DETERMINING MOOD FROM FACIAL EXPRESSIONS

A Project Report submitted in partial fulfilment of the requirements for the award of the degree of

Bachelor of Technology

in

Computer Science and Engineering

By

Raman Shukla (181500554)
Ritika Agrawal (181500581)
Vanshika Mahle (181500781)
Vimal Mishra (181500792)

Under the Guidance of
Mrs. Gunjan Bhardwaj

Department of Computer Engineering & Applications

Institute of Engineering & Technology

GLA University

Mathura- 281406, INDIA

Dec, 2020

Department of Computer Engineering and Applications
GLA University, 17 km. Stone NH#2, Mathura-Delhi Road,
Chaumuha, Mathura – 281406 U.P (India)

DECLARATION

We hereby declare that the work which is being presented in the B.Tech. Project **"Determining mood from Facial Expressions"**, in partial fulfillment of the requirements for the award of the ***Bachelor of Technology*** in Computer Science and Engineering and submitted to the Department of Computer Engineering and Applications of GLA University, Mathura, is an authentic record of our own work carried under the supervision of Mrs. Gunjan Bhardwaj.

The contents of this project report, in full or in parts, have not been submitted to any other Institute or University for the award of any degree.

Sign ______________________

Name of Candidate: Raman Shukla
University Roll No.: 181500554

Sign ______________________

Name of Candidate: Ritika Agrawal
University Roll No.: 181500581

Sign ______________________

Name of Candidate: Vanshika Mahle
University Roll No.: 181500781

Sign ______________________

Name of Candidate: Vimal Mishra
University Roll No.: 181500792

CERTIFICATE

This is to certify that the above statements made by the candidate are correct to the best of my/our knowledge and belief.

Supervisor
Mrs. Gunjan Bhardwaj
Designation of Supervisor

Project Coordinator	**Program Coordinator**
(Mr. Mayank Srivastava)	(Dr Anant Ram)

Date:

ACKNOWLEDGEMENT

Presentation inspiration and motivation have always played a key role in the success of any venture.

This project is itself an acknowledgement for all those people who have given us their heartfelt co-operation in making this project a grand success.

We extend our sincere thanks to ***Mrs. Gunjan Bhardwaj,*** for providing valuable guidance at every stage of this project work. We are profoundly grateful towards the unmatched services rendered by her.

This project cannot be completed without the effort and co-operation from our group members, Raman, Ritika, Vanshika, Vimal.

Last but not the least, we would like to express our deep sense of gratitude and earnest thanks giving to our dear parents for their moral support and heartfelt cooperation in doing the main project.

Group Members:

Raman Shukla (181500554)
Ritika Agrawal (181500581)
Vanshika Mahle (181500781)
Vimal Mishra (181500792)

ABSTRACT

Facial expressions are a form of non-verbal communication that conveys emotional state of a person. The Facial Expression Recognition (FER) system is the process of identifying the emotional state of a person. Facial Expression Recognition usually performed in four-stages consisting of pre-processing, face detection, feature extraction, and expression classification. FER can be widely applied to various research areas, such as mental diseases diagnosis and human social/physiological interaction detection. With the emerging advanced technologies in hardware and sensors, FER systems have been developed to support real-world application scenes, instead of laboratory environments. FER is an important and promising field of computer vision and artificial intelligence.

In this project, we use Convolution Neural Network (CNN), a deep learning algorithm to classify each face based on the emotion shown in the facial expression into one of seven categories (0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise, 6=Neutral). Kaggle facial expression dataset (fer2013) with seven facial expression labels as happy, sad, surprise, fear, anger, disgust, and neutral is used in this project.
Convolutional Neural Networks achieve better accuracy with big data. The proposed method achieves 58.12% accuracy and 0.68 precision on training dataset.
From the results our project it can be concluded that happiness and surprise can be easily expressed and identified than the other emotions. Sadness, quite a common emotion, may end up being misclassified most often and the most notable ones may be anger, disgust, contempt and fear.

CONTENTS

List of Figures

List of Tables

INTRODUCTION

Overview and Motivation

Human facial expressions can be easily classified into 7 basic emotions: happy, sad, surprise, fear, anger, disgust, and neutral. Our facial emotions are expressed through activation of specific sets of facial muscles.

Image processing is the field of signal processing where both the input and output signals are images. One of the most important application of Image processing is Facial expression recognition. Our emotion is revealed by the expressions in our face. Facial Expressions plays an important role in interpersonal communication. Facial expression is a nonverbal scientific gesture which gets expressed in our face as per our emotions. Automatic recognition of facial expression plays an important role in artificial intelligence and robotics and thus it is a need of the generation. Some application related to this include Personal identification and Access control, Videophone and Teleconferencing, Forensic application, Human-Computer Interaction, Automated Surveillance, Cosmetology and so on.
FER systems have broad applications in various areas, such as computer interactions, health-care systems and social marketing.

We have also been motivated observing the benefits of physically handicapped people like deaf and dumb. But if any normal human being or an automated system can understand their needs by observing their facial expression then it becomes a lot easier for them to make the fellow human or automated system understand their needs.

Objective

Our aim is to develop a method of face mood detection from facial expressions that is fast, robust, reasonably simple and accurate with a relatively simple and easy to understand Machine Learning algorithms and techniques. The key elements of face are considered for detection of face and prediction of expressions or emotions of face. For detection and classification of different classes of facial expressions, machine learning algorithms are used by training of different set of images.

The scope of this system is to tackle with the problems that can arise in day to day life. Some of the scopes are:

1. The system can be used to detect and track a user's state of mind.
2. The system can be used in mini-marts, shopping center to view the feedback of the customers to enhance the business,
3. The system can be installed at busy places like airport, railway station or bus station for detecting human faces and facial expressions of each person. If there are any faces that appeared suspicious like angry or fearful, the system might set an internal alarm.
4. The system can also be used for educational purpose such as one can get feedback on how the student is reacting during the class.
5. This system can be used for lie detection amongst criminal suspects during interrogation.
6. This system can help people in emotion related -research to improve the processing of emotion data.

Summary of Similar Application

Face detection uses computer learning to detect the location of any faces within an image. Image processing consists of scaling and image rendering to prepare the face for identification. Face identification uses mathematical techniques on the pixel values or features in the facial area of an image to determine who the face belongs to. The most useful applications contain crowd surveillance, video content indexing, personal identification (ex. driver's license), mug shots matching, entrance security, etc.

In recent years, much research has been done on machine recognition of human facial expressions Face recognition has been studied extensively for more than 40 years. Face recognition is a technology which recognizes the human by his/her faces image. Recent advances in image analysis and pattern recognition open up the possibility of automatic detection and classification of emotional and conversational facial signals.

Before a facial expression can be analyzed, the face must be detected in a scene. Next is to devise mechanisms for extracting the facial expression information from the observed facial image or image sequence. In the case of static images, the process of extracting the facial expression information is referred to as localizing the face and its features in the scene. In the case of facial image sequences, this process is referred to as tracking the face and its features in the scene. At this point, a clear distinction should be made between two terms, namely, facial features and face model features. The facial features are the prominent features of the face eyebrows, eyes, nose, mouth, and chin. The face model features are the features used to represent (model) the face. The face can be represented in various ways, e.g., as a whole unit (holistic representation), as a set of features (analytic representation) or as a combination of these (hybrid approach). The applied face representation and the kind of input images determine the choice of mechanisms for automatic extraction of facial expression information.

Issues and Challenges

Face recognition being the most important biometric trait it still **faces** many **challenges**, like pose variation, illumination variation etc. When such variations are present in both pose and illumination, all the algorithms are greatly affected by these variations and their performance gets degraded.

In recent years, facial expression analysis and recognition (FER) have emerged as an active research topic with applications in several different areas, including the human-computer interaction domain. Solutions based on 2D models are not entirely satisfactory for real-world applications, as they present some problems of pose variations and illumination related to the nature of the data. Thanks to technological development, 3D facial data, both still images and video sequences, have become increasingly used to improve the accuracy of FER systems.

Despite the advance in 3D algorithms, these solutions still have some drawbacks that make pure three-dimensional techniques convenient only for a set of specific applications; a viable solution to overcome such limitations is adopting a multimodal 2D+3D analysis. In this paper, we analyze the limits and strengths of traditional and deep-learning FER techniques, intending to provide the research community an overview of the results obtained looking to the next future. Furthermore, we describe in detail the most used databases to address the problem of facial expressions and emotions, highlighting the results obtained by the various authors. The different techniques used are compared, and some conclusions are drawn concerning the best recognition rates achieved.

Organization of the Project

Humans show a great deal of variability in their abilities to recognize emotion. A key point to keep in mind when learning about automated emotion recognition is that there are several sources of "ground truth," or truth about what the real emotion is. Suppose we are trying to recognize the emotions of Alex. One source is "what would most people say that Alex is feeling?" In this case, the 'truth' may not correspond to what Alex feels, but may correspond to what most people would say it looks like Alex feels. For example, Alex may actually feel sad, but he puts on a big smile and then most people say he looks happy. If an automated method achieves the same results as a group of observers it may be considered accurate, even if it does not actually measure what Alex truly feels. Another source of 'truth' is to ask Alex what he truly feels. This works if Alex has a good sense of his internal state, and wants to tell you what it is, and is capable of putting it accurately into words or a number. However, some people are alexithymic and do not have a good sense of their internal feelings, or they are not able to communicate them accurately with words and numbers. In general, getting to the truth of what emotion is actually present can take some work, can vary depending on the criteria that are selected, and will usually involve maintaining some level of uncertainty.

Research has shown that over 90 percent of our communication can be non-verbal, but technology has struggled to keep up, and traditional code is generally bad at understanding our intonations and intentions. But emotion recognition—also called Affective Computing—is becoming accessible to more types of developers.

Facial emotion recognition is the process of detecting human emotions from facial expressions. The human brain recognizes emotions automatically, and software has now been developed that can recognize emotions as well. This technology is becoming more accurate all the time, and will eventually be able to read emotions as well as our brains do.

AI can detect emotions by learning what each facial expression means and applying that knowledge to the new information presented to it. Emotional artificial intelligence, or emotion AI, is a technology that is capable of reading, imitating, interpreting, and responding to human facial expressions and emotions.

SOFTWARE REQUIREMENT ANALYSIS

Problem Statement

Human facial expressions can be easily classified into 7 basic emotions: happy, sad, surprise, fear, anger, disgust, and neutral. Our facial emotions are expressed through activation of specific sets of facial muscles. These sometimes subtle, yet complex, signals in an expression often contain an abundant amount of information about our state of mind. Through facial emotion recognition, we are able to measure the effects that content and services have on the audience/users through an easy and low-cost procedure. For example, retailers may use these metrics to evaluate customer interest. Healthcare providers can provide better service by using additional information about patients' emotional state during treatment. Entertainment producers can monitor audience engagement in events to consistently create desired content.

Humans are well-trained in reading the emotions of others, in fact, at just 14 months old, babies can already tell the difference between happy and sad. But can computers do a better job than us in accessing emotional states? To answer the question, We designed a deep learning neural network that gives machines the ability to make inferences about our emotional states. In other words, we give them eyes to see what we can see.

Technical Feasibility

System Specifications:

- Core i5 processor
- 4GB RAM (minimum)
- 512 GB Hard Disk

Software Specifications:

- Python modules - Numpy, Pandas, Tensorflow, keras
- Online Python IDE – Jupyter / Spyder / Google Colab

SYSTEM DESIGN

System design shows the overall design of system.

Data flow diagram

A data flow diagram (DFD) is a graphical representation of the "flow" of data through an information system, modelling its process aspects. A DFD is often used as a preliminary step to create an overview of the system without going into great detail, which can later be elaborated. DFDs can also be used for the visualization of data processing.

Level-0

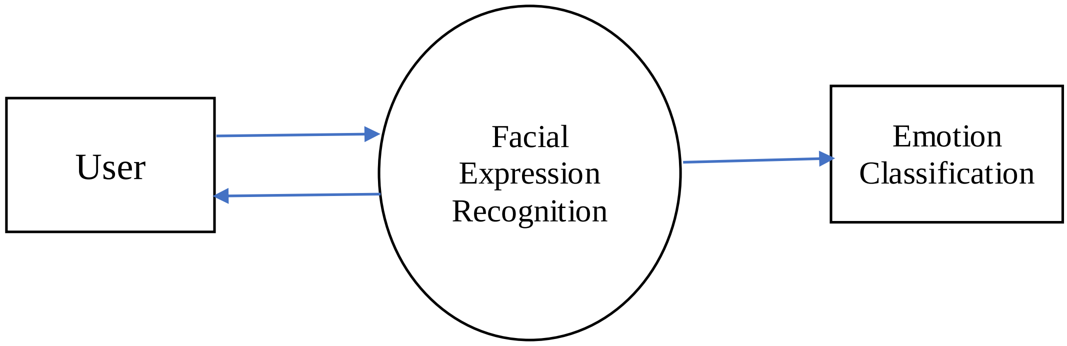


Fig 1

Level-1

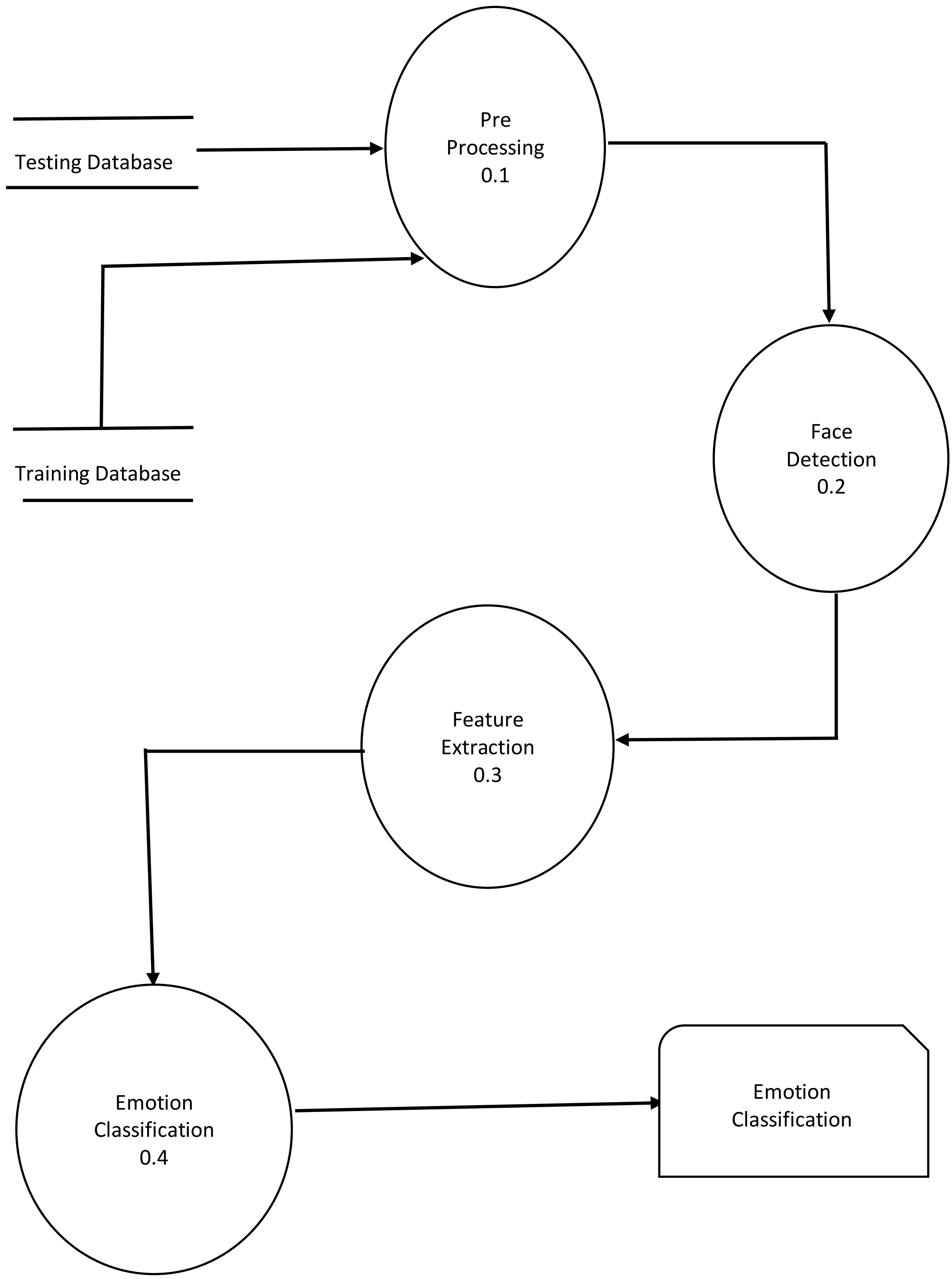

Testing Database
Training Database
Pre
Processing
0.1
Face
Detection
0.2
Feature
Extraction
0.3
Emotion
Classification
0.4
Emotion
Classification

Fig 2

UML Diagram

Emotion Classification

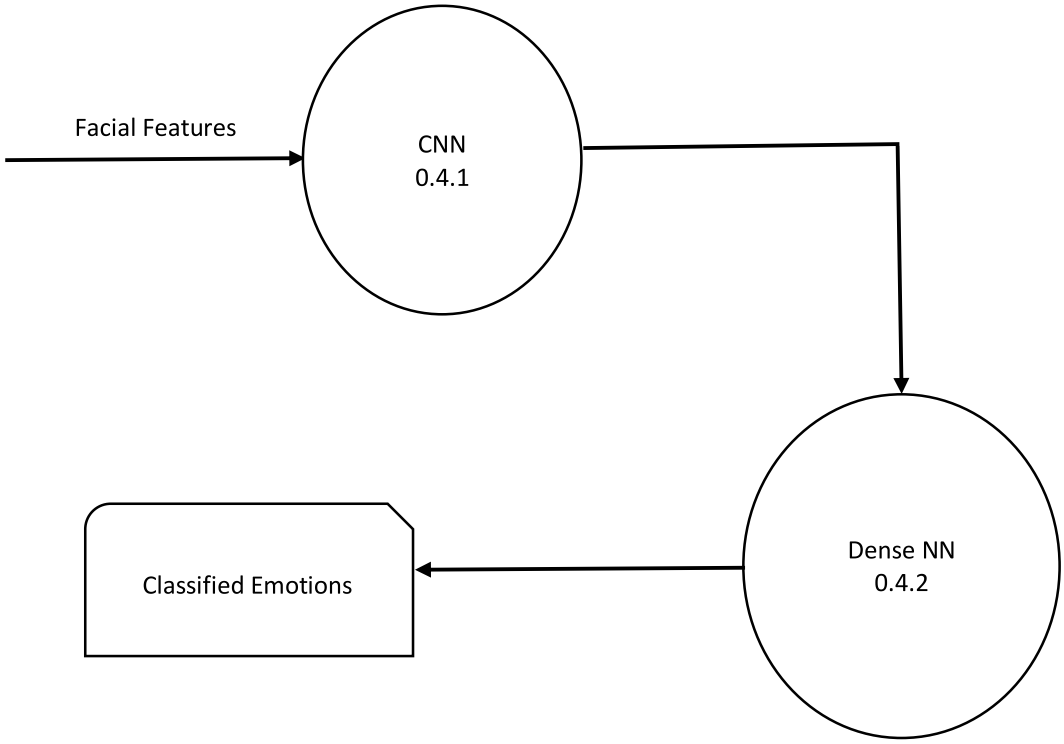


Fig 3

IMPLEMENTATION AND USER INTERFACE

Dataset

The dataset has been taken from an online source named Kaggle.com which was published on International Conference on Machine Learning (ICML) 5 years ago, to recognize the facial expression.

The data consists of 48x48 pixel grayscale images of faces. The faces have been automatically registered so that the face is more or less centered and occupies about the same amount of space in each image. This dataset consists of 35,887 images with various emotions -7 emotions, all labeled-.

Emotion labels in the dataset:

0: -4593 images- *Angry*

1: -547 images- *Disgust*

2: -5121 images- *Fear*

3: -8989 images- *Happy*

4: -6077 images- *Sad*

5: -4002 images- *Surprise*

6: -6198 images- *Neutral*

Fig 4

Data contains two columns, "emotion" and "pixels". The "emotion" column contains a numeric code ranging from 0 to 6, inclusive, for the emotion that is present in the image. The "pixels" column contains a string surrounded in quotes for each image. The contents of this string a space-separated pixel values in row major order.

CNN

- CNN is a type of neural network model which allows us to extract higher representations for the image content.

- Unlike the classical image recognition where you define the image features yourself, CNN takes the image's raw pixel data, trains the model, then extracts the features automatically for better classification.

- A convolution sweeps the window through images then calculates its input and filter dot product pixel values.

- CNN uses max pooling to replace output with a max summary to reduce data size and processing time. This allows you to determine features that produce the highest impact and reduces the risk of overfitting.

Algorithm

Step 1: Collection of a data set of images. (In this case we are using FER2013 database of 35887 pre-cropped, 48-by-48-pixel grayscale images of faces each labeled with one of the 7 emotion classes: anger, disgust, fear, happiness, sadness, surprise, and neutral.
Step 2: Pre-processing of images.
Step 3: Detection of a face from each image.
Step 4: The cropped face is converted into grayscale images.
Step 5: The pipeline ensures every image can be fed into the input layer as a (1, 48, 48) numpy array.
Step 6: The numpy array gets passed into the Convolution2D layer.
Step 7: Convolution generates feature maps.
Step 8: Pooling method called MaxPooling2D that uses (2, 2) windows across the feature map only keeping the maximum pixel value.
Step 9: During training, Neural network Forward propagation and Backward propagation performed on the pixel values.
Step 10: The Softmax function presents itself as a probability for each emotion class. The model is able to show the detail probability composition of the emotions in the face.

Method

The facial expression recognition system is implemented using convolutional neural network.
The block diagram of the system is shown in following figures:

Raw Image Normalization CNN Train CNN Weights

Training Phase

Raw Image Normalization CNN Facial Expression

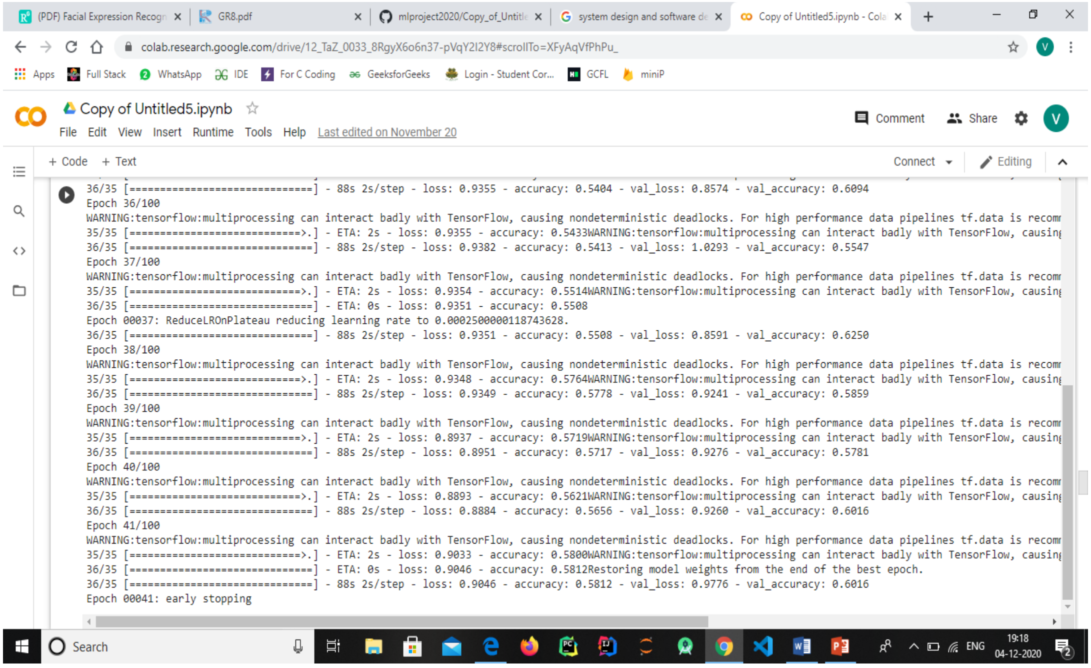
Testing Phase

During training, the system received a training data comprising grayscale images of faces with their respective expression label and learns a set of weights for the network. The training step took as input an image with a face. Thereafter, an intensity normalization is applied to the image. The normalized images are used to train the Convolutional Network. To ensure that the training performance is not affected by the order of presentation of the examples, validation dataset is used to choose the final best set of weights out of a set of trainings performed with samples presented in different orders. The output of the training step is a set of weights that achieve the best result with the training data. During test, the system received a grayscale image of a face from test dataset, and output the predicted expression by using the final network weights learned during training. Its output is a single number that represents one of the seven basic expressions.

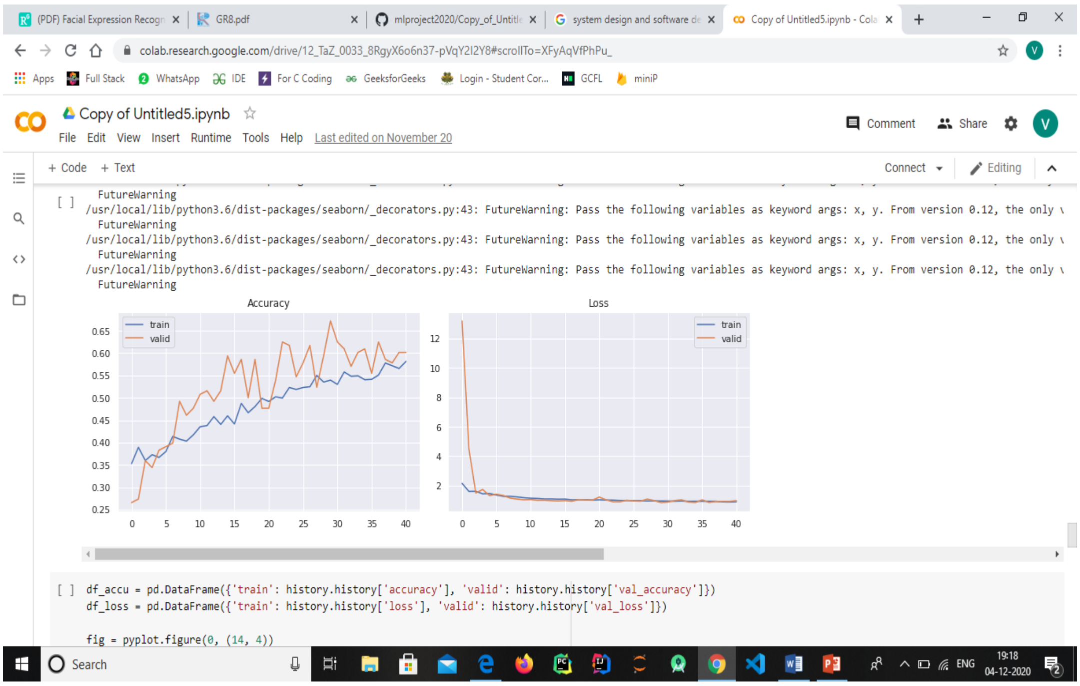
SOFTWARE TESTING

Outcome

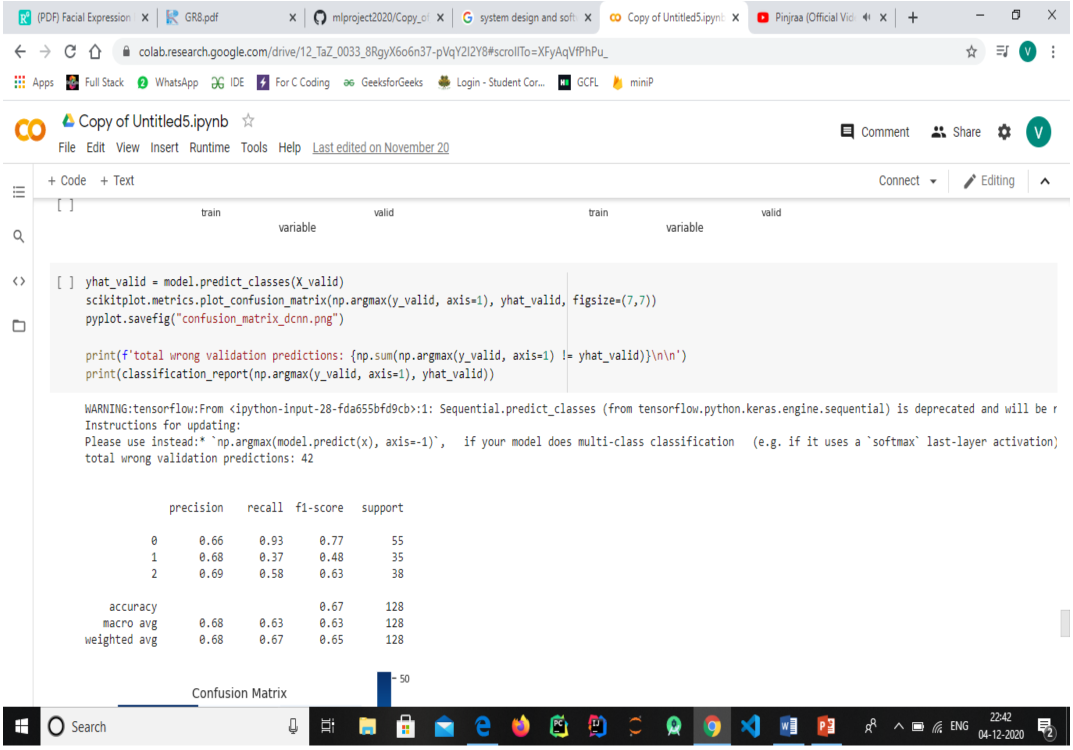
Fig 5


Copy of Untitled5.ipynb
File Edit View Insert Runtime Tools Help Last edited on November 20
36/35 [==============================] - 88s 2s/step - loss: 0.9355 - accuracy: 0.5404 - val_loss: 0.8574 - val_accuracy: 0.6094
Epoch 36/100
36/35 [==============================] - 88s 2s/step - loss: 0.9382 - accuracy: 0.5413 - val_loss: 1.0293 - val_accuracy: 0.5547
Epoch 37/100
Epoch 00037: ReduceLROnPlateau reducing learning rate to 0.0002500000118743628.
36/35 [==============================] - 88s 2s/step - loss: 0.9351 - accuracy: 0.5508 - val_loss: 0.8591 - val_accuracy: 0.6250
Epoch 38/100
36/35 [==============================] - 88s 2s/step - loss: 0.9349 - accuracy: 0.5778 - val_loss: 0.9241 - val_accuracy: 0.5859
Epoch 39/100
36/35 [==============================] - 88s 2s/step - loss: 0.8951 - accuracy: 0.5717 - val_loss: 0.9276 - val_accuracy: 0.5781
Epoch 40/100
36/35 [==============================] - 88s 2s/step - loss: 0.8884 - accuracy: 0.5656 - val_loss: 0.9260 - val_accuracy: 0.6016
Epoch 41/100
36/35 [==============================] - 88s 2s/step - loss: 0.9046 - accuracy: 0.5812 - val_loss: 0.9776 - val_accuracy: 0.6016
Epoch 00041: early stopping

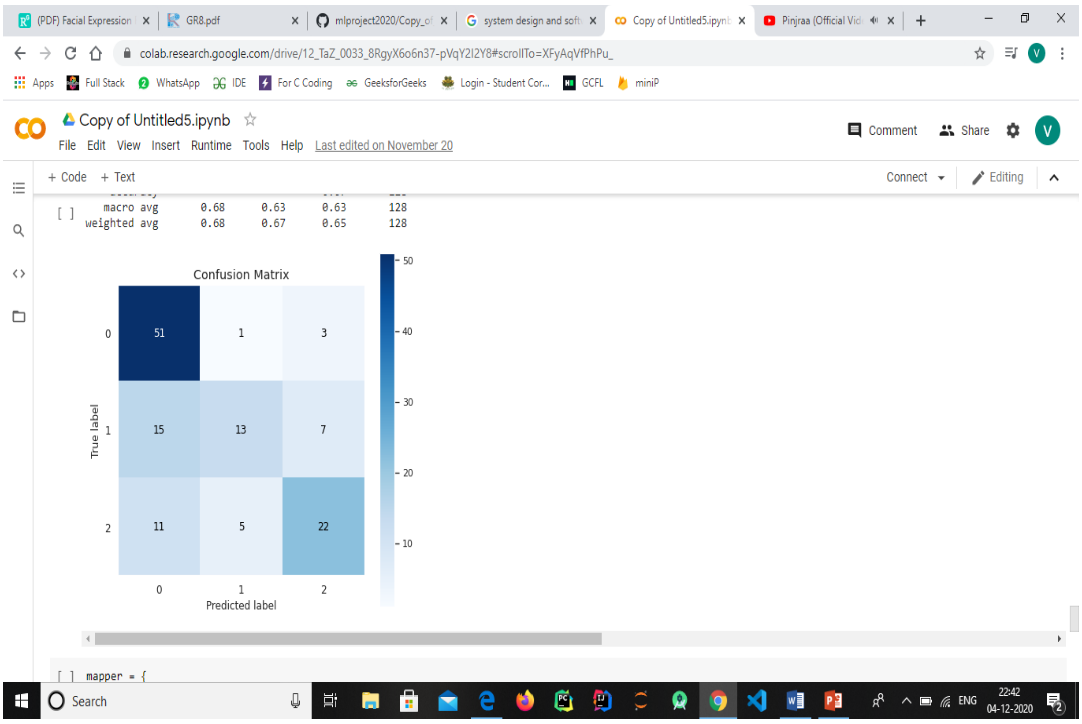
(1)


Copy of Untitled5.ipynb
Accuracy
Loss
train
valid
df_accu = pd.DataFrame({'train': history.history['accuracy'], 'valid': history.history['val_accuracy']})
df_loss = pd.DataFrame({'train': history.history['loss'], 'valid': history.history['val_loss']})
fig = pyplot.figure(0, (14, 4))

(2)


Copy of Untitled5.ipynb
File Edit View Insert Runtime Tools Help Last edited on November 20
yhat_valid = model.predict_classes(X_valid)
scikitplot.metrics.plot_confusion_matrix(np.argmax(y_valid, axis=1), yhat_valid, figsize=(7,7))
pyplot.savefig("confusion_matrix_dcnn.png")
print(f'total wrong validation predictions: {np.sum(np.argmax(y_valid, axis=1) != yhat_valid)}\n\n')
print(classification_report(np.argmax(y_valid, axis=1), yhat_valid))
WARNING:tensorflow:From <ipython-input-28-fda655bfd9cb>:1: Sequential.predict_classes (from tensorflow.python.keras.engine.sequential) is deprecated and will be r
Instructions for updating:
Please use instead:* `np.argmax(model.predict(x), axis=-1)`, if your model does multi-class classification (e.g. if it uses a `softmax` last-layer activation)
total wrong validation predictions: 42
precision recall f1-score support
0 0.66 0.93 0.77 55
1 0.68 0.37 0.48 35
2 0.69 0.58 0.63 38
accuracy 0.67 128
macro avg 0.68 0.63 0.63 128
weighted avg 0.68 0.67 0.65 128
Confusion Matrix

(3)


Copy of Untitled5.ipynb
macro avg 0.68 0.63 0.63 128
weighted avg 0.68 0.67 0.65 128
Confusion Matrix
51 1 3
15 13 7
11 5 22
True label
Predicted label
mapper = {

(4)

CONCLUSION

In this project, a LeNet architecture based six layer convolution neural network is implemented to classify human facial expressions i.e. happy, sad, surprise, fear, anger, disgust, and neutral.
The system has been evaluated using Accuracy, Precision, Recall and F1-score.
The classifier achieved:

- Accuracy of 58.12%
- Precision of 0.68
- Recall 0.67
- F1-score 0.65

In this case, when the model predicts incorrectly, the correct label is often the second most likely emotion. The facial expression recognition system presented in this research work contributes a resilient face recognition model based on the mapping of behavioral characteristics with the physiological biometric characteristics. The physiological characteristics of the human face with relevance to various expressions such as happiness, sadness, fear, anger, surprise and disgust are associated with geometrical structures which restored as base matching template for the recognition system.

FUTURE SCOPE

It is important to note that there is no specific formula to build a neural network that would guarantee to work well. Different problems would require different network architecture and a lot of trail and errors to produce desirable validation accuracy. This is the reason why neural nets are often perceived as "black box algorithms.".

In this project we got an accuracy of almost 70% which is not bad at all comparing all the previous models. But we need to improve in specific areas like-

- number and configuration of convolutional layers
- number and configuration of dense layers
- dropout percentage in dense layers

But due to lack of highly configured system we could not go deeper into dense neural network as the system gets very slow and we will try to improve in these areas in future.

We would also like to train more databases into the system to make the model more and more accurate but again resources becomes a hindrance in the path and we also need to improve in several areas in future to resolve the errors and improve the accuracy.

Having examined techniques to cope with expression variation, in future it may be investigated in more depth about the face classification problem and optimal fusion of color and depth information. Further study can be laid down in the direction of allele of gene matching to the geometric factors of the facial expressions. The genetic property evolution framework for facial expressional system can be studied to suit the requirement of different security models such as criminal detection, governmental confidential security breaches etc.

SUMMARY

A Facial expression is the visible manifestation of the affective state, cognitive activity, intention, personality and psychopathology of a person and plays a communicative role in interpersonal relations. Automatic recognition of facial expressions can be an important component of natural human-machine interfaces; it may also be used in behavioral science and in clinical practice. An automatic Facial Expression Recognition system needs to perform detection and location of faces in a cluttered scene, facial feature extraction, and facial expression classification.

Automatic recognition of facial expressions can be an important component of natural human-machine interfaces; it may also be used in behavioral science and in clinical practice. It have been studied for a long period of time and obtaining the progress recent decades. Though much progress has been made, recognizing facial expression with a high accuracy remains to be difficult due to the complexity and varieties of facial expressions.

Facial expression recognition system is implemented using Convolution Neural Network (CNN). CNN model of the project is based on LeNet Architecture. Kaggle facial expression dataset with seven facial expression labels as happy, sad, surprise, fear, anger, disgust, and neutral is used in this project. The system achieved 58.12 % accuracy and 0.68 precision on testing dataset.

APPENDICES

Appendix 1

Dataset Link:

https://www.kaggle.com/deadskull7/fer2013?select=fer2013.csv

Code Link of GitHub Repository:

https://github.com/VanshikaMahle/mlproject2020

Appendix 2

References

- Research Paper https://s.docworkspace.com/d/AEggGVv16qUt0ezIyuSdFA

- Convolutional Neural Networks (CNN) With TensorFlow by Sourav from Edureka[https://www.youtube.com/watch?v=umGJ30-15_A]

- https://www.kaggle.com/

- https://github.com/

- https://towardsdatascience.com/

- https://github.com/atulapra/Emotion-detection