International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 05 | May 2020 www.irjet.net p-ISSN: 2395-0072

ARTIFICIAL INTELLIGENCE BASED DEPRESSION RECOGNITION SYSTEM

Lija Joy[1], Nissy Alex[2], Sherin Thomas[3], Teenu Sunny[4], Angitha George[5]

[1,2,3,4]UG Students, Dept. of Computer Science and Engineering, St. Joseph's College of Engineering and Technology, Palai, Kerala, India

[5]Assistant Professor, Dept. of Computer Science and Engineering, St. Joseph's College of Engineering and Technology, Palai, Kerala, India

---

Abstract - *Prediction of depression manually is a time consuming process. Apps and chat boxes with cognitive capabilities can recognize the mental state of a person. This project proposes a method to predict the level of depression using deep learning and machine learning algorithms. Audio and video input is collected for this purpose. Both these inputs are analyzed separately to get accurate results. From video, emotion is detected and using that result together with the audio input, we analyze the level of depression. The output is the predicted value of the depression.*

Key Words: **Haar Cascade Classifier, Naïve- Bayes, CNN, DNN, Machine learning**

1. INTRODUCTION

Depression. That word we commonly hear nowadays. It is a health issue that persists among a large number of people. According to the World Health Organization (WHO) 2018 report, depression affects more than three hundred million people worldwide. Depression is characterized by a number of symptoms. These may include persistent sadness, anxiety or a feeling of emptiness. Symptoms may vary from person to person. It is classified as a mood disorder that interfere with a person's everyday activities. Depression can be sometimes associated with a deficit of cognitive control. Depression, in psychology is defined as "a mood or emotional state that is marked by feelings of low self-worth or guilt and a reduced ability to enjoy life".

American Psychiatric Association states that depression causes feelings of sadness and/or loss of interest in the activities once enjoyed [3]. It often leads to various emotional and physical problems. It decreases a person's ability to function at work or home. Studies show that women and elderly people are more commonly affected than men. In a recent report it is found that women at the age between 40 to 59 are more likely to fall into depression. Various researches are still going on in order to find the root cause of depression. Scientists think that it may be due to the imbalance of brain's signaling chemicals called neurotransmitters such as serotonin, dopamine etc., Neuro transmitters carry signals in the brain, which the body uses to control the mood. Artificial intelligence and machine learning techniques have been used in recent time to progressively predict the mental state. Facial expression and emotion detection is an effective method to partition the depression level into different levels of depression. This method aims to tackle with the problem of early. Less time consuming and efficient recognition of depression will reduce the risk of human's life. These techniques efficiently help the psychologist to analyse the depression level and is less time consuming. Here to take visual and audio input we use the system's webcam and a microphone of any type .The obtained input data can then be used to extract the audio and visual features and then be fed to the trained model. The aim of the paper is to describe a method to automatically predict the type of the depression using the visual and audio data of the patient.

2. RELATED WORKS

2.1 Artificial Intelligent Systems for Automatic Depression Level Analysis through Visual and Vocal Expressions

In this proposed method, depression is monitored using an artificial intelligence system. A session of questionnaires is provided and a webcam is used to record the responses of a person in order to analyse whether he/she is suffering from depression. Beck Depression Inventory ll (BDl-ll) is the scale used here to plot depression from vocal and visual expressions [7].First, Deep learning method is used to extract key visual features from facial expressions. Second, in order to capture vocal expressions spectral low-level descriptors and mel-frequency cepstral coefficient features are extracted from short audio segments. Third,the temporal movement in feature space is captured using feature dynamic history histogram (FDHH).Finally, regression techniques are used for fusing these FDHH and audio features for the prediction of the BDl-ll scales. Thus the level of depression is predicted. The results indicates that the proposed method works well on the AVEC2014 dataset. The AVEC2014 dataset contains 300 video clips [5].These clips are of length 6s to 4min.People between the age group 18 to 63 years are considered here. These recordings are then used to map on BDl-ll depression scale. The range of the scale is[0,63],where 0-10 is classified as normal,11-16 is mild mood disturbance,17-20 is a borderline clinical depression,21-30 is moderate,31-40 is severe and above 40 is extreme depression. This paper provides a general framework that can be used for automatic detection of depression level using facial and vocal expressions. The proposed system shows a remarkable performance an

application that has slow changing facial expressions by focusing on the small changes of pattern within the deep/handcrafted descriptors [4].

2.2 Depression Detection Based on Deep Distribution Learning

In this paper, a deep learning architecture is used for predicting depression levels accurately. It relies on a new expectation loss function that allows to estimate the underlying data distribution over depression levels, here the expected values of the distribution are optimized to approach the ground-truth levels. This approach can predict depression levels more accurately. Expectation loss is proposed to train the model in order to estimate a data distribution over depression levels where the expected value of the estimated depression distribution gives the predicted value. This distribution allows to examine the ordinal relationships between facial images and depression levels, and thus improves the accuracy of the model without additional streams [1].

2.3 Depression Detection using Emotion Artificial Intelligence

This paper aims to apply natural language processing on twitter feeds for emotion analysis. Here individual tweets are analyzed and is classified as neutral or negative based on the words used there. This is done to identify depression tendencies. The paper focus on text based emotion detection. Support vector machine and Naive -Bayes classifier is used for this purpose. It uses the twitter API for the generation of the dataset. After the dataset creation, there comes data pre-processing module which systematically churns the data. The data pre-processing includes tokenization, stemming, stop words removal, POS tagger. The POS tagger, then identifies the pieces of text which should be analyzed thoroughly. Then in the training phase the text classifier is trained on the processed text data. In the testing phase, class is predicated on the test dataset to identify potential tweets demonstrating depression tendencies. Thus the tweets supporting depression could be identified [6].

2.4 Detecting Depression severity from vocal prosody

In this paper, Hamilton Rating Scale for Depression (HRSD) is used to analyse participants undergoing a clinical trial for treatment of depression[8] .This aims to investigate the relation between vocal prosody and change in depression severity. HRSD is the most widely used clinician administered depression assessment scale which determines a patient's level of depression before, during, and after treatment. Making use of both perceptual judgments by naive listeners and quantitative analysis of vocal timing and fundamental frequency three hypotheses were tested. 1) First, Naive listeners can perceive the severity of depression from vocal recordings of depressed participants and interviewers. 2) It is found that the Quantitative features of vocal prosody in depressed participants reveal changes in symptom severity over the course of depression.3) Also, Interpersonal effects occur such that vocal prosody in interviewers shows corresponding effects. These hypotheses marked variation in the depression scores and has detected the range of severity. Thus the analysis of vocal prosody could be a powerful tool to detect depression.

3. PROPOSED MODEL

Normally, people don't consider depression as a serious medical illness and leave it untreated. But this is not the case. It should be identified and treated at an earlier stage itself otherwise it may affect a person's peaceful state of mind. Fig1 shows how the model works.

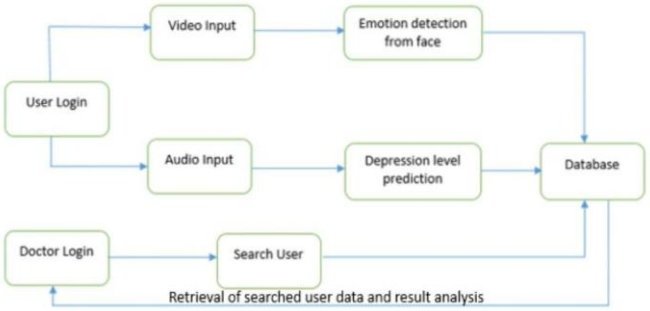

Fig -1: Work flow of the system

Here both the patient and the doctor have to create and login to their accounts. When the patient login he/she is asked to attend either audio or video session. Tkinter is used for the GUI design.

4. SYSTEM DESIGN AND IMPLEMENTATION

4.1 Dataset collection

Haar Cascade classifier is used for face detection and thereby emotion analysis. Haar Cascade is an object detection algorithm where a lot of positive and negative images are used to train the classifier. This dataset contains a number of XML files to detect different body parts. Haarcascade_frontalface_default.xml is used to detect frontal face. Also, here we use a CSV file to train the system. This file contains dataset classifying depressive and non depressive text.

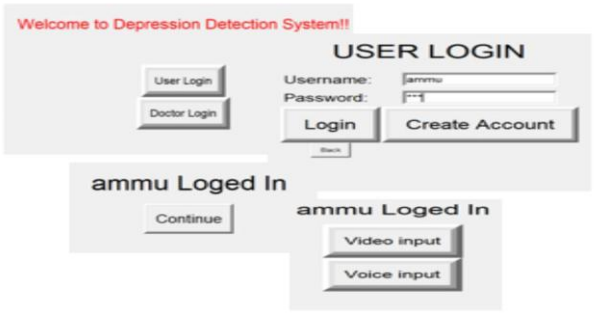

Fig -2: User Login

4.2 Face detection and emotion recognition

The real time video is captured by using a system's webcam. OpenCV is used for detecting faces from this captured video. It is a library which aims at real-time computer vision and is used for image processing. It helps in the detection of shapes in an image. Then the CNN algorithm is used to read the Haar Cascade xml files to recognize facial features. CNN is mainly used for image classification [2].Thus CNN together with the XML files recognize the facial features. These facial features are then converted to grey scale images. Then we use a trained model for emotion detection, which contains a large set of data related to emotions such as happy, sad, neutral, fear, angry. We use DNN to read this trained model. The grey scale images obtained from the face detected are then compared with this trained model. Thus the emotions are identified.

4.3 Audio recognition and depression level prediction

The user is provided with a set of Questionnaires containing 20 questions. For each question, he/she has to respond using a microphone. The speech recognition module is used for this purpose. It is a library which can be used to recognize speech input from the microphone. For each question, the responses made are classified into depression and normal by using Naive-Bayes classifier, provided the CSV file as the training set.

Fig -3: Questionnaires

Naive-Bayes classifier supports text classification. For each response made, a score is assigned based on the emotion detected. After the completion of the last question, the total score is evaluated.

4.4 Result Analysis

The scores obtained from the audio input along with the emotion detected from the face are then used to find out whether a patient is normal or what type of depression a patient is having. Thus the system predicts 5 major levels of depression. They are major, chronic, manic, seasonal and situational depression. All the results are then stored in a database.

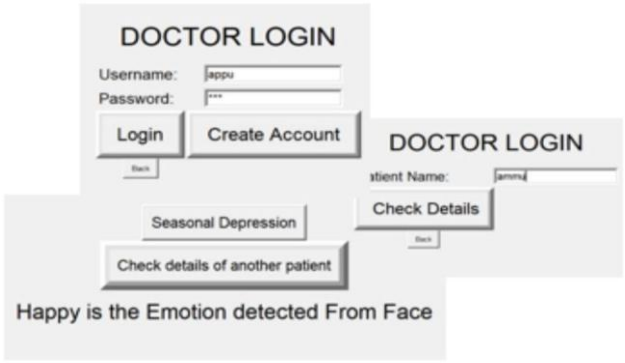

Fig -4: Results

The doctor can retrieve the results from the database by searching the name of the patient. Thus he could obtain the results, i.e. the patient's type of depression and can thereby suggest further medications. This could save a lot of his time.

5. CONCLUSION AND FUTURE LIKELIHOODS

In this paper, an artificial intelligence system was proposed for automatic depression detection. This could save a lot of the doctor's time and effort. We have studied and analyzed various tools and technologies for the successful completion of the project. This work is based on facial and audio expressions. Deep learning techniques are used here. The system identifies five major levels of depression. We have come across certain issues while running the audio and video input together. That was one of the biggest problems we have encountered with. There are limitations within the system that affects its performance. Here, depression is measured by the limited number of questions asked and the responses made. Responses made using different microphones also contribute to the error rate. The system developed is a desktop one. We would like to convert it into an application in order to make it accessible to a large crowd of people. All these things can be considered for further improvement of the system.

6. REFERENCES

[1] Wheidima Carneiro de Melo, Eric Granger, Abdenour Hadid ; Depression Detection Based on Deep Distribution Learning ,IEEE International Conference on Image Processing (ICIP)(2019).

[2] Le Yang; Multi-Modal Depression Detection and Estimation, 8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW) (2019).

[3] Prajakta Bhalchandra Kulkarni ,Minakshee M. Pati ;Clinical Depression Detection in Adolescent by Face, International Conference on Smart City and Emerging Technology(ICSCET)(2018).

[4] Asim Jan, Hongying Meng, Yona Falinie Binti A. Gaus and Fan Zhang, Artificial Intelligent System for Automatic Depression Level Analysis through Visual and Vocal Expressions, IEEE Transactions on Cognitive and Developmental Systems (2018).

[5] Mohamad Al Jazaery, Guodong Guo; Video-Based Depression Level Analysis by Encoding Deep Spatiotemporal Features, IEEE Transactions on Affective Computing (2018).

[6] Mandar Deshpande, Vignesh RaoMa Xianjun ; Depression detection using emotion artificial intelligence, International Conference on Intelligent Sustainable Systems (ICISS)(2017).

[7] Shamla Mantri, Dipti Patil,Pankaj Agrawal, Vijay Wadhai; Cumulative video analysis based smart framework for detection of depression disorders, International Conference on Pervasive Computing(ICPC)(2015).

[8] Ying Yang, Catherine Fairbairn, Jeffrey F. Cohn; Detecting Depression Severity from Vocal Prosody, IEEE Transactions on Affective Computing (2013).