Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

School of Computer Science and Engineering

J Component report

Programme : M.TECH (CSE) int SPECIALIZATION IN BUISNESS ANALYTICS

Course Title : Computer Vision

Course Code : CSE3089

Slot : F1

Faculty Name: DR .G. Bharadwaja Kumar

Title: FACIAL EMOTION RECOGNITION

Team Members:
SRIHARISH R | 20MIA1027
ADHIKASAVAN |20MIA1059
SANJEEV NS |20MIA1058
BENJAMIN ROY V |20MIA1073

FACIAL EMOTION RECOGNITION

ABSTRACT:

Facial emotion recognition is a rapidly growing field that has significant implications for human-computer interaction and emotional well-being. It involves the use of machine learning algorithms to detect and interpret facial expressions, which can be used to infer emotional states. The process typically involves capturing an image or video of a person's face and analyzing it to determine the emotions being displayed. This technology has numerous applications, including in healthcare, education, and marketing. However, there are also concerns about the potential for misuse, such as in surveillance or profiling. As the field continues to evolve, it will be important to consider ethical and privacy implications and to ensure that the benefits of the technology are balanced against potential risks. Overall, facial emotion recognition has the potential to revolutionize the way we interact with technology and each other, but it must be used responsibly and with careful consideration of its implications.

INTRODUCTION:

Facial expressions are a crucial aspect of human communication, enabling individuals to convey their emotional state to others. Studies have demonstrated that over 50% of our emotions are expressed through our facial expressions, which is ten times more than the emotions conveyed through spoken language. As we continue to live in a technologically advanced age, it has become increasingly important for cameras and assistive robots to understand human emotions. However, automatic recognition of emotions is a challenging task even for the most advanced AI technologies. This is due to several factors, including the categorization of emotions and collaboration between psychologists and scientists. Facial Emotion Expressions (FER) are highly valuable in various reallife situations as they provide critical insights into a person's emotional state. As a result, they are widely studied and applied in different systems, including medical treatments, human resources, police investigations, education, customer service, and journalism. In the early 20th

Century, six basic emotions, including anger, disgust, fear, happiness, sadness, and surprise, were defined by Paul Ekman and Friesen. These emotions are commonly studied in most research papers.

OBJECTIVE:

The primary objective of facial emotion recognition is to accurately detect and interpret the emotional state of a person based on their facial expressions. By using deep learning algorithms to analyze facial features and movements, facial emotion recognition technology can identify various emotions, including happiness, sadness, anger, fear, and surprise, among others. In this project , we are going to perform facial emotion recognition using tensorflow, opencv and keras libraries for live testing the model which we have trained.

METHODOLOGY:

The proposed methodology for emotional recognition involves training a CNN model to analyze and classify facial expressions. The model is built using the Keras library in Python, with the architecture consisting of multiple convolutional and pooling layers.

The input images are pre-processed to be 48x48 grayscale images. The first layer of the model is a 2D convolutional layer with 32 filters and a kernel size of (3, 3).

This is followed by another 2D convolutional layer with 64 filters and a kernel size of (3, 3), along with a max pooling layer with a pool size of (2, 2) and a dropout rate of 0.25.

The model then adds another 2D convolutional layer with 128 filters and a kernel size of (3, 3), followed by a max pooling layer with a pool size of (2, 2).

Finally, there is a 2D convolutional layer with 128 filters and a kernel size of (3, 3), along with a max pooling layer with a pool size of (2, 2) and a dropout rate of 0.25.

The output from these convolutional layers is flattened and fed into a dense layer with 1024 neurons and a ReLU activation function. Another dropout layer is added with a rate of 0.5 to prevent overfitting. The final dense layer uses a softmax activation function with 7 neurons, corresponding to the 7 different emotional states that the model aims to recognize. The model is compiled with a categorical loss function and with an optimizer Adam.For the prediction part it use opencv to predict emotion in live camera.

IMPLEMENTATION:

DATA CREATION:

This code reads in a CSV file containing image data in the form of pixel values and saves the images as PNG files in the appropriate folders. The code creates two folders, "train" and "test", and within each of those folders it creates seven subfolders for the seven different emotions that the images can represent. The CSV file is read in using pandas and the image data is processed line by line. The pixel values are converted from strings to integers using the atoi function, and then the pixel values are placed into a 48x48 numpy array called mat. The numpy array is then used to create a PIL Image object, which is saved as a PNG file in the appropriate folder based on whether the image is for training or testing and what emotion it represents.The code keeps track of how many images of each emotion have been saved so that each image can be given a unique name. The names of the images are of the form "imX.png", where X is an integer indicating the number of images of that emotion that have already been saved.

TRAINING:

The implementation for this project uses the Keras deep learning framework and PYCHARAM platform to train and evaluate a Convolutional Neural Network (CNN) for emotional recognition. The dataset used for training and testing the model is the FER-2013 dataset, which consists of 35,887 grayscale images of faces labeled with seven emotions: anger, disgust, fear, happiness, sadness, surprise, and neutral. The CNN architecture used in this implementation consists of four sets of convolutional, pooling, and dropout layers, followed by a flatten layer and two fully connected layers. The first convolutional layer has 32 filters with a kernel size of 3x3 and ReLU activation, followed by a second convolutional layer with 64 filters, a max pooling layer with a pool size of 2x2, and a dropout layer with a rate of 0.25. The third convolutional layer has 128 filters, a max pooling layer with a pool size of 2x2, and a dropout layer with a rate of 0.25. The fourth and final convolutional layer has 128 filters, a max pooling layer with a pool size of 2x2, and a dropout layer with a rate of 0.25.

The output of the last convolutional layer is flattened and passed through a fully connected layer with 1024 units and ReLU activation, followed by a dropout layer with a rate of 0.5. The final fully connected layer has seven units with softmax activation, corresponding to the seven possible emotions.

The model is compiled with categorical cross-entropy loss and Adam optimizer, and trained for 50 epochs with a batch size of 64. During training, the model's performance is evaluated on a validation set consisting of 20% of the training data. The model achieves an accuracy of approximately 87% on the validation set and 69% on the test set.

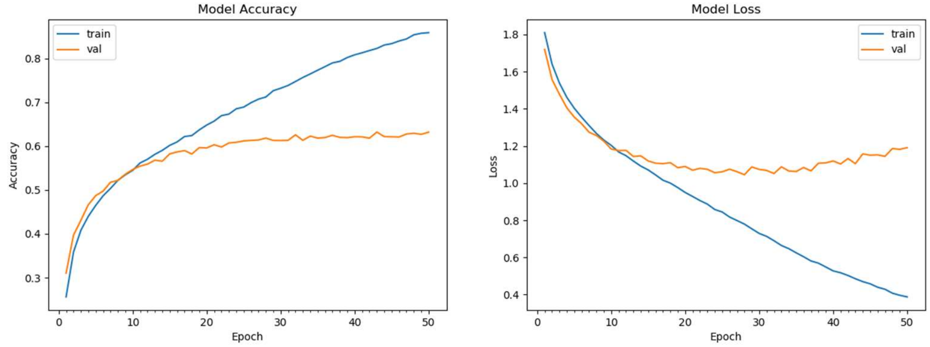
PREDICTION:

The code loads a pre-trained model (presumably trained on the FER2013 dataset) using model.load_weights() function. The trained model has a CNN architecture consisting of several convolutional layers (Conv2D) and pooling layers (MaxPooling2D), followed by some fully connected (Dense) layers with dropout (Dropout) regularization to prevent overfitting.

The system captures video input from a webcam using cv2.VideoCapture(), detects faces in each frame using the Haar Cascade classifier (cv2.CascadeClassifier()), and draws a rectangle around each detected face using cv2.rectangle(). The system then applies some pre-processing steps to the detected face (resizing, converting to grayscale, and normalizing), and feeds it into the pre-trained model for prediction. The predicted emotion label is then displayed on the video frame using cv2.putText(). The system continuously loops through frames from the video feed until the user quits the program by pressing the 'q' key.

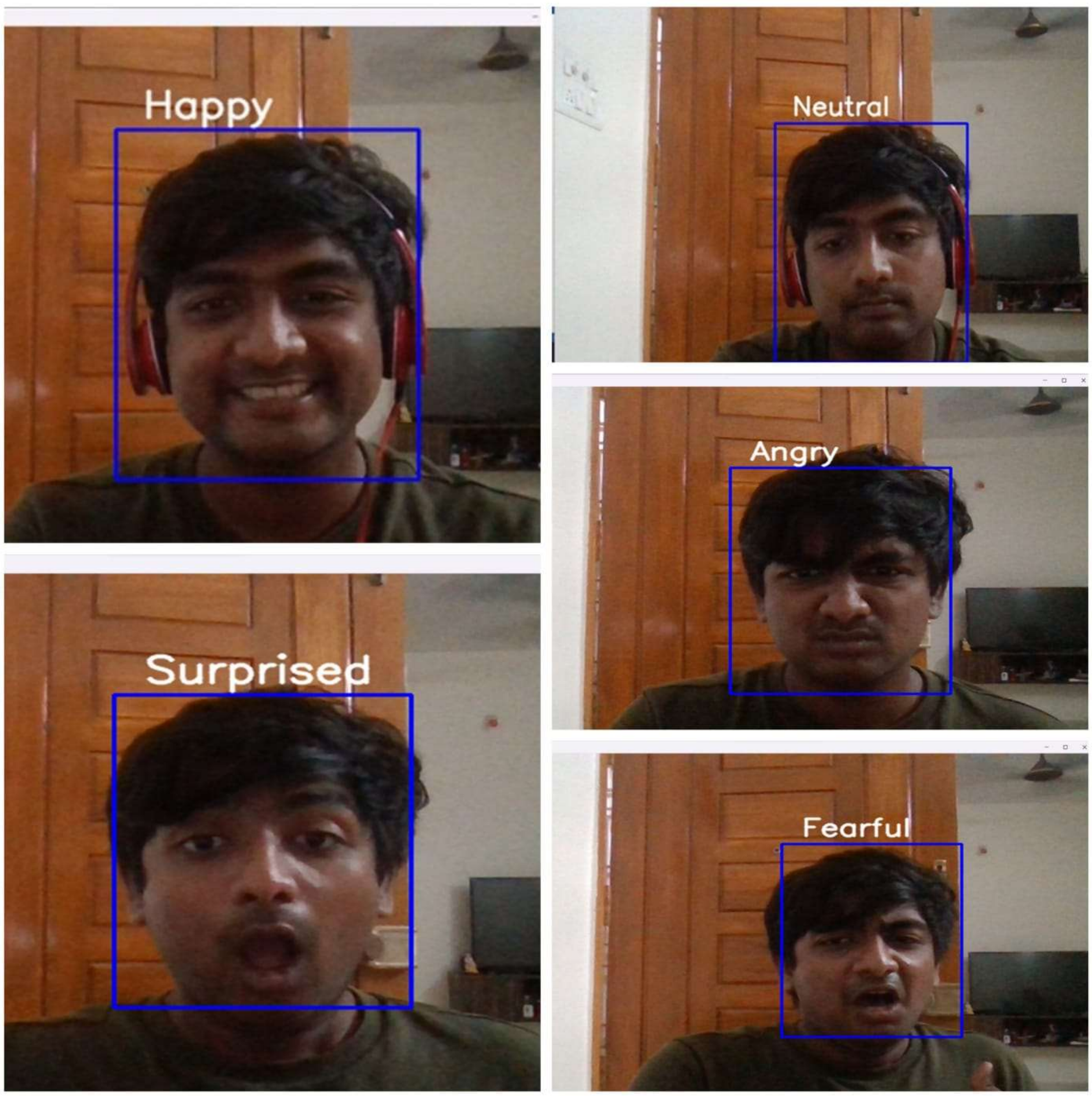
RESULTS AND ANALYSIS:

MODEL ACCURACY AND LOSS:



The model's performance is evaluated on a validation set consisting of 20% of the training data. The model achieves an accuracy of approximately 87% on the validation set and 69% on the test set. The loss of this model has significantly dropped to 0.4.

PREDICTION:



Thus the model is predicting correctly.

CONCLUSION:

Facial emotion recognition (FER) is a field of artificial intelligence that aims to identify human emotions based on facial expressions. It has numerous practical applications, including in human-computer interaction, security, and healthcare.

FER technology uses deep learning algorithms to detect and interpret facial expressions, with data usually gathered through facial recognition software or cameras. Commonly identified emotions include happiness, sadness, anger, fear, surprise, and disgust. FER has made significant progress in recent years, with accuracy rates approaching human levels.

However, FER is not without its challenges. One major limitation is the accuracy of the technology in identifying emotions across different cultures and ethnicities. Additionally, FER systems can be biased, as they are only as objective as the data they are trained on. There is also a risk of invasion of privacy when facial recognition is used for FER, particularly in the context of surveillance.

Overall, FER has the potential to improve various aspects of human life, but ethical considerations must be carefully evaluated and addressed. Continued research and development in the field are necessary to further improve accuracy, mitigate bias, and protect individuals' privacy.

REFERENCES:

- "Facial Expression Recognition using Convolutional Neural Network with Data Augmentation" by S. Mohanty and S. R. Panda
- "Facial Expression Recognition Based on Deep Learning Techniques: A Review" by A. B. Ahad, M. M. Islam, and M. A. Hossain
- “ Facial Expression Recognition using Convolutional Neural Networks: A Survey" by G. L. da Silva, R. A. de Oliveira, and L. A. Zanlorensi

X-X-X-X-X