KARNATAK LAW SOCIETY'S
GOGTE INSTITUTE OF TECHNOLOGY
UDYAMBAG, BELAGAVI-590008
(An Autonomous Institution under Visvesvaraya Technological University, Belagavi) **(APPROVED BY AICTE, NEW DELHI)**

Department of Computer Science Engineering

Course Activity Report

Machine Learning Laboratory

Submitted in the partial fulfillment for the academic requirement of
6th Semester B.E.

TITLE: Detect Emotion from Face using Deep Learning

Model Submitted by

SL.No.	Batch Members Names	USN
3	Tarun Bagewadi	2GI18CS170

Problem Statement:
Detect Emotion from Face using Deep Learning Model.

Abstract:
Automatic emotion recognition based on facial expression is an interesting research field, which has been presented and applied in several areas such as safety, health and in human machine interfaces. Researchers in this field are interested in developing techniques to interpret, code facial expressions and extract these features to have a better prediction by computer. With the remarkable success of deep learning, the different types of architectures of this technique are exploited to achieve a better performance. The purpose of this project is to make a study on recent works on automatic facial emotion recognition FER via deep learning.

Introduction:
Automatic emotion recognition is a large and important research area that addresses two different subjects, which are psychological human emotion recognition and artificial intelligence (AI). The emotional state of humans can obtain from verbal and non-verbal information captured by the various sensors, for example from facial changes, tone of voice and physiological signals. In 1967, Mehrabian showed that 55% of emotional information were visual, 38% vocal and 7% verbal. Face changes during a communication are the first signs that transmit the emotional state, which is why most researchers are very interested by this modality.
Extracting features from one face to another is a difficult and sensitive task to have a better classification. In 1978 Ekman and Friesen are among the first scientific interested in facial expression which are developed FACS (Facial Action Coding System) in which facial movements are described by Action Units AUs, they are broken down the human face into 46 AUs action units each AU is coded with one or more facial muscles.

The automatic FER is the most studied by researchers compared to other modalities to statistics, but it is task that is not easy because each person presents his emotion by his way. Several obstacles and challenges are present in this area that one should not neglect like the variation of head pose, luminosity, age, gender and the background, as well as the problem of occlusion caused by Sunglasses, scarf, skin illness…etc.

Several traditional methods exist are used for the extraction facial features such as geometric and texture features; for example, local binary patterns LBP, facial action units FAC, local directional patterns LDA, Gabor wavelet. In recent years, deep learning has been very successful and efficient approach thanks to the result obtained with its architectures which allow the automatic extraction of features and classification such as the convolutional neural network CNN and the recurrent neural network RNN; here what prompted researchers to start using this technique to recognize human emotions. Several efforts are made by researchers on the development of deep neural network architectures, which produce very satisfactory results in this area.

Facial Available Databases:

One of the success factors of deep learning is the training the neuron network with examples, several FER databases now available to researchers to accomplish this task, each one different from the others in term of the number and size of images and videos, variations of the illumination, population, and face pose. Some presented in the table in which we will note its presence in the works cited in the following section.

Databases	Descriptions	Emotions
MultiPie [10]	More than 750,000 images captured by 15 view and 19 illumination conditions	Anger, Disgust, Neutral, Happy, Squint, Scream, Surprise
MMI [11]	2900 videos, indicate the neutral, onset, apex and offset	Six basic emotions and neutral
GEMEP FERA [12]	289 images sequences	Anger, Fear, Sadness, Relief, Happy
SFEW [13]	700 images with different ages, occlusion, illumination and head pose.	Six basic emotions and neutral
CK+ [14]	593 videos for posed and non-posed expressions	Six basic emotions, contempt and neutral
FER2013 [15]	35,887 grayscale images collect from google image search	Six basic emotions and neutral
JAFFE [16]	213 grayscale images posed by 10 Japanese females	Six basic emotions and neural
BU-3DFE [17]	2500 3D facial images captured on two view -45°, +45°	Six basic emotions and neutral
CASME II [18]	247 micro-expressions sequences	Happy, Disgust, Surprise, Regression and others
Oulu-CASIA [19]	2880 videos captured in three different illumination conditions	Six basic emotions
AffectNet [20]	More than 440,000 images collected from the internet	Six basic emotions and neutral
RAFD-DB [21]	30000 images from real world	Six basic emotions and neutral

Facial Emotion Recognition using Deep Learning:

Despite the notable success of traditional facial recognition methods through the extracted of handcrafted features, over the past decade researchers have directed to the deep learning approach due to its high automatic recognition capacity. In this context, we will present some recent studies in FER, which show proposed methods of deep learning to obtain better detection. Train and test on several static or sequential databases.

Mollahosseini propose deep CNN for FER across several available databases. After extracting the facial landmarks from the data, the images reduced to 48x 48 pixels. Then, they applied the augmentation data technique. The architecture used consist of two convolution-pooling layers, then add two inception styles modules, which contains convolutional layers size 1x1, 3x3 and 5x5. They present the ability to use technique the network-in-network, which allow increasing local performance due to the convolution layers applied locally, and this technique also make it possible to reduce the over-fitting problem.
Lopes Studied the impact of data pre-processing before the training the network to have a better emotion classification. Data augmentation, rotation correction, cropping, down sampling with 32x32 pixels and intensity normalization are the steps that were applied before CNN, which consist of two convolution-pooling layers ending with two fully connected with 256 and 7 neurons. The best weight gained at the training stage are used at the test stage. This experience was evaluated in three accessible databases: CK+, JAFFE, BU-3DFE. Researchers shows that combining all these pre-processing steps is more effective than applying them separately.

These pre-processing techniques also implemented by Mohammadpour. They propose a novel CNN for detecting AUs of the face. For the network, they use two convolution layers, each followed by a max pooling and ending with two fully connected layers that indicate the numbers of AUs activated.

In 2018, for the disappearance or explosion gradient problem Cai propose a novel architecture CNN with Sparse Batch normalization SBP. The property of this network is to use two convolution layers successive at the beginning, followed by max pooling then SBP, and to reduce the over-fitting problem, the dropout applied in the middle of three fully connected. For the facial occlusion problem Li present a new method of CNN, firstly the data introduced into VGGNet network, then they apply the technique of CNN with attention mechanism ACNN. This architecture trained and tested in three large databases FED-RO, RAF-DB and AffectNet.

Source Code (test.py):

```
from keras.models import load_model
from time import sleep
from keras.preprocessing.image import img_to_array
from keras.preprocessing import image
import cv2
import numpy as np

face_classifier =
cv2.CascadeClassifier('./haarcascade_frontalface_default.xml') classifier
=load_model('./Emotion_Detection.h5')
class_labels = ['Angry','Happy','Neutral','Sad','Surprise']

cap = cv2.VideoCapture(0)

while True:
 ret, frame = cap.read()
 labels = []
 gray = cv2.cvtColor(frame,cv2.COLOR_BGR2GRAY)
 faces = face_classifier.detectMultiScale(gray,1.3,5)

 for (x,y,w,h) in faces:
 cv2.rectangle(frame,(x,y),(x+w,y+h),(255,0,0),2)
 roi_gray = gray[y:y+h,x:x+w]
 roi_gray = cv2.resize(roi_gray,(48,48),interpolation=cv2.INTER_AREA)


 if np.sum([roi_gray])!=0:
 roi = roi_gray.astype('float')/255.0
 roi = img_to_array(roi)
 roi = np.expand_dims(roi,axis=0)

            preds = classifier.predict(roi)[0]
 print("\nprediction = ",preds)
 label=class_labels[preds.argmax()]
```

```
 print("\nprediction max = ",preds.argmax())
 print("\nlabel = ",label)
 label_position = (x,y)

cv2.putText(frame,label,label_position,cv2.FONT_HERSHEY_SIMPLEX,2,(0,2
5 5,0),3)
 else:
 cv2.putText(frame,'No Face
Found',(20,60),cv2.FONT_HERSHEY_SIMPLEX,2,(0,255,0),3
)  print("\n\n")
 cv2.imshow('Emotion Detector',frame)
 if cv2.waitKey(1) & 0xFF == ord('q'):
 break

cap.release()
cv2.destroyAllWindows()
```

Source Code (train.py):

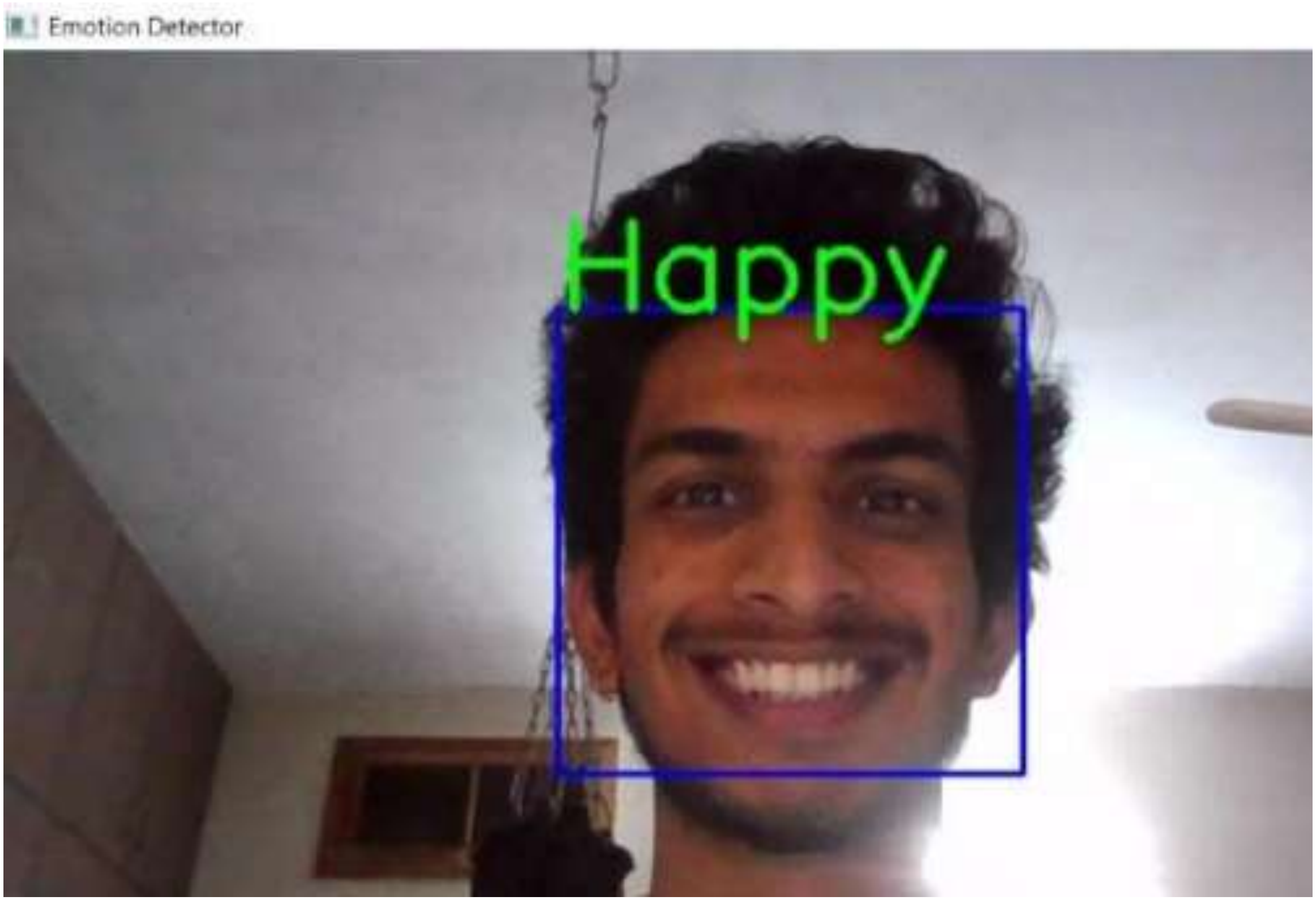
```
from keras.applications import MobileNet
from keras.models import Sequential,Model
from keras.layers import
Dense,Dropout,Activation,Flatten,GlobalAveragePooling2D
from keras.layers import
Conv2D,MaxPooling2D,ZeroPadding2D from
keras.layers.normalization import BatchNormalization
from keras.preprocessing.image import ImageDataGenerator

img_rows,img_cols = 224,224

MobileNet =
MobileNet(weights='imagenet',include_top=False,input_shape=(img_rows,img_c
o ls,3))

for layer in MobileNet.layers:
```

```
 layer.trainable = True
for (i,layer) in enumerate(MobileNet.layers):
 print(str(i),layer.__class__.__name__,layer.trainable)

def addTopModelMobileNet(bottom_model, num_classes):
 """creates the top or head of the model that will be
 placed ontop of the bottom layers"""

 top_model = bottom_model.output
 top_model = GlobalAveragePooling2D()(top_model)
 top_model = Dense(1024,activation='relu')(top_model)

 top_model = Dense(1024,activation='relu')(top_model)
 top_model = Dense(512,activation='relu')(top_model)
 top_model = Dense(num_classes,activation='softmax')(top_model)

return top_model

num_classes = 5

FC_Head = addTopModelMobileNet(MobileNet, num_classes)

model = Model(inputs = MobileNet.input, outputs = FC_Head)

print(model.summary())

train_data_dir = '/Users/durgeshthakur/Deep Learning Stuff/Emotion
Classification/fer2013/train'
validation_data_dir = '/Users/durgeshthakur/Deep Learning Stuff/Emotion
Classification/fer2013/validation'

train_datagen = ImageDataGenerator(
 rescale=1./255,
 rotation_range=30,
 width_shift_range=0.3,
```

```
 height_shift_range=0.3,
 horizontal_flip=True,
 fill_mode='nearest'
 )

validation_datagen =

ImageDataGenerator(rescale=1./255) batch_size = 32

train_generator = train_datagen.flow_from_directory(
 train_data_dir,
 target_size = (img_rows,img_cols),
 batch_size = batch_size,
 class_mode = 'categorical'
 )

validation_generator =
validation_datagen.flow_from_directory(  validation_data_dir,
 target_size=(img_rows,img_cols),
 batch_size=batch_size,
 class_mode='categorical')

from keras.optimizers import RMSprop,Adam
from keras.callbacks import ModelCheckpoint,EarlyStopping,ReduceLROnPlateau

checkpoint = ModelCheckpoint(
 'emotion_face_mobilNet.h5',
 monitor='val_loss',
 mode='min',
 save_best_only=True,
 verbose=1)

earlystop = EarlyStopping(
 monitor='val_loss',
 min_delta=0,
```

```
 patience=10,
 verbose=1,restore_best_weights=True)

learning_rate_reduction = ReduceLROnPlateau(monitor='val_acc',
patience=5,
                                            verbose=1,
                                            factor=0.2,
                                            min_lr=0.0001)
callbacks = [earlystop,checkpoint,learning_rate_reduction]

model.compile(loss='categorical_crossentropy'
,  optimizer=Adam(lr=0.001),
 metrics=['accuracy']
 )

nb_train_samples = 24176
nb_validation_samples = 3006

epochs = 25

history = model.fit_generator(
 train_generator,
 steps_per_epoch=nb_train_samples//batch_size,
epochs=epochs,
 callbacks=callbacks,
 validation_data=validation_generator,
 validation_steps=nb_validation_samples//batch_size)
```

Output:



Conclusion:

FER are one of the most important ways of providing information about the emotional state, but they are always limited by learning only the six-basic emotion plus neutral. It conflicts with what is present in everyday life, which has emotions that are more complex. This will push researchers in the future work to build larger databases and create powerful deep learning architectures to recognize all basic and secondary emotions.

References:

www.sciencedirect.com