



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 12, Issue 5, May 2023

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.423

Depression Detection Using Text and Facial Expression

Dr. K. Mariyappan, Pratyush Pandey, Shanu Babu, Abhinav Shrivastava, M.Manisha

Associate Professor, Department of Information Science and Engineering, Faculty of Engineering and Technology, Jain University, Bangalore, India

Department of Information Science and Engineering, Faculty of Engineering and Technology, Jain University, Bangalore, India

ABSTRACT: Depression is a common illness worldwide. According to WHO, with an estimated 3.8% of the population affected, including 5.0% among adults and 5.7% among adults older than 60 years . Approximately 280 million people in the world have depression . Depression is different from usual mood fluctuations and short-lived emotional responses to challenges in everyday life. Especially when recurrent and with moderate or severe intensity, depression may become a serious health condition. It can cause the affected person to suffer greatly and function poorly at work, at school and in the family. At its worst, depression can lead to suicide. Over 700 000 people die due to suicide every year. Suicide is the fourth leading cause of death in 15-29-year-olds. Although there are known, effective treatments for mental disorders, more than 75% of people in low- and middle-income countries receive no treatment. Barriers to effective care include a lack of resources, lack of trained health-care providers and social stigma associated with mental disorders. In countries of all income levels, people who experience depression are often not correctly diagnosed, and others who do not have the disorder are too often misdiagnosed and prescribed antidepressants. Depression is a major mental health disorder that not only affects the emotional but also the physical and psychological state of the person. To face this problem we are planning to build a system which will help us to predict whether a person has any symptoms of depression or not as predicting the actual depression state of a person is very difficult. So, this model will study the basic moods of a person like happiness, sadness, fear, and anger using facial recognition and ask some basic questions about there mood, appetite, interest in doing things etc.and try to analyse the sentiment of there answers which will eventually help us to get more accurate answer. For facial recognition we will use CNN(convolutional neural network) and for answer we will use roBERTa model.

I.INTRODUCTION

Depression (major depressive disorder) is a common and serious medical illness that negatively affects how you feel, the way you think and how you act. Fortunately, it is also treatable. Depression causes feelings of sadness and/or a loss of interest in activities you once enjoyed. It can lead to a variety of emotional and physical problems and can decrease your ability to function at work and at home.

Depression symptoms can vary from mild to severe and can include:

- Feeling sad or having a depressed mood
- Loss of interest or pleasure in activities once enjoyed
- Changes in appetite — weight loss or gain unrelated to dieting
- Trouble sleeping or sleeping too much
- Loss of energy or increased fatigue
- Increase in purposeless physical activity (e.g., inability to sit still, pacing, handwringing) or slowed movements or speech (these actions must be severe enough to be observable by others)
- Feeling worthless or guilty
- Difficulty thinking, concentrating or making decisions
- Thoughts of death or suicide

Types of depression

Healthcare providers name depression types according to symptoms and causes. These episodes often have no obvious cause. In some people, they can linger much longer than in others for no clear reason. Types of depression include:Depression



- **Major depressive disorder (MDD):** Major depression (clinical depression) has intense or overwhelming symptoms that last longer than two weeks. These symptoms interfere with everyday life.
- **Bipolar depression:** People with bipolar disorder have alternating periods of low mood and extremely high-energy (manic) periods. During the low period, they may have depression symptoms such as feeling sad or hopeless or lacking energy.
- **Perinatal and postpartum depression:** “Perinatal” means around birth. Many people refer to this type as postpartum depression. Perinatal depression can occur during pregnancy and up to one year after having a baby. Symptoms go beyond “the baby blues,” which causes minor sadness, worry or stress.
- **Persistent depressive disorder (PDD):** PDD is also known as dysthymia. Symptoms of PDD are less severe than major depression. But people experience PDD symptoms for two years or longer.
- **Premenstrual dysphoric disorder (PMDD):** Premenstrual dysphoric disorder is a severe form of premenstrual disorder (PMS). It affects women in the days or weeks leading up to their menstrual period.
- **Psychotic depression:** People with psychotic depression have severe depressive symptoms and delusions or hallucinations. Delusions are beliefs in things that are not based in reality, while hallucinations involve seeing, hearing, or feeling touched by things that aren’t actually there.
- **Seasonal affective disorder (SAD):** Seasonal depression, or seasonal affective disorder, usually starts in late fall and early winter. It often goes away during the spring and summer. To detect depression we will make a fusion model which will include examining the facial expression of the person sitting in front of camera. To examine the facial expression we will study the basic moods of a person like sadness, happiness, fear, anger, surprise and disgust using CNN (convolutional neural network). After studying the facial expression, we will ask some basic questions to the person “What doctor use to ask while counselling about depression” like question about their appetite or interest in doing things etc. Then we will perform the sentimental analysis on those answers using roBerta model(Robustly Optimized BERT Pre-training Approach) and check the negativity of those answers. Based on these two factors we will try to predict, what are the percentage of a person that he/she can be depressed as detecting the actual mental state is not easy

CNN(convolutional neural network)

A **Convolutional Neural Network(ConvNet/CNN)** is a Deep Learning algorithm that can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image, and be able to differentiate one from the other. The pre-processing required in a ConvNet is much lower as compared to other classification algorithms. While in primitive methods filters are hand-engineered, with enough training, ConvNets have the ability to learn these filters/characteristics.

RoBERTa model(Robustly Optimized BERT Pre-training Approach)

RoBERTa stands for Robustly Optimized BERT Pre-training Approach. It was presented by researchers at Facebook and Washington University. The goal of this paper was to optimize the training of BERT architecture in order to take lesser time during pre-training. RoBERTa has almost similar architecture as compare to BERT, but in order to improve the results on BERT architecture, the authors made some simple design changes in its architecture and training procedure.

II. RELATED WORK

Automatic recognition of facial gestures (i.e., facial muscle activity) is rapidly becoming an area of intense interest in the research field of machine vision. In this paper, we present an automated system that we developed to recognize facial gestures in static, frontal- and/or profile-view color face images. A multidetector approach to facial feature localization is utilized to spatially sample the profile contour and the contours of the facial components such as the eyes and the mouth. From the extracted contours of the facial features, we extract ten profile-contour fiducial points and 19 fiducial points of the contours of the facial components. Based on these, 32 individual facial muscle actions (AUs) occurring alone or in combination are recognized using rule-based reasoning. With each scored AU, the utilized



algorithm associates a factor denoting the certainty with which the pertinent AU has been scored. A recognition rate of 86% is achieved.

The fact that depression tendency detection is not equivalent to sentiment analysis makes it difficult to find a unified standard for depression tendency detection and to accurately mine depression emotions in image and text. A depression tendency detection model based on fusion of image and text is proposed, which is mainly composed of a textual depression tendency detection model and an image depression tendency detection model. These two models extract text emotion feature through a pre-trained BERT model and use pre-trained VGG19 to learn image emotion feature. Then use a BiGRU-based classifier to obtain the corresponding depression tendency polarity probability. Finally, a model fusion formula is designed in order to play the complementary role of text and image, according to the idea of late fusion. The model combines the text depression tendency detection model and the image depression tendency detection model to perform comprehensive depression tendency detection. The experimental results on the simplified WU3D dataset show that the proposed model has better performance in each index evaluation compared with the single-modal depression tendency detection model.

Face emotion is an utmost efficient way of nonverbal communication and indication of sentimental status, mentality and intent. In this thesis, we focused on computer vision-based schemes to identify and recognize the faces and expressions, respectively. In this work a CNN based prototype is presented to estimate the age and to classify the gender of detected faces along with its emotions. The proposed model is implemented in three stages namely a hybrid feature extraction technique to identify the facial expressions by using SVM classifier. A combination of SIFT and Deep learning-based model is presented for identifying facial emotions. Whale optimization algorithm is applied to optimize the performance of face emotion recognition. This model gives a convolution neural network and introduces a novel technique for classification of facial emotions, age prediction and male-female identification. The data set for the analysis is considered from WideResNet architecture and Caffe models. The proposed model is tested on a static emotion identification benchmark dataset, and empirical investigation demonstrates that it improves emotion recognition significantly. An accuracy of 98.79% is reported and compared with the existing methods. The reported model is found to be suitable for many applications which involves human computer interaction with the adoption of emerging technologies AI and ML.

Recognizing human emotions from partial facial features is quite hard to achieve reasonable accuracy. In this paper, we propose to use a tree structure representation to simulate as human perceiving the real human face and both the entities and relationship could contribute to the facial expression features. Moreover, a new structural connectionist architecture based on a probabilistic approach to adaptive processing of data structures is presented to generalize the Face Emotion Tree Structures (FEETS). We demonstrated the robustness of our proposed system in recognizing the correct emotion based on partial face features. The system yields an accuracy of about 90% for subjects with partial face covered by artifacts. FACE reveals the emotion states as well as cognitive activities, faster than people verbalize or even realize their emotional state. Psychologists have also found that the cognitive interpretations of emotions from facial expressions were innate and universal to all humans regardless of cultures. Paul Ekman has identified six basic categories of emotions (i.e. fear, anger, sadness, surprise, disgust and happiness). He has created the Facial Action Coding System (FACS) to recognize the emotional states from the measurement of facial movements. It is a leading method used in behavioral science, but for a human expert to identify the 46 defined Action Units a minimal 100 hours of training is needed. Each of the defined Action Units corresponds into each independent motion of the face.

III. METHODOLOGY

The is the architecture diagram of our system which we are planning to build. This is a fusion of two different techniques. To detect depression we will first make a system which will examine the facial expression of the person sitting in front of camera. To examine the facial expression we will study the basic moods of a person like sadness, happiness, fear, anger, surprise and disgust using CNN (convolutional neural network).



CNN consist of 3 main type of layers:

Convolutional layer

In this layer the input image will be divided into the matrix of millions of block. To extract the features of the we have to go through the each and every pixel of the image. But reading these many pixel is not an easy task. So in convolutional layer will apply some filters which are known as kernals. These kernals will be applied to a set of pixel and extract only important features by ignoring some of the unwanted pixels. So, we can say that the convolution is searching the image by sliding a filter (kernel) across the image to find different features of the image. Kernels are just 2D matrices with different weights in them. Basically, this kernel will pass over the image replacing the pixel values with the average of the sum of its weight on the respective part of the image. These kernels are an amazing way to find the most important features in the image.

Pooling layer

Now that you have found the important features of the image, still, the amount of input is very large and our machine could not be able to handle this amount of inputs. So here is where pooling comes. Pooling is just reducing the size of the image without losing the features that we found with convolution. For example, a MaxPooling method will take in a shape of a matrix and return the larger value in that range. Bydoing this we can compress the image without losing the important features of this image.

Flattening

Flattening is nothing but converting a 3D or 2D matrix into a 1D input for the model this will be our last step to process the image and connect the inputs to a fully connected dense layer for further classification. Based up on this one dimensional result we can say that what might be the expression of the person. After studying the facial expression, we will ask some basic questions to the person “What doctor use to ask while counselling about depression” like question about their appetite or interest in doing things etc. Then we will perform the sentimental analysis on those answers using roBerta model(Robustly Optimized BERT Pre-training Approach) and check the negativity of those answers. Roberta model is pre-trained model which was developed by Facebook. This model is based on the Bert masking language technique. First it will perform the text embedding. In this it will encode the sentence by adding some embeddings like:

Token embeddings: A [CLS] token is added to the input word tokens at the beginning of the first sentence and a [SEP] token is inserted at the end of each sentence.

Segment embeddings: A marker indicating Sentence A or Sentence B is added to each token. This allows the encoder to distinguish between sentences.

Positional embeddings: A positional embedding is added to each token to indicate its position in the sentence. Depression

5.1 embedding in Bert masking

Then it will perform pre-training. In this it will perform 2 natural language processing tasks:

Masked Language Modeling

BERT is designed as a **deeply bidirectional** model. The network effectively captures information from both the right and left context of a token from the first layer itself and all the way through to the last layer. Let’s say we have a sentence – “I love to data science ”. We want to train a bi directional language model. Instead of trying to predict the next word in the sequence, we can build a model to predict a missing word from within the sequence itself. Let’s replace “data” with “[MASK]”. This is a token to denote that the token is missing. We’ll then train the model in such a way that it should be able to predict “data” as the missing token: “I love to read [MASK] science ”. The masked words were not always replaced by the masked tokens [MASK] because the [MASK] token would never appear during fine-tuning. So, the researchers used the below technique:

- 80% of the time the words were replaced with the masked token [MASK]
- 10% of the time the words were replaced with random words
- 10% of the time the words were left unchanged



Next Sentence Prediction

Masked Language Models (MLMs) learn to understand the relationship between words. Additionally, BERT is also trained on the task of Next Sentence Prediction for tasks that require an understanding of the relationship between sentences.

A good example of such a task would be question answering systems. Since it is a binary classification task, the data can be easily generated from any corpus by splitting it into sentence pairs. Just like MLMs, the authors have added some caveats

here too. Let's take this with an example:

Consider that we have a text dataset of 100,000 sentences. So, there will be 50,000 training examples or pairs of sentences as the training data.

- For 50% of the pairs, the second sentence would actually be the next sentence to the first sentence
- For the remaining 50% of the pairs, the second sentence would be a random sentence from the corpus
- The labels for the first case would be 'IsNext' and 'NotNext' for the second case

Using these factors it will give the polarity of the sentences and using that polarity we can calculate the percentage of negativity.

With the help of negativity percentage of sentences and facial expressions we will calculate the estimate by taking the average of both. Thus, based on the average we will try to predict whether the respected person is depressed or not.

IV. PROPOSED SYSTEM

In our proposed system, to detect depression we will make a fusion model which will include examining the facial expression of the person sitting in front of camera. To examine the facial expression we will study the basic moods of a person like sadness, happiness, fear, anger, surprise and disgust using CNN (convolutional neural network). After studying the facial expression, we will ask some basic questions to the person "What doctor use to ask while counselling about depression" like question about their appetite or interest in doing things etc. Then we will perform the sentimental analysis on those answers using roBerta model(Robustly Optimized BERT Pre-training Approach) and check the negativity of those answers. Based on these two factors we will try to predict, what are the percentage of a person that he/she can be depressed as detecting the actual mental state is not easy.

V. RESULT

We have developed a system for detection of early stage depression using "Deepface" for detection for emotions which has an accuracy of 89.9% and "roBerta model" which has an accuracy of 92.5%. By combining these to result our fusion model will give us the accuracy of 91.2%.

VI. CONCLUSION

In this paper, we have developed a system for detection of early stage depression by detecting the emotion of the person using a convolutional neural network called "Deepface" and by checking the negativity of the answers using roBerta model(Robustly Optimized BERT Pre-training Approach). In future we can develop a complete user friendly application which can be use by anyone at their homes. Our system should

REFERENCES

1. Published in: 2019 Eleventh International Conference on Ubiquitous and Future Networks (ICUFN),Date of Conference: 02-05 July 2019,Date Added to IEEE *Xplore*: 22 August 2019,INSPEC Accession Number: 18940575 ,DOI: 10.1109/ICUFN.2019.8806182
2. Published in: 2020 IEEE International Conference on Big Data (Big Data),Date of Conference: 10-13 December 2020,Date Added to IEEE *Xplore*: 07 April 2021,INSPEC Accession Number: 20533475,DOI:10.1109/BigData50022.2020.9391769
3. Published in: 2021 6th International Conference on Inventive Computation Technologies (ICICT), Date of Conference: 20-22 January 2021,Date Added to IEEE *Xplore*: 26 February 2021,INSPEC Accession Number: 20431038,DOI:10.1109/ICICT50816.2021.9358772



4. Published in: 2020 IEEE 9th Joint International Information Technology and Artificial Intelligence Conference (ITAIC), Date of Conference: 11-13 December 2020, Date Added to IEEE *Xplore*: 03 February 2021, INSPEC Accession Number: 20405641, DOI:10.1109/ITAIC49862.2020.9339058, Publisher: IEEE
5. Published in: 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), Date of Conference: 08-10 April 2021, Date Added to IEEE *Xplore*: 06 May 2021, INSPEC Accession Number: 20676714, DOI:10.1109/ICCMC51019.2021.9418320, Publisher: IEEE
6. Published in: 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM) Date of Conference: 27-30 August 2019 Date Added to IEEE *Xplore*: 23 April 2020 ISBN Information: Electronic ISBN:978-1-4503-6868-1 Print on Demand (PoD) ISBN:978-1-4503-6868-1 ISSN Information: INSPEC Accession Number: 19557027 DOI: 10.1145/3341161.3343854 Publisher: IEEE Conference Location: Vancouver, BC, Canada
7. <https://www.who.int/news-room/fact-sheets/detail/depression>
8. <https://online.uwa.edu/infographics/basicemotions/#:~:text=The%20Six%20Basic%20Emotions&text=They%20include%20sadness%2C%20happiness%2C%20fear,%2C%20anger%2C%20surprise%20and%20disgust.>
9. <https://my.clevelandclinic.org/health/diseases/9290-depression>
10. <https://towardsdatascience.com/acomprehensive-guide-to-convolutional-neuralnetworks-the-eli5-way-3bd2b1164a53>



INNO SPACE
SJIF Scientific Journal Impact Factor
Impact Factor: 8.423



ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com



www.ijirset.com

Scan to save the contact details