Proceedings of the Third International Conference on Intelligent Sustainable Systems [ICISS 2020]
IEEE Xplore Part Number: CFP20M19-ART; ISBN: 978-1-7281-7089-3

# Mental Health State Detection Using Open CV and Sentimental Analysis

Manav Pradeep Jain
*Department Of Information Technology*
*Dwarkadas J.Sanghvi College Of Engineering*
Mumbai, India
manavmessi25@gmail.com

Soumyaprakash Sribash Dasmohapatra
*Department Of Information Technology*
*Dwarkadas J.Sanghvi College Of Engineering*
Mumbai, India
som01123579@gmail.com

Professor Stevina Correia
*Department Of Information Technology*
*Dwarkadas J.Sanghvi College Of Engineering*
Mumbai, India
Stevina.Dias@djsce.ac.in

2020 3rd International Conference on Intelligent Sustainable Systems (ICISS) | 978-1-7281-7089-3/20/$31.00 ©2020 IEEE | DOI: 10.1109/ICISS49785.2020.9315984

***Abstract*—The Mental health of an individual refers to his/her's cognitive, behavioral, and emotional well-being. In simple terms, it can be described as how people feel, behave, think, and how they handle situations. Mental health is an important aspect of an individual's well being. Several factors damage mental health, namely stress, anxiety, etc. If proper care is not given to a person suffering from mental illness the consequences may be bad. With technology at its peak nowadays one can harness the powers of a computer to predict the onset of such mental illnesses. Applications implemented using machine learning and Artificial intelligence algorithms can be used to predict such onsets and can serve as a monitor for any unusual behavior in individuals. The main aim of the paper is to predict Mental Health state of an individual. The paper discusses the system in detail which helps in doing so. The system consists of 4 modules namely -1) Pulse-based Depression detection 2) Facial Emotion Detection 3) Questionnaire 4) BOT Assistant. The output of all the models is combined and by further analysis ,the final outcome is given in the form of prediction of the user's state of mind .**

***Index Terms*—Machine Learning, CBT(Cognitive Behavioral Therapy),OpenCV, Mental Disorder, Sentimental Analysis,**

## I. Introduction

Health is a state of complete physical, mental, and social well-being and not merely the absence of disease or infirmity [7]. Humans will simply focus on their physical health. Unfortunately, in many parts of the planet, mental state and mental disorders are not given similar importance as physical health. Rather they have been for the most part unnoticed or neglected. Psychological disorder, if not taken care of, will cause disability, permanent loss of memory, manipulation, or perhaps self-harm.

According to the World Health Organization(WHO) Report in 2019, Globally, an estimated 264 million people are affected by depression[8]. The recognition of mental disorders as a major health problem has increased in primary care settings[6]. These disorders have a huge impact on those who are suffering from it as well as on their families[6]. They diminish the quality of life as they result in impaired functioning as well as are accompanied by an increased need for health care[6]. Depression is among the leading causes of disability in the world .Doctors are still not able to properly diagnose depression, with roughly half the cases that are diagnosed by primary care physicians receive adequate minimal treatment.Several low- and middle-income countries allocate less than 2% of their health budget to the treatment and prevention of mental disorders[1].

The system explained in this paper is not only required in India but also around the world. Now take the example of India, where due to excessive population the ratio of doctors to a patient is 1:1456 [9] and the amount of the time that doctors have to treat a patient is less than five minutes. To detect any mental disorders, this amount of time is very small and hence a reliable system is required which is accurate in screening, takes lesser time, and can be used by a large cohort. Recently, in India, the death of Famous Bollywood actor Sushant Singh Rajput triggered everyone to think in this direction. Going to a psychiatrist is considered a stigma and that is why many people refrain to visit one. There is also a taboo that a person dealing with depression or stress is considered a dumb or psyche. The system explained in the paper helps a person to treat themselves by knowing their mental health state condition. It also helps in a way that an individual doesn't have to go to a doctor and would not be outcast by society. It can be easily used as a preliminary test for the detection of Mental State and can act as an early detection of symptoms leading to mental disorders.

The system aims at helping the basic investigation to detect the individual's state of mind. This will save the doctor's time being invested in preliminary tests. It has a wide range of applications like testing the mental health of students in school and colleges, assisting doctors in preliminary tests, testing mental stability in the organization, and also in the common

public. It can also be used to aid the medical check-up for other related diseases. For e.g. stress and depression can also lead to a rise in blood pressure. Once tested that the mental health of an individual is not good, he shall be advised to visit a psychologist in person or through our website in virtual mode thereby reducing the risk of mental disorder at early stages.

## II. Related Work

This section will go through the different approaches of two main modules in our system such as Facial Emotion Detection and Pulse Rate Detection . There is a comparative study done for different approaches available for Facial Emotion Detection which is shown in Table 1.Based on it, Histogram of Oriented Gradient (HOG) + CNN approach is used in our project . Also , for the second module Pulse Rate Detection different approaches are mentioned:- Costa et al. [1] made use of camera images.They tried to make use of color variation of the skin to extract the physiological parameters.These approaches did not produce any quantitative results. Additionally,it also failed to show any kind of correlation with ECG signals. Kenneth et al [2] developed a non-contact system capable of capturing two PPG (Photoplethysmogram) signals that are at different wavelengths at the same time.The experiment was conducted on ten test subjects.The camera and the PPG sensors were used for data collection.The system extracted oxygen saturation efficiently but its efficiency is unknown. Verkruysse et al. [3] presented a system that made use of simple,inexpensive digital cameras to extract the heart rate and the respiratory rate of the individual from his face.The system takes a time of about thirty seconds to give the output and therefore it can be used in real time applications.The non-contact based method were improved further .The most recent one is in 2015 , Rahman et al. in 2015 [12] developed a simple web camera based method to detect Heart Rate. It uses ROI technique, where RGB color values of each pixels of facial image frame is detected.Finally ,Fast Fourier Transform(FFT) is used to obtain Power Spectrum which corresponds to highest power of spectrum within an operational frequency band.The benefit of it is that , it helps in monitoring the physiological signals such as heart rate in a unobtrusive manner.This kind of non-invasive methods are also cost effective as they can reduce the cost of cabling and equipment which is required in a regular procedure.It can be used efficiently for neonatal monitoring,sleep studies etc where the continuous measurement of heartbeat is important. Since it monitors Heart Rate for unlimited time , it overcomes the drawbacks of previous works which monitor for certain amount of time and in a controlled environment .The system uses this method to detect pulse rate from web camera.

TABLE I
Comparison of Algorithms

| Module | Approaches | Comments | Limitations |
|---|---|---|---|
| Facial Emotion Detection | 3D motion based feature + Hidden Markov Model (HMM). | 1) Features are extracted and classified using Gentleboost classifier such as AdaBoost and the result is used to build temporal models of each expression using an HMM .<br>2) The obtained results show that use of 3D information enhances the accuracy when compared to 2D data. | As HMM is a generative model, it has a problem of Local Optimal solution Trap. |
| | HMM+SVM | 1) The Limitation of the HMM model was improved by hybridizing a discriminant classifier.<br>2) HMM was hybridized to capture the temporal dynamics and employed SVM as a multiclass classifier of the features on a frame by using frame basis. | It performed better on binary problems rather than multiclass problems like facial emotion recognition. |
| | Local Binary Pattern(LBP) +Linear Programming | 1) It's benefit lies in its simplicity in computation, high tolerance for low image resolution, invariant to illumination changes.<br>2) It improved the accuracy for the JAFFE database but it could not be generalized as it was implemented only on one database. | It is challenged with factors like rotation, increase in computational complexity, small sample size and limited information representation. |
| | Hand crafted Preprocessing stages + CNN | 1) It achieved a better accuracy than the other approaches mentioned above because of enhancing preprocessing stages like resizing, face detection, cropping, adding noises, data normalization which consist of local, global contrast and histogram equalization.<br>2) It was trained on CK (posed) ,JAFFE (posed) and MUG(posed) database and achieved an average accuracy of 97.06%. | It tends to over fit because of overly significance to preprocessing stages. |
| | Histogram of Oriented Gradient (HOG) +CNN | 1) It extracted features from the active facial patches and fed them to the convolutional neural network.<br>2) It gives better performance than other models and it showed an average accuracy of 95% | It increases the time of computation as it is computationally intensive |
| | Hybrid(LBP and HOG)+ Softmax | 1) The fusion of extracted LBP and HOG features reduced the extracted features dimensionality with PCA.<br>2) The fused feature was permuted on several classifiers and achieved a maximum accuracy of 98.3% using Softmax classifier.<br>3) The result is evidence that hybrid features could enhance the performance significantly. | Research is still going on in finding the best combination of features. |

## III. Proposed Solution

The Proposed Solution consists of deriving results using OpenCV and Sentimental Analysis. In OpenCV two modules are being used, One is Heart pulses rate module where heart rate is measured of an intended user.The principle behind it is, to extract heart rate data from facial skin color variation caused by blood circulation. The heart rate is measured in BPM(beats per minute). This rate measured is then used to verify the user's state with the assistance of another term, HRV (Heart Rate Variability) this is often not a brand new development and is an accepted term in the medical fraternity[11]. Heart rate variability or HRV is that the physiological development of the variation within the measure between consecutive heartbeats in milliseconds. A normal, healthy heart doesn't tick equally sort of a pendulum, however instead, once viewing the milliseconds between heartbeats, there's constant variation. In general, it tends to acutely tuned in to this variation; it's not an equivalent because the heart rate(beats per minute) increasing and decreasing as it tends to set about our daily business. Reliable HRV analysis needs the correct measure of every heartbeat and therefore the time between beats. There are completely different technologies for conniving HRV and one among them is ECG-based methodology wherein you'll be needing user ECG report for locating HRV and therefore the different methodology additionally needs different reports which may solely be deduced from hospitals and labs. Since, this method is dealing with BPM, the relationship between HRV and BPM is used to find out values corresponding to BPM and then evaluating threshold values for detecting mental disorders [11].

The second module in Open CV is Facial Emotion Detection. This module uses deep convolutional neural networks to classify the emotion detected on the person's face.The FER-2013 dataset which is present in Kaggle is used to train our model. This dataset consists of face images portraying seven

emotions - Angry, Disgusted,Sad,Neutral, Happy,Very Happy and Joyful. It detects emotions on all faces in front of the camera. A 4 layer CNN is used with the help of which a test accuracy of 63.2 percent is achieved in 50 epochs. The Algorithm used is:- 1)First, the faces that are present in each frame of the webcam feed are detected using the haar cascade method. 2)The region of the image that contains the face is given as an input to the CNN. 3)The CNN gives a list of softmax scores for the seven classes of emotions as the output. 4)The emotion which has the maximum softmax score is displayed on the screen. The result of this is just used in the analytical process to improve the accuracy of the overall process[4].

Next is the sentimental analysis which is done by using a questionnaire and a chatbot. In the questionnaire, there are ten exclusive questions that are used to find an individual's state of mind[5]. These questions are in the form of MCQ's having 3 answers to them. The answers are ranked from 0-3 [10]. Now, the answers to the question are stored and given one of the point from 0-3 depending on the answer chosen by the individual. Since there are ten questions the sum of the points of all the questions answer will be averaged and the final result is stored. The questions are part of CBT(Cognitive behavioral therapy)which are used by most Psychological doctors to check the state of mind of an individual. After this, there will be a chatbot, which tries to further understand the individual's psychology. The chatbot is trained using social media dataset like twitter using NLP and bags of word is created. This created bag has words with positive, neutral and negative emotions. Thus, the overall chat of the individual is taken as an input in the form of a paragraph and skimmed through to check with the words present in the bags of word. The number of positive, negative, and neutral words are counted for an individual.

## IV. Architecture

The architecture consist of 4 parts which come under the broad category of Psychological Screening test. These parts are basically 4 steps in screening test . Each step is a module in itself and the result derived from individual module is used for further analysis to derive the final result. The following block diagram in Figure 1 shows the Architecture and each module will be explained in detail.

## V. Pulse Based Depression Detection

It uses Open CV to search out the situation of the user's face, then isolates the forehead region. Information is collected from this location over time to estimate the user's rate, this is often done by measurement average optical intensity within the forehead location, within the sub image's inexperienced channel alone (a higher color mixture quantitative relation could exist, however, the blue channel tends to be noisy). Physiological information is calculable in this manner due to the optical absorption characteristics of (oxy-) Hb.

With better lighting and lowest noise, a stable heartbeat ought to be isolated in about fifteen seconds, an alternative

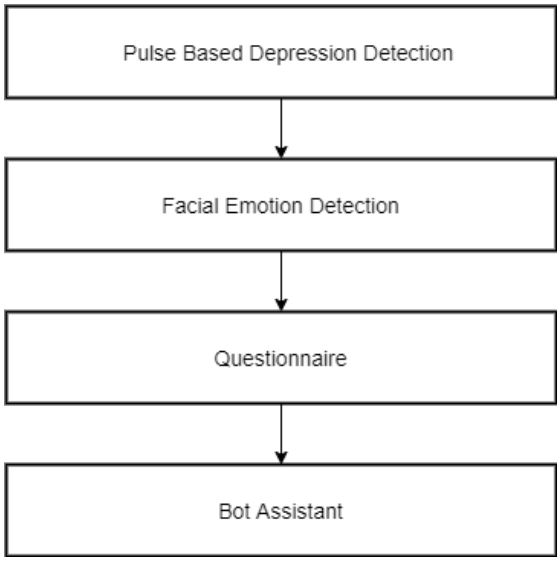


Fig. 1. Architecture of Proposed Model

physiological waveform (such as Mayer waves) ought to be visible within the information stream. Once the user's rate has been calculable, the period of time section variation related to this frequency is additionally computed. This permits for the heartbeat to be exaggerated within the post-process frame rendering, inflicting the highlighted forehead location to pulse incorrect with the user's heartbeat. There is support for the detection of multiple simultaneous people in a very single camera's image stream, however, at any instant, only the data from one face is extracted for analysis.

Implementing the above principle using a simple method shown in Figure 2: 1) RGBsplitter and Grayscale with contrast equations helps in detecting the face , aligning and highlighting it.

2) Region of Interest (ROI) is obtained using the technique of facial landmarks.

3)Then applying band pass filter as a threshold with fl = 0.8 Hz and fh = 3 Hz, which are 48 and 180 bpm respectively.

4) After calculating the mean color value of ROI in each frame, provide it to a data buffer which is 150 in length.

5)Fast Fourier Transform (FFT) the data buffer, the greatest peak is nothing but Heart Rate. The result of this implementation is shown in Figure 3. The heart rate of an individual is scaled to an integer value from 0-3 using the Table 2.

## VI. Facial Emotion Detection

The first thing is to collect the dataset. The model is trained and build on the FER-2013[10] data set available on Kaggle. This dataset contains face images with seven emotions - Angry,Disgusted,Sad,Neutral,Happy,Very Happy and Joyful. It is a three-step process to implement it. The first step is the implementation of OpenCV HAAR CASCADES. The"Frontal Face Alt" Classifier is used for detecting the presence of Face in the Webcam. The Second step is Training the Network with the help of the TensorFlow Image Classifier. The Third step is

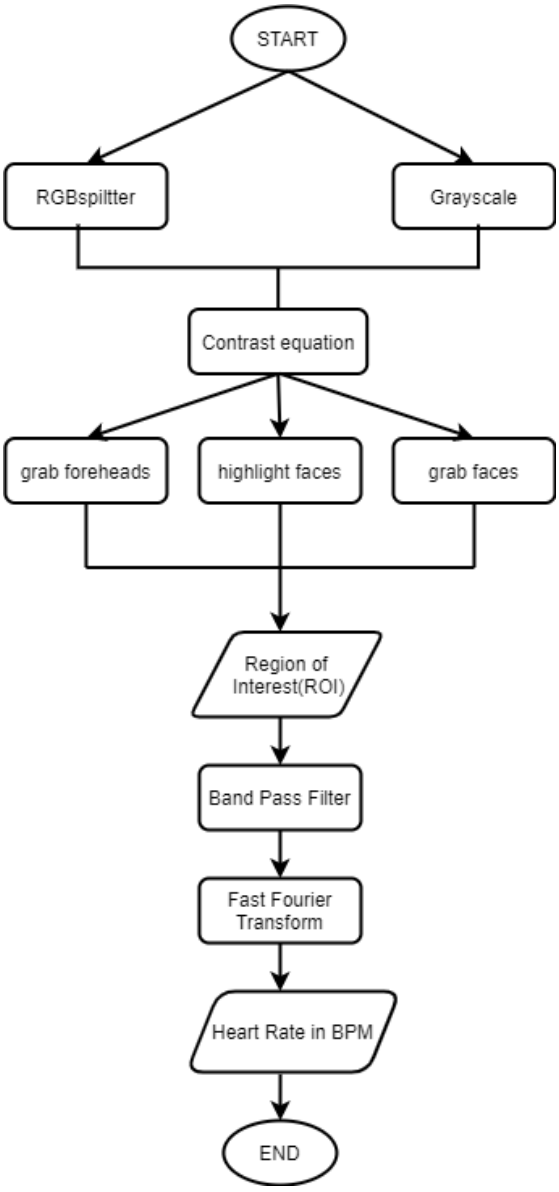


Fig. 2. Detection of Heart Rate

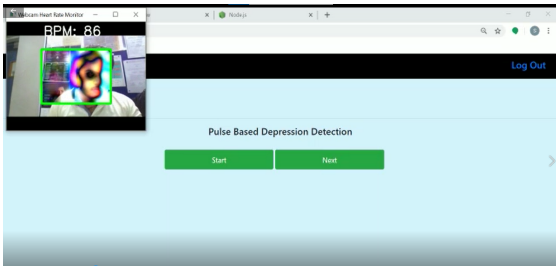


Fig. 3. Output of Heart Rate in BPM

TABLE II
SCALED VALUE FOR HEART RATE

| Heart Rate in BPM | Integer Equivalent |
|---|---|
| Below 60 | 0 |
| 60-70 | 0.5 |
| 70-80 | 1 |
| 80-90 | 1.5 |
| 90-95 | 2 |
| 95-100 | 2.5 |
| Above 100 | 3 |

Running the trained model. The algorithm is described in the proposed solution in depth. Using opencv,the face is detected, then by using dlib, the face landmarks are then extracted. Also the HOG features are extracted and inputs the raw image data with the face landmarks+hog into a convolutional neural network as shown in Figure 4. The output from this module is shown in Figure 5. The output in the form of label is obtained and then converted into an equivalent integer form using the information from Table 3.

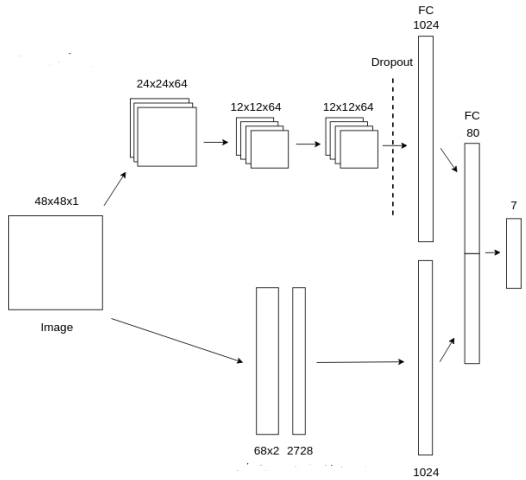


Fig. 4. CNN Architecture

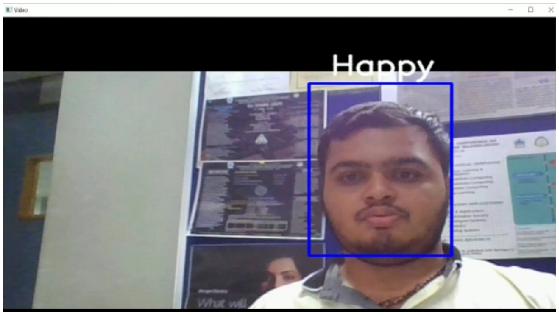


Fig. 5. Output of Facial Emotion Detection

TABLE III
SCALED VALUE IN EMOTIONS

| Value | Emotions |
|---|---|
| 0 | Angry |
| 0.5 | Disgusted |
| 1 | Sad |
| 1.5 | Neutral |
| 2 | Happy |
| 2.5 | Very Happy |
| 3 | Joyful |

## VII. QUESTIONNAIRE

The questionnaire forms the part of rule based learning technique where 10 questions were selected after research which is mostly used by Psychological doctors to check the state of mind of an individual[5]. The questions are in an MCQ format and have three answers. The answers have been given a predefined range from (0-3) [10]. Based on the answers to the question by an individual, the value is given. The sum of values from 10 questions is then averaged to get the final result. The final value will be in the range of 0-3 and the values are rounded off to the nearest half value. E.g. if the value is 1.37 it will be round off to the nearest half value that is 1.5. Now this final value from Questionnaire module will be used in further Analysis as shown in Table 5 for deducing the mental state of an individual.

## VIII. BOT ASSISTANT

The bot forms the part of Context based learning .Now, the bot was developed using dialog flow, which again checks the individual's state of mind but the process is different from the questionnaire. The bot's responses are programmed in such a way as to reflect properly on the user's chat and understand clearly what the user's intention is and proceed with that intention in its responses. In this, the answers given by the user is split into word by word. The responses are then converted into words. All the unnecessary words like (the, it) are removed and the responses are trimmed and cut short. Now using sentimental analysis on data from twitter, reviews, and web crawling, a bag of words is created which consists of both positive and negative words. This bag of words acts as a trained dataset that is useful for deducing sentiment from the user's chat. The modified response from the user is checked with the bag of words. Since we are concerned with finding the mental state of an individual, we focus on the negative words rather than positive. A word count for negative words is created as seen in Table 4. The total no of word count for negative words from each sentence is obtained and the final value is obtained in the form of integer from range of 0-3. This

TABLE IV
NEGATIVE WORDS COUNT

| Post | Recognized word | Wordcount |
|---|---|---|
| Let me struggle this lack alone | **Struggle** | 3 |
| | **Lack** | |
| | **Alone** | |
| Having anxiety and depression is like being scared and tired | **Anxiety** | 4 |
| | **Depression** | |
| | **Scared** | |
| | **Tired** | |
| Wanting to be alone, but not wanting to be lonely | **Alone** | 2 |
| | **lonely** | |

value will then be used as shown in table 5 for predicting final result.

## IX. RESULT

The web application is created for the same. It has in total, 6 features as shown in Figure 6 that is 1)Psychological Screening test - this feature is clearly explained in detail in the above paper. 2)Spell the color! not the word, this is nothing but a game that will help relax a person and also distract his mind so that there will be no external factor into consideration while taking the screening test. 3)Local Mental Health Centers - this is a feature to locate nearby Mental health Center and it uses the googles map API .4)CBT techniques - this is to educate users about different CBT techniques to reduce stress and increase positivity in life .5)Motivate Yourself - this feature includes motivational videos, songs, and quotations to motivate the user not to feel low in life .6)Talk to Psychiatrist - this gives the user the ability to directly talk to a psychiatrist, it uses webRTC in NodeJS for audio and video communication in real-time directly between peer to peer without any third-party plugins. The interface is depicted in Figure 6. Once the user does the screening test which includes 4 steps, the analysis is done on the output of each step, an average integer value is derived which isthen scaled using table 3 . Now, different values on this scale represents different emotions.The final output in the form mental state of an individual is obtained as shown in table 5.

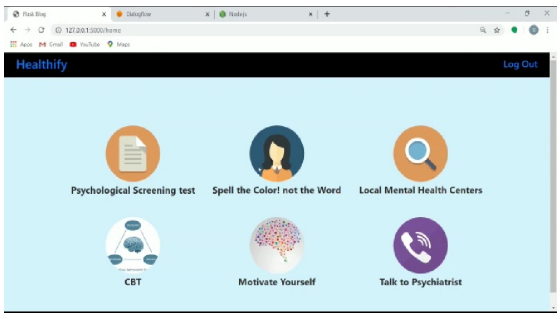


Fig. 6. Interface of Web Application

TABLE V
FINAL RESULT

| Pulse based Detection | Facial emotion Detection | Questionnaire | Bot Assistant | Average | Emotion Detected |
|---|---|---|---|---|---|
| 2.5 | 2 | 2 | 1.5 | 2 | Happy |
| 2 | 2.5 | 2 | 3 | 2.375~2.5 | Very Happy |
| 1.5 | 1 | 1 | 2 | 1.375~1.5 | Neutral |
| 0.5 | 1 | 0.5 | 1 | 0.75~1 | Sad |

## X. FUTURE SCOPE

1) The work on improving the accuracy of the sentimental analysis that is, giving weights to negative words so that we can get different values for negative words according to the context in which they are used and the strength of negativeness they contain. This would help us to analyze in a more enhanced way.

2) This scope is more on the feature side that is to include a mood-based recommendation of songs according to the analysis done on individual's state of mind after going through the Psychological Screening Test.

## XI. CONCLUSION

In conclusion, this project helps to detect the Mental state of an individual. The paper gives an analysis of different techniques and algorithms used in this project to deduce the results. It also demonstrates the related work on the same topic and their approaches in solving along with the unique approach used in the project. It gives a detailed overview of the proposed solution and architecture. The paper also describes each and every module used in the Psychology Screening test. This paper helps us in understanding the features, its result, and the overall project..

## REFERENCES

[1] G. D. Costa, "Optical remote sensing of heartbeats," Optics Communications, vol. 117, pp. 395-398, 6/15/ 1995.
[2] [5]K. Humphreys, T. Ward, and C. Markham, "Noncontact simultaneous dual wavelength photoplethysmography:noncontact pulse oximetry," Rev Sci Instrum, vol. 78, p. 044304, Apr 2007.
[3] W. Verkruysse, L. O. Svaasand, and J. S. Nelson, "Remote plethysmographic imaging using ambient light," Optics express, vol. 16, pp. 21434-21445, 2008
[4] PawełTarnowski, Marcin Kołodziej, Andrzej Majkowski, Remigiusz J. Rak,Emotion recognition using facial expressions,Procedia Computer Science,Volume 108,2017,Pages1175-1184,ISSN 1877-0509,https://doi.org/10.1016/j.procs.2017.05.025.
[5] Kaj Sparle Christensen, Per Fink, Tomas Toft, Lisbeth Frostholm, Eva Ørnbøl, Frede Olesen, A brief case-finding questionnaire for common mental disorders: the CMDQ, Family Practice, Volume 22, Issue 4, August 2005, Pages 448–457, https://doi.org/10.1093/fampra/cmi025
[6] Ogunsemi, O. O., Oluwole, F. A., Abasiubong, F., Erinfolami, A. R., Amoran, O. E., Ariba, A. J., Alebiosu, C. O., & Olatawura, M. O. (2010). Detection of mental disorders with the Patient Health Questionnaire in primary care settings in Nigeria. Mental illness, 2(1), e10.https://doi.org/10.4081/mi.2010.e10.
[7] Preamble to the Constitution of WHO as adopted by the International Health Conference, New York, 19 June - 22 July 1946.
[8] World Health Organisation, Mental Disorder, 28th November 2019, https://www.who.int/news-room/fact-sheets/detail/mental-disorders.
[9] Samiksha Goel, DHNS,Economic Survey: Doctor-population ratio is 1:1456,Bengaluru, JAN 31 2020 ,https://www.deccanherald.com/business/budget-2020/the-doctor-population-ratio-in-india-is-11456-against-who-recommendation-800034.html
[10] National Institute of Mental Health. (2019). Anxiety. Retrieved from https://www.nimh.nih.gov/health/publications/anxiety/index.shtml on May 22, 2020.
[11] Jerzy Sacha, M.D,Interaction between Heart Rate and Heart Rate Variability,2014,https://onlinelibrary.wiley.com/doi/pdf/10.1111/anec.12148.
[12] H. Rahman, M. U. Ahmed, and S. Begum, "Noncontact Physiological Parameters Extraction using Camera," in The 1st Workshop on Embedded Sensor Systems for Health through Internet of Things (ESS-H IoT), Oct., 2015.