SCTR's Pune Institute of Computer Technology
Dhankawadi, Pune

AN INTERNSHIP REPORT ON

Face feature analysis using Machine Learning

SUBMITTED BY

Name: Rajas Chitale
Class: TE-2
Roll no: 31218

Under the guidance of

Dr. Archana Ghotkar

DEPARTMENT OF COMPUTER ENGINEERING

ACADEMIC YEAR 2021-22

Internship Place Details

No.	Section	Details
1	Name of the internship offering industry/institution with address	Computational Linguistics Domain Lab, PICT Address: Survey No. 27, Near Trimurti Chowk, Bharati Vidyapeeth Campus, Dhankawadi, Pune, Maharashtra 411043
2	Industrial/institutional contact person's name, designation, and contact details	Dr. Archana Ghotkar, Incharge of CL lab Associate professor, aaghotkar@pict.edu 9422357775
3	Type of Internship (Onsite/online work from home)	Online
4	Internship work details (Scope and project assigned)	Title: Face feature analysis using Machine Learning Scope: Recognize face features for 6-7 emotions
5	Duration of internship	2 months (Starting 08/01/22 and excluding exam leaves)
6	Internship stipend	Unpaid internship
7	CGPA and attendance of previous semesters	CGPA: 9.435 Attendance: 93-100% in all semesters

Contents

List of Figures

1 Title

Face feature analysis using Machine Learning

2 Introduction

Computational Linguistics Lab in PICT is a lab where students carry out research and projects under the guidance of Dr. Archana Ghotkar, Dr. Sheetal Sonawane and other professors. My problem statement is a part of a big project, *"Multi-modal Preventive Approach and Detection of Stress Level Using Machine Learning Algorithm/s"*, of a PhD student. I am working on a Computer Vision module which can detect the facial expressions of the user to analyse the stress level on a self-made scale. The score of this scale will be then integrated with the scores of the other score generated from other modes of detection like Natural language processing, audio-processing, heart beat interval analysis, etc.

3 Problem Statement

To create a CNN model by using deep learning on facial expression datasets which can predict emotions (happy, sad, angry, neutral, surprise, disgust, fear) by taking real-time images as input.

4 Objectives

- To read about different emotions that can be helpful for analysing stress level. Read research papers regarding expression analysis and stress level analysis using facial expressions.
- To train and test a facial image dataset on some popular CNNs e.g. AlexNet, DeepFace Library
- To create or find a dataset having Indian facial images so that the model we train will be able to analyse the emotions on Indian faces more efficiently.
- To build a deep learning model to detect facial expressions and should learn to increase the accuracy of the model by analysing results and experimenting.

5 Motivation

Now a days, a lot of people suffer from mental illnesses.
Anxiety is the most common mental illness in the world, affecting 284 million people. (Our World in Data, 2018)
According to the WHO, 56 million Indians suffer from depression and another 38 million Indians suffer from anxiety disorders as of 2020. A lot of these people don't take treatment before their situation worsens and this can lead to deadly consequences. The main reason for people not opening up to anyone or any expert (Psychiatrists, Psychologists, etc.) is the feeling of shame and fear of other people getting to know about their disorder. For solving this problem using technology, we plan to develop an application which will take user's input in multiple modes (Facial image, audio, chatting text, etc.) to predict the stress/depression level of the user and inturn prescribe him/her a session with a psychologist, psychiatrist or some other mental health professional.

6 Scope

This project when complete can make a great self-test kit for mental health issues. The users will be able to take the test from their phone or computer and get the results without any human interference. A chat-bot integrated with video capturing (for stress level detection) can be the future where the patient talks or types his/her problems openly to get emotional support, consolidation from the artificial speaker and at the same time the user's mental state can be analysed and he/she can be suggested some helpful measure like visitiing a medical professional.

7 Methodological Details

The work started with the literature survey of the topic. As the project needed facial images of people showing different emotions, we had to search for such datasets. We came across quite a few datasets like AffectNet, FER-2013, ISED, Emotic, AVEC2014, etc. From these, we selected FER-2013 as our main dataset.
Facial Emotion Recognition -2013 (FER-2013):-

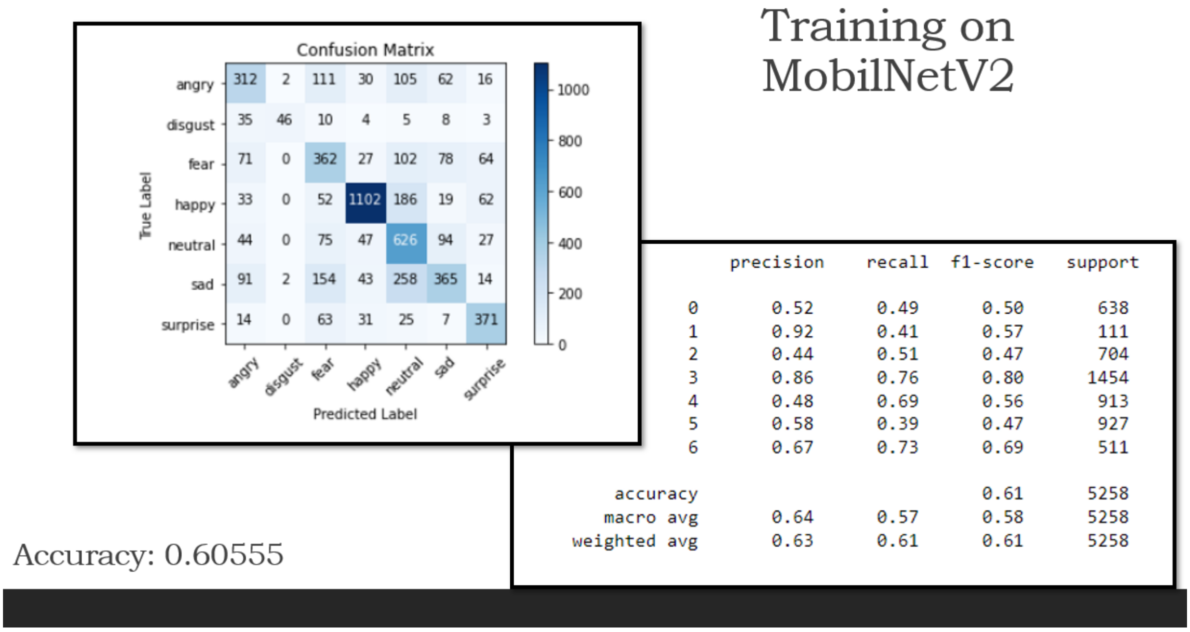
- FER-2013 has 35000 images
- The images cover 7 emotion labels - happy, sad, angry, neutral, surprise, disgust, fear
- This dataset is biased towards happy emotion hence some over-under sampling was performed on the dataset (making them 37936 images)

Figure 1: FER-2013 Sample images

In order to train the model over Indian facial images too, we merged the FER-2013 dataset with ISED. **ISED (India Spontaneous Expression Database):-**

- ISED has 428 images, approximately 500 videos of 51 different people
- ISED has 4 emotion labels - happy, sad, disgust, surprise
- We augmented the ISED dataset by flipping a few images horizontally and by over-sampling images under labels where there were less images

Train-test split use is 86%-14%.
Few known networks like AlexNet, MobilNetv2 were trained on the FER-2013 dataset to check what accuracy we are able to reach. The model which was built and trained by us was built using Convolutional Neural Networks.



	precision	recall	f1-score	support
0	0.52	0.49	0.50	638
1	0.92	0.41	0.57	111
2	0.44	0.51	0.47	704
3	0.86	0.76	0.80	1454
4	0.48	0.69	0.56	913
5	0.58	0.39	0.47	927
6	0.67	0.73	0.69	511
accuracy			0.61	5258
macro avg	0.64	0.57	0.58	5258
weighted avg	0.63	0.61	0.61	5258

Figure 2: MobilNetv2 Trained on FER-2013

```
model = Sequential()
model.add(Conv2D(32, kernel_size=(3,
3),activation='relu',input_shape=(224,224,1)))
model.add(Conv2D(64, (3, 3), activation='relu'))
model.add(MaxPooling2D(pool_size=(2, 2)))
model.add(Dropout(0.25))
model.add(Flatten())
model.add(Dense(128, activation='relu'))
model.add(Dropout(0.5))
model.add(Dense(7, activation='softmax'))
```

Figure 3: Experimental model 1

```
model = Sequential([
    Conv2D(32, kernel_size=(3, 3),
      activation='relu',input_shape=(224,224,1)),
    MaxPooling2D(pool_size=(2, 2)),

    Conv2D(64, (3, 3), activation='relu'),
    MaxPooling2D(pool_size=(2, 2)),
    Dropout(0.2),

    Conv2D(128, (3, 3), activation='relu'),
    Conv2D(128, (3, 3), activation='relu'),
    MaxPooling2D(pool_size=(2, 2)),
    BatchNormalizationV2(),

    Conv2D(256, (3, 3), activation='relu'),
    MaxPooling2D(pool_size=(2, 2)),
    Dropout (0.2),

    Conv2D (512, (3, 3), activation='relu'),
    MaxPooling2D (pool_size= (2, 2)),
    BatchNormalizationV2(),

    Dropout(0.25),
    Flatten(),
    Dense(256, activation='relu'),
    Dropout(0.5),
    Dense(7, activation='softmax')
])
```

Figure 4: Experimental model 2

```
model = Sequential([
    Conv2D(32, kernel_size=(3, 3),activation='relu',input_shape=(224,224,1)),
    MaxPooling2D(pool_size=(2, 2)),

    Conv2D(64, (3, 3), activation='relu'),
    Conv2D(64, (3, 3), activation='relu'),
    MaxPooling2D(pool_size=(2, 2)),
    Dropout(0.5),

    Conv2D(128, (3, 3), activation='relu'),
    Conv2D(128, (3, 3), activation='relu'),
    MaxPooling2D(pool_size=(2, 2)),
    BatchNormalizationV2(),

    Conv2D(256, (3, 3), activation='relu'),
    MaxPooling2D(pool_size=(2, 2)),
    Dropout(0.5),

    Conv2D(512, (3, 3), activation='relu'),
    MaxPooling2D(pool_size=(2, 2)),
    BatchNormalizationV2(),

    Dropout(0.5),
    Flatten(),
    Dense(256, activation='relu'),
    Dropout(0.5),
    Dense(7, activation='softmax')
])
```

Figure 5: Experimental model 3

P.T.O.

This project is ultimately to be deployed on a mobile application or a website in order to collect realtime data and give the depression analysis score to the user instantly. For the same reason, a first step in that direction, the model I built and trained was run using OpenCV to take in realtime data. The results are in the image below.

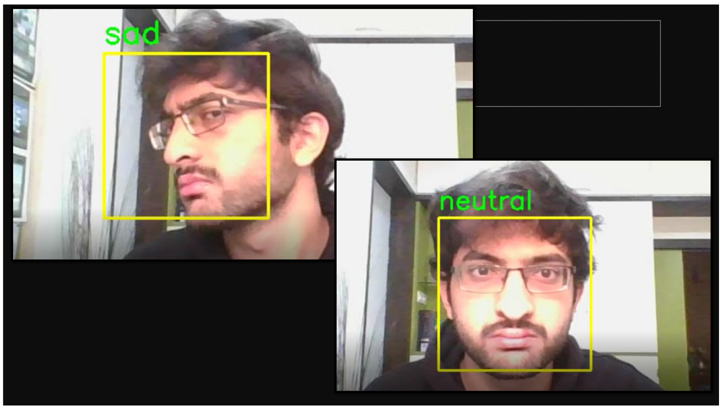


Figure 6: Real-time prediction results

8 Modern Engineering tools used

The modern engineering tools used in this project are:

1. **Tensorflow**
   TensorFlow is a free and open-source software library for machine learning and artificial intelligence. It can be used across a range of tasks but has a particular focus on training and inference of deep neural networks.

2. **Keras**
   Keras is an open-source software library that provides a Python interface for artificial neural networks. Keras acts as an interface for the TensorFlow library. Up until version 2.3, Keras supported multiple backends, including TensorFlow, Microsoft Cognitive Toolkit, Theano, and PlaidML.

3. **OpenCV** OpenCV is a library of programming functions mainly aimed at real-time computer vision. Originally developed by Intel, it was later supported by Willow Garage then Itseez. The library is cross-platform and free for use under the open-source Apache 2 License.

9 Result

The following table shows the accuracy scores when the dataset is trained on different CNNs or same same CNN but by changing the number of epochs.

Training results of my custom model

Sr. no.	Model	Epochs	Dataset	Accuracy
1	Custom Model /custModel1.h5	20	FER 2013 Oversampled	0.43952
2	Custom Model /custModel2.h5	20	FER13_ISED_OUS dataset	0.44237
3	Custom Model /custModel3.h5	20	FER13_ISED_OUS dataset	0.54545
4	Custom Model /custModel4.h5	30	FER13_ISED_OUS dataset	0.59147
5	Custom Model /custModel5.h5	40	FER13_ISED_OUS dataset	0.57874
6	Custom Model /custModel6.h5	35	FER13_ISED_OUS dataset	0.55534
7	Custom Model /custModel6.h5	35	FER13_ISED_OUS dataset	0.54089

Figure 7: Training results of custom CNN models

Models 1,2 refer to the CNN-1 [Figure 3]
Models 3,4,5,6 refer to the CNN-2 [Figure 4]
Models 7 refer to the CNN-3 [Figure 5]

We can see that the 4th model has the highest accuracy of 59.147%.

10 Analysis

The highest accuracy a Neural Network has been able to reach in the task of emotion recognition is 70%. This accuracy was achieved by WSCNet (Weakly Supervised Convolutional Network). As compared to this accuracy, what we have achieved(59% accuracy) is not bad but still there is a lot of scope for improvement.

11 Inference/Conclusion

As learned from the analysis section, we have successfully completed the problem statement but we need to optimize the CNN to get a better accuracy. This can be done by using more complex CNN or increasing the number of epochs.

References

[1] L. Yang, "Multi-Modal Depression Detection and Estimation", 2019 8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW), 2019, pp. 26-30, doi: 10.1109/ACIIW.2019.8925288.

[2] A. Jan, H. Meng, Y. F. B. A. Gaus and F. Zhang, "Artificial Intelligent System for Automatic Depression Level Analysis Through Visual and Vocal Expressions," in IEEE Transactions on Cognitive and Developmental Systems, vol. 10, no. 3, pp. 668-680, Sept. 2018, doi: 10.1109/TCDS.2017.2721552.

[3] K. Sato, H. Otsu, H. Madokoro and S. Kadowaki, "Analysis of psychological stress factors by using bayesian network",2013 IEEE International Conference on Mechatronics and Automation, 2013, pp. 811-818, doi: 10.1109/ICMA.2013.6618020.

[4] L. He, D. Jiang and H. Sahli, "Multimodal depression recognition with dynamic visual and audio cues," 2015 International Conference on Affective Computing and Intelligent Interaction (ACII), 2015, pp. 260-266, doi: 10.1109/ACII.2015.7344581.

[5] S. Alghowinem et al., "Multimodal Depression Detection: Fusion Analysis of Paralinguistic, Head Pose and Eye Gaze Behaviors," in IEEE Transactions on Affective Computing, vol. 9, no. 4, pp. 478-490, 1 Oct.-Dec. 2018, doi: 10.1109/TAFFC.2016.2634527.

[6] L. Zahara, P. Musa, E. Prasetyo Wibowo, I. Karim and S. Bahri Musa, "The Facial Emotion Recognition (FER-2013) Dataset for Prediction System of Micro-Expressions Face Using the Convolutional Neural Network (CNN) Algorithm based Raspberry Pi," 2020 Fifth International Conference on Informatics and Computing (ICIC), 2020, pp. 1-9, doi: 10.1109/ICIC50835.2020.9288560.

[7] S. L. Happy, P. Patnaik, A. Routray and R. Guha, "The Indian Spontaneous Expression Database for Emotion Recognition," in IEEE Transactions on Affective Computing, vol. 8, no. 1, pp. 131-142, 1 Jan.-March 2017, doi: 10.1109/TAFFC.2015.2498174.

12 Outcome of the internship in the form of deployment of Paper Presentation/Research Publication

After addition of some more features like realtime emotion detection, depression level scale and incremental dataset, this project will become worth presenting in a Conference.