DEPRESSION DETECTION SYSTEM USING PYTHON

Prof.G.Harinatha Reddy[1], G.Meghana[2], D.Panidhar[3],D.Narayana Nanda Neeraj[4], Ch.Vinaykumar[5]
[1,2,3,4,5] *Department of ECE , N.B.K.R. Institute of science and and technology,*
Tirupati district, Andhra Pradesh, India

Abstract- Depression is a leading cause of mental illness and is associated with an increased risk of premature death. In addition, it is a major contributor to suicidal ideation and causes significant disruption to daily life. Every year, one in 15 adults suffer from depression, which affects 300 million people worldwide. Several previous empirical studies have shown that certain language characteristics can be analyzed and correlated with likely symptoms of depression as well as help predict self-destructive behavior. This depression detection system detects the percentage of depression and recommends a consultation with a psychiatrist at nearby clinics. In addition, the user must talk about himself for 1 minute while his facial expression is recorded. A CNN and a unique dataset for Depression faces will be created for Image & Video. The programming language used in this system is Python. Html, CSS and JavaScript are used on the front end and MySQL is used on the back end. Here we used the flask framework. Humans have a stronger sense of emotion and feeling that can be combined with technology to create useful tools. This depression detection system is extremely useful for facilitating self-assessment of depression and improving diagnostic accuracy.

Keywords: Depression, Machine learning, accuracy, CNN

1. INTRODUCTION

1.1 Depression:

Depression is a state of mental illness. It is characterized by deep, long-lasting feelings of sadness or despair. Depression can change an individual's thinking/feelings and also affects their social behavior and sense of physical well-being. It can affect people of any age group, including young children and adolescents. It can run in families and usually starts between the ages of 15 and 30.

Women and the elderly are affected more often than men. There are several types of depression, such as major depression, it is a change in mood that lasts for weeks or months. It is one of the most severe types of depression. Dysthymia (chronic depression) is a less severe form of depression but usually lasts for several years. Psychotic depression a severe form of depression associated with hallucinations and delusions (feelings that are untrue or unfounded). Seasonal depression, occurring only at a certain time of the year, usually winter, also known as the "winter blues".

Depression is thought to be caused by an imbalance of certain brain chemicals called 'neurotransmitters', which carry signals in the brain that the body uses to control mood. Some of the common factors that can cause depression are genetics (inherited), trauma and high levels of stress, mental illness such as schizophrenia and substance abuse, postpartum depression (women can develop depression after having a baby), serious medical conditions , such as heart disease, cancer and HIV, the use of certain medications, alcohol and drug abuse, individuals with low self-esteem, trauma and high levels of stress due to financial problems, the breakdown of a relationship or the loss of a loved one.

Signs and symptoms of depression include feeling sad and lonely, losing interest in activities you once found enjoyable, feelings of hopelessness, worthlessness or excessive guilt, tiredness or loss of energy, too little or too much sleep, loss of appetite, restlessness and being easily angered.

If Left untreated, depression can lead to some serious complications such as emotional, behavioral, health and even legal/financial problems, relationship difficulties, social isolation and even suicide.

1.2 Convolutional Neural Network:

In the field of deep learning, the convolutional neural network belongs to the class of deep neural networks that have been most widely deployed in the field of image analysis/recognition. Convolutional neural uses a very special kind of method known as convolution. The mathematical definition of convolution is a mathematical operation applied to two functions that produces an output in the form of a third function that shows how the shape of one function is affected, modified by another function.

Convolutional neural networks consist of different layers of artificial neurons. Artificial neurons, similar to those neuron cells that the human brain uses to relay various sensory input signals and other responses, are mathematical functions that are used to calculate the sum of various inputs and provide an output in the form of an activation value.
The behavior of each CNN neuron is defined by the value of its weights. When artificial neurons of CNN are fed with (pixel) values, they recognize different visual features and specifications, when we give an input image to CNN, each of its inner layers generates different activation maps. Activation maps draw attention to relevant features of a given input image. Each of the CNN neurons generally receives an input in the form of a group/area of a pixel, multiplies their values (colors) by the value of their weights, sums them, and passes them through the appropriate activation function. the first (or perhaps bottom) layer of the CNN typically recognizes various features of the input image, such as edges horizontally, vertically, and diagonally. The output of the first layer is fed as the input of the next layer, which in turn extracts other complex features of the input image, such as corners and edge combinations. The deeper one gets into a convolutional neural network, the more the layers begin to detect various higher-level features such as objects, faces, etc.

2. LITERATURE SURVEY

Aliaa A. A. Youssif, Wesam A. A. Asker presents a computer vision system for automatic facial expression recognition (AFER). AFER has three main steps, the first step is face detection in the scene. The second step is to extract the facial features that depict the facial expression, and the third step is to classify the facial image displayed on the face. Face detection uses an open source code library (OpenCV) that uses a face detection algorithm based on Viola & Jones features. After that, facial feature extraction is performed, where a segmentation process is first performed to divide the face image into three regions: mouth, nose, and two eyes and two eyebrows. Second, facial feature points (FCPs) are in each part of the face using FCP extraction techniques from the mouth, nose, eyes, and eyebrows. A feature extraction process is applied to the face image to create a feature vector that consists of two types of features: geometric features and appearance features that depict the pattern for facial expression classes. Then the feature vector is given as input to a radial basis neural network for facial expression recognition The results show that the AFER system classifies facial expressions accurately with a recognition rate between 90% and 99% in the person-dependent dataset and between 83% and 100% in data file independent of the person.

Authors Enrique Correa, Arnoud Jonker, Michael Ozo, and Rob Stolk designed their paper on emotion recognition using a convolutional neural network. This method ranges from several hundred high-resolution photos to tens of thousands of smaller images. In order to increase the accuracy of the detected emotions, the size of the training dataset must be increased from 9000 images to 20000 images from FERC. The obtained results are compared with other methods such as SVM and LVQ. It achieves an accuracy of 90% happy, 80% neutral and 77% surprised.

Authors Kartika Candra Kirana, Slamet Wibawanto and Heru Wahyu Herwanto proposed emotion detection using the Viola Jones Algorithm in their paper. Although Viola Jones is commonly used for face detection, here the Viola Jones algorithm is used for both face detection and emotion recognition. The rectangular element and the cascaded AdaBoost algorithm are used as the main concept of the Viola-Jones algorithm in both processes. These processes use Russel's Circumplex to classify emotions because it has better efficiency in classifying emotions. This method consists of 3 stages: initially, a frame is captured from the video; the unwanted rectangular areas are deleted and then the emotion in the image is recognized. The prediction provided an accuracy of 74%.

Yogita et al. presented a multilingual speech-to-text system using the Mel-Frequency Cepstral Coefficient (MFCC) feature extraction technique and the minimum distance classifier method. Support vector machine (SVM) for speech classification.

3. REQUIREMENTS SPECIFICATION

3.1 Hardware Requirements:

Processor: Any Processor above 500 MHz
Ram: 8 GB
Hard Disk: 4 GB
Dedicated Graphics Card
Input device: Mouse and Camera.
Output device: High Resolution Monitor.

3.2 Software Requirements:

Operating System: Windows 7 or higher
Programming language: Python and related libraries.
Dataset is obtained from Kaggle. The data consists of 48x48 pixel grayscale images of faces.
The training set consists of 28,709 examples.

4.IMPLEMENTATION

4.1 Convolutional Neural Network

The architecture of a ConvNet is analogous to that of the connectivity pattern of Neurons in the Human Brain and was inspired by the organization of the Visual Cortex. The Convolutional Layer and the Pooling Layer together form the i-th layer of a Convolutional Neural Network. Depending on the complexities in the images, the number of such layers may be increased for capturing low levels details even further, but at the cost of more computational power.

1.Facial Expression dataset:

There are many open accesses facial expression dataset in literature. We have used dataset for facial expression from Kaggle and the data has 48x48 pixel grayscale images of faces. The training set consists of 28,709 examples with seven emotions (happy, sad, surprised, fearful, angry, disgusted, and neutral).

2. Image Preprocessing:

Face perimeter was detected using Haar Cascade library from images. Then these detected rectangular facial expressions were cropped and recorded. The images were also converted to grayscale and placed into neural networks. This process was done to avoid unnecessary density in the neural networks.

3.**Convolutional neural network architecture:**

The proposed CNN architecture is focused on learning pixel values in a rectangular region containing facial expressions. This occurs in 3 stages, after which it is fed into fully bonded layers. The CNN structure consists of facial expression data and includes 2 stages. The first stage consists of 2 convolutional layers with relay activation and a max pooling layer, and the 2nd stage consists of 2 convolutional and max pooling layers with Relu activation. After all the operations of convolution layers and max-pooling layers, each image is transferred to fully connected layers, and the prediction of the images was processed by a classifier as seven different facial emotional states

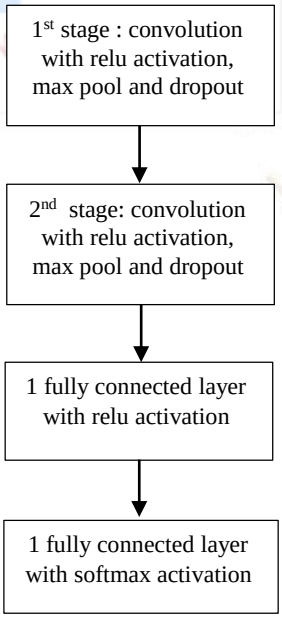

Fig:CNN model architecture - the layers used in the construction of CNN model

4.Network training:

The neural networks were implemented using Keras with a TensorFlow backend running in Python. The model was trained for 50 epochs.

5.Speech to text conversion:

Speech was converted to text using assemblyAI. AssemblyAI is the fastest way to build with AI for audio. AssemblyAI's API can be used to transcribe and understand audio/video files with AI models.

6.Real time testing:

After the training of proposed CNN architecture, the trained model was tested in real time. First, human faces were detected by the computer camera using Haar Cascade library. After that, the detected images were sent to the model and the classes they belong to were queried.

As a result of the predictions, the possibility of belonging to which class the facial expression was shown on the webcam screen.

OpenCV is used to draw a rectangular boundary around the face detected and the emotions recognised is displayed on the screen with an emoticon indicator on the window with the percentage confidence of that emotion.

Meanwhile in the video, the speech was converted to text

Table 4.1.1 - Modules used

Modules used	**Functionality**
Data Pre-processing	The data required to train and test the model is processed and divided into training and test dataset. The images are modified for keras model.
Model definition	The CNN Model code is defined using the required ML Libraries with 2 fully formed layers after the operations of 4 convolutional and 3 max pooling layers.
Real time Emotion Recognition	The real time classification of emotion is done with the help of OpenCV to access video with a webcam and the emotion recognition classification is shown on the screen

Analysis:

The video will be recorded and it will be uploaded into the webpage where the system takes 50% of prediction data from images and 50% of prediction data from audio.By taking the average from those two values, the depression percentage will be presented on the monitor. If the depression percentage exceeds 60% then the green colour in the screen turns red and if it is less than 60% the screen remains green. In addition to this, doctor details will be provided in the webpage.

5. RESULTS AND DISCUSSION

We trained our Convolutional Neural Network model using Kaggle database which includes seven emotions (happiness, anger, sadness, disgust, neutral, fear and surprise) The detected face images are resized to 48×48 pixels, and converted to grayscale images then were used for inputs to the CNN model. We achieved an accuracy rate of 85% for the emotion recognition.

In this paper, a link will be provided by running the source code which redirects the user to the webpage in which he/she can record a live video or upload an existing video from their files. After uploading the video to the webpage, the data will be extracted and depression percentage will be presented on the screen along with the text.By clicking on find a doctor icon ,the doctors details will be visible.

Fig: Working of depression detection system on a webpage

6.CONCLUSION

A depression detection System has a wide range of applications in psychological research and human interaction applications. The system plays a communicative role in interpersonal relations because they can reveal the affective state, cognitive activity, personality, intention, and psychological state of a person. The system has 3 modules-face detection that is implemented by Haar Cascade, emotion recognition which is implemented by CNN using Keras that mainly focuses on detecting emotions that can reflect depression in an individual,speech to text conversion which can be done by AssemblyAI API.

7.FUTURE WORK

The project mainly deals with depression percentage calculation by using emotion detection and speech conversion. As of the current working, the system takes considerable time to show the depression percentage. In the future days, the time taken can be reduced by modifying the modules and layers used.

The current system is basically a screening test before consulting a doctor. In the future days, a video consultancy to doctor can be arranged if the user is detected to be depressed.

.

REFERENCES

[1] Automatic Facial Expression Recognition System Based on Geometric and Appearance Features- Aliaa A. A. Youssif, Wesam A. A. Asker

[2] Enrique Correa, Arnoud Jonker, Michael Ozo and Rob Stolk proposed their paper of emotion recognition using Convolutional Neural Network.

[3] Emotion Detection using Viola Jones Algorithm- Kartika Candra Kirana, Slamet Wibawanto and, Heru Wahyu Herwant

[4] A Survey Paper on Chatbots, Aafiya Shaikh, Dipti More, Ruchika Puttoo, Sayli Shrivastav, Swati Shinde

[5] Video based Emotion Recognition using Deeply Supervised Neural Networks by Yingruo Fan, Jacqueline C.K Lam, Victor O.K Li

[6] Dr. D.Venkataraman, Namboodri Sandhya Parameswaran, "Extraction of Facial Features for Depression Detection among Students",International Journal of Pure and Applied Mathematics,2018.

[7] Ma Xiaoxi Lin Weisi,Huang Dongyan,Dong MinGhui,Haizhou Li, "Facial Emotion Recognition",IEEE,2017.

[8] Facial expression recognition based on Local Binary Patterns: A comprehensive study by Caifeng Shana, Shaogang Gongb, Peter W. McOwanb

[9] Nicu Sebe, Michael S,Lew ,Ira Cohen, Ashutosh Garg, Thomas S Huang,"Emotion Recognition using Cauchy Naive Bayes Classifier",IEEE,2002.

[10] Kia-Biao He,Jing Wen,Bin Fang, "Adaboost algorithm using MB-LBP and skin color segmentation",IEEE,2011.

[11] Head Pose and Movement Analysis as an Indicator of Depression, Sharifa Alghowinem, Roland Goecke, Michael Wagner, Gordon Parker, Michael Breakspear

[12] E.M Bouhabba, A.A Shafie, R.Akmeliawati, "Support Vector Machine for Face Emotion Detection on real time basis",IEEE,2011.