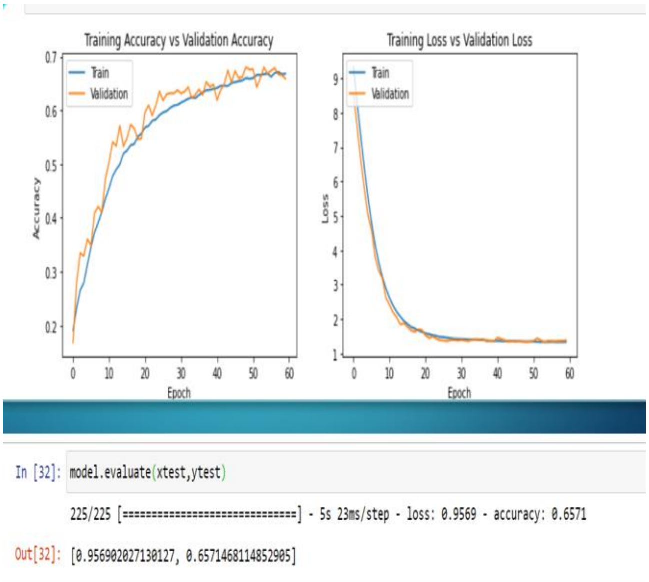

Fig7. Performance Evaluation

The Model accuracy is observed to be 65.71 percentage as shown in the above figure. Below are the results obtained.

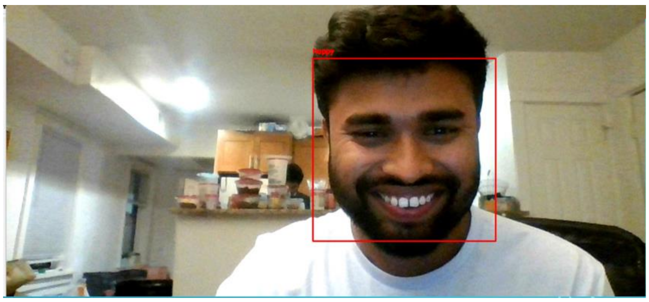

Fig8. Emotion recognition on Single Image.

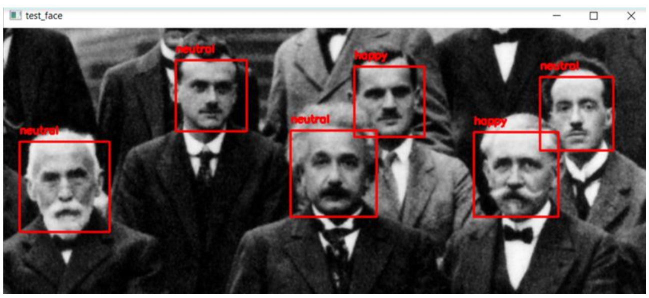

Fig9. Emotion recognition on multiple Images.

8. Applications

- Can be employed for employee wellness who are WORKING FROM HOME
- Psychometric examinations
- Deepfake Detection.

9. REFERENCES

[1] Matthew N Dailey, Garrison W Cottrell, Curtis Padgett, and Ralph Adolphs. Empath: A neural network that categorizes facial expressions. Journal of cognitive neuroscience, 14(8):1158–1173, 2002.

[2] Samira Ebrahimi Kahou, Christopher Pal, Xavier Bouthillier, Pierre Froumenty, C¸ aglar Gulc¸ehre, ¨ Roland Memisevic, Pascal Vincent, Aaron Courville, Yoshua Bengio, Raul Chandias Ferrari, et al. Combining modality specific deep neural networks for emotion recognition in video. In Proceedings of the 15th ACM on International conference on multimodal interaction, pages 543–550. ACM, 2013.

[3] Thai Hoang Le. Applying artificial neural networks for face recognition. Advances in Artificial Neural Systems, 2011:15, 2011.

[4] Yi Sun, Xiaogang Wang, and Xiaoou Tang. Deep learning face representation from predicting 10,000 classes. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 1891–1898, 2014.