

Automatic Mood and Gloom Detection using Visual Inputs

Aniket Gholve¹, Raja Kumar², Ritik Raj³, Om Sharma⁴

Students, Department of Computer Engineering^{1,2,3,4}
Sinhgad College of Engineering, Pune, Maharashtra, India

Abstract: *In natural psychological equilibrium, tension may be generally perceived as disturbance. If a user is unable to reconcile the expectations imposed on him/ her with user capacity to deal to them, so it generates tension and produces burden on mental health. Gloom may be generally described as psychological equilibrium disruption. One of major research fields of biomedical engineering is Gloom detection, as proper Gloom prevention could be easy. Facial expression recognition is the process of identifying human emotion. This is both something that humans do automatically but computational methodologies have also been developed. Several bio signals are available. Which are useful in identifying levels of Mood and Gloom since these signals indicate distinctive changes in the induction of Mood and Gloom. In this project, because of the easily accessible datasets on Kaggle, image processing is used as the primary candidate and the CNN model types have been formed which is used to predict the mood and gloom of persons.*

Keywords: CNN, Psychological Equilibrium, Image Processing, Kaggle

I. INTRODUCTION

Gloom and anxiety disorders are highly prevalent worldwide. Attention to the adverse effects of Gloom on patient health, as well as its associated economic burden has been warranted. To support objective Gloom assessment, the affective computing community engaged signal processing, computer vision and machine learning approaches for analyzing verbal and non-verbal behavior of Gloom patients and made predictions about what patterns should be indicative of Gloom state. These studies have analyzed the relationship between objective measures of voice, speech, non-verbal behavior and clinical subjective ratings of severity of Gloom for the purpose of automatic Gloom assessment. Despite major advances have been achieved in recent years, there are still several open research directions to be solved in the study of Gloom: Audio and video features from individual only concern the paralinguistic information, such as speaking rate, facial action units (AUs), etc. rather than the linguistic information from the speaking content, which can reflect the sleep status, emotional status, feeling and other life status of the individual. It is important to explore more effective audio, visual, linguistic and other multi-modal features, and design multi-modal fusion framework for Gloom recognition.

Due to the privacy issues, only limited Gloom datasets are currently available, and there are barely pre-training models for Gloom. Moreover, these common-used Gloom datasets also lack consistency. They have different languages, different durations, different data types and different targets, which make them difficult to be combined to increase the number of samples, therefore difficult to take advantage of deep models. Adopting some data augmentation approaches to increase the number of samples are requisite to improve the model performance.

Gloom is a state of low mood and aversion to activity. From this perspective, the study of Gloom should be closely related to affective state. However, the current researches on Gloom and affective state are relatively independent. We hypothesize that combining Gloom estimation and dimensional affective analysis simultaneously would yield more powerful Gloom analysis

II. LITERATURE SURVEY

Humans utilize emotion as a means of communicating their sentiments. It can be communicated by facial expressions, body language, and tone of speech. Because the most powerful, natural, and universal signal to express humans' emotion condition is their facial expression, it is a primary way of transmitting emotion. Human facial expressions, on the other hand, have similar patterns, making it difficult to recognize them with the naked eye. For example, the emotions of fear and



surprise are extremely similar. As a result, determining the face expression will be difficult. As a result, the goal of this research is to create a mobile-based emotion identification application that can determine emotion based on facial expression in real time. Convolutional Neural Network (CNN), a Deep Learning-based approach, is used in this research. The Mobile Net technique is used to train the recognition model. There are four different forms of face expressions: pleased, sad, surprised, and disgusted. As a result, this study's recognition accuracy was 85 percent. The built application could be improved in the future by including more face expression types.[1].

Machine learning has been implemented in the medical profession to improve diagnosis accuracy, precision, and analysis while minimizing tiresome tasks. Machine learning now has the ability to recognize mental illness such as depression, according to accumulating research. Because depression is the most common mental condition in our society today, and practically everyone suffers from it. As a result, depression detection models that provide a support system and early identification of depression are in high demand. This study is based on a machine learning-based image and video-based depression detection model. This study examines data collection strategies as well as databases. [2].

Facial expressions are important in social communication because they convey a lot of information about people, including their moods, feelings, and other characteristics. Many researchers achieved ideal accuracy in most widely used facial recognition datasets, while the best model accuracy in FER2013 is around 74%. The goal of this paper is to use deep learning-based models to solve this problem. [3].

Facial expression recognition is a hot topic in a variety of sectors, including artificial intelligence, gaming, marketing, and healthcare. The purpose of this paper is to sort human face photos into one of seven primary emotions. Before arriving at a final Convolutional Neural Network (CNN) model, a variety of models were tested, including decision trees and neural networks. Since of their huge number of filters, CNNs are excellent for image identification jobs because they can catch distinctive aspects of the inputs. Six convolutional layers, two max pooling layers, and two fully connected layers make up the suggested model.[4].

III. OBJECTIVES

- The main goal of this project is to Detect Moods and Gloom of Person
- The aim of the system is to continuously track human Mood and Gloom and try to contain it within normal levels.
- Using CNN Classifier to predict Mood and Gloom of user

IV. IMPLEMENTATION DETAILS OF MODULE

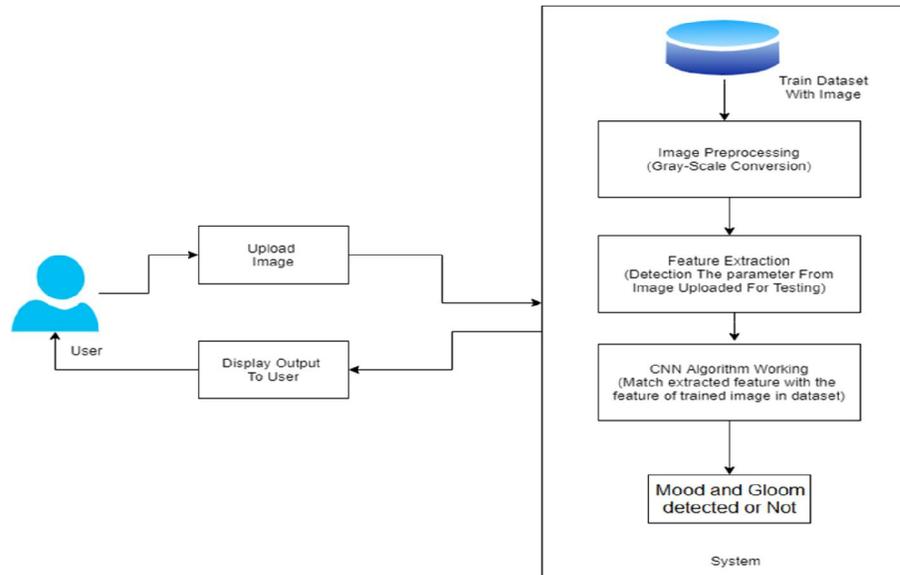


Figure: System Architecture



4.1 Convolutional Neural Network (CNN)

Step 1: Convolution Operation

Convolutional layer is the core part of the Convolutional neural network, which has local connections and weights of shared characteristics. The aim of Convolutional layer is to learn feature representations of the inputs. Convolutional layer consists of several feature maps. Each neuron of the same feature map is used to extract local characteristics of different positions in the former layer, but for single neurons, its extraction is local characteristics of same positions in former different feature map. In order to obtain a new feature, the input feature maps are first convolved with a learned kernel and then the results are passed into a nonlinear activation function. We will get different feature maps by applying different kernels.

Step 2: Pooling

The sampling process is equivalent to fuzzy filtering. The pooling layer has the effect of the secondary feature extraction, it can reduce the dimensions of the feature maps and increase the robustness of feature extraction. It is usually placed between two Convolutional layers. The size of feature maps in pooling layer is determined according to the moving step of kernels. The typical pooling operations are average pooling and max pooling. We can extract the high-level characteristics of inputs by stacking several Convolutional layer and pooling layer

Step 3: Full Connection

In general, the classifier of Convolutional neural network is one or more fully-connected layers. They take all neurons in the previous layer and connect them to every single neuron of current layer. There is no spatial information preserved in fully-connected layers. The last fully-connected layer is followed by an output layer. For classification tasks, SoftMax regression is commonly used because of it generating a well-performed probability distribution of the outputs. Another commonly used method is SVM, which can be combined with CNNs to solve different classification tasks

V. CONCLUSION

This study was undertaken for finding out the level of moods in five different videos of user. The presence of 'Happy', 'Neutral', 'Fearful', 'Sad' facial features, which are found prominent in videos were found out and analyzed. The dataset for training and testing was captured separately and the facial features of the same were classified using a CNN classifier. The amount of the positive and negative emotions in each image was analyzed and the image were predicted as image with 'Happy and Not Gloom', 'Sad and Mild Gloom' or 'High Symptoms of Gloom'.

REFERENCES

- [1]. Nur Alia Syahirah Badrulhisham and Nur Nabilah Abu Mangshor, "Emotion Recognition Using Convolutional Neural Network (CNN)". The 1st International Conference on Engineering and Technology (ICoEngTech), doi:10.1088/1742-6596/1962/1/012040, Journal of Physics: Conference Series 2021.
- [2]. D Y Liliana, "Emotion recognition from facial expression using deep convolutional neural network" 2018 International Conference of Computer and Informatics Engineering (IC2IE). doi:10.1088/1742-6596/1193/1/012004, IOP Conf. Series: Journal of Physics: Conf. Series 1193 (2019) 012004 I
- [3]. Rahul Chauhan, Kamal Kumar Ghanshala and R.C Joshi, "Convolutional Neural Network (CNN) for Image Detection and Recognition" First International Conference on Secure Cyber Computing and Communication (ICSCCC) 2018
- [4]. Akash Saravanan, Gurudutt Perichetla, Dr. K.S.Gayathri, "Facial Emotion Recognition using Convolutional Neural Networks" doi : arXiv:1910.05602v1 [cs.CV] 12 Oct 2019
- [5]. Lang Hea, Cui Caob, "Automated depression analysis using convolutional neural networks from speech," Journal of Biomedical Informatics Volume 83, Pages 103-111, July 2018.
- [6]. Ninad Mehendale "Facial emotion recognition using convolutional neural networks (FERC)". Received: 16 July 2019 / Accepted: 12 February 2020 / Published online: 18February 2020 © Springer Nature Switzerland AG 2020.
- [7]. Erhan Tiriyaki Akshay Sonawane, Lakshman Tamil. "Real-Time CNN Based ST Depression Episode Detection Using Single-Lead ECG". 021 22nd International Symposium on Quality Electronic Design (ISQED) | 978-1-7281-7641-3/20/\$31.00 ©2021 IEEE | DOI: 10.1109/ISQED51717.2021.9424275
- [8]. M. Pushpa Valli, "Image Processing Technique for Fruit Grading", International Journal of Engineering and Advanced Technology (IJEAT) 2019.



- [9]. Milan Tripathi. "Facial Emotion Recognition Using Convolutional Neural Network" Issn: 0976-9102 (Online) Ictact Journal On Image And Video Processing, Volume: 12, ISSUE: 01 DOI: 10.21917/ijivp.2021.0359, AUGUST 2021.
- [10]. Arselan Ashraf, Teddy Surya Gunawan, Bob Subhan Riza, Edy Victor Haryanto, Zuriati Janin "On the review of image and video-based depression detection using machine learning." Indonesian Journal of Electrical Engineering and Computer Science Vol. 19, No. 3, pp. 1677~1684 ISSN: 2502-4752, DOI: 10.11591/ijeecs. v19.i3. pp1677-1684, September 2020.