IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

IJARSCT
Impact Factor: 6.252

Volume 2, Issue 7, May 2022

Depression Detection Using Deep Learning

Prasad Khodave[1], Sanket Godge[2], Shreyas Jadhav[3], Prof. Mr. Uttam Patole[4]
Students[1] & [4]Prof. of Department of Computer Engineering,
Sir Visvesvaraya Institute of Technology, Nashik, Maharashtra, India
Savitribai Phule Pune University, Pune, Maharashtra, India

Abstract: *The proportion of individuals with depression has rapidly increased along with the growth of the global population. Depression has been the currently most prevalent mental health disorder. An effective depression recognition system is especially crucial for the early detection of potential depression risk. A depression-related dataset is also critical while evaluating the system for depression or potential depression risk detection. Due to the sensitive nature of clinical data, the availability and scale of such datasets are scarce. Depression is classified as a mood disorder. It may be described as feelings of sadness, anger, or loss that interfere with a person's everyday activities. People experience depression in different ways. In certain cases, depression may lead to fatal cases. To avoid all of these, depression must be detected at the earliest and victim must be treated with appropriate remedies. The objective of the project is to analyze the emotion of a user using real-time video. This is achieved using Convolutional Neural Networks [CNN]. The final decision result comes from the combination of the two models. Finally, we evaluate all proposed deep models on our built dataset. The experimental results demonstrate that (1) our proposed method can identify patients/users with potential depression risk; (2) the recognition performance of combined 2D and 3D features model outperforms using either 2D or 3D features model only; (3) the performance of depression recognition is higher in the positive and negative emotional stimulus. Meanwhile, we compare the performance with other methods on the same dataset.*

Keywords: Convolutional Neural Networks [CNN], Machine Learning, Depression Detection, Face Recognition, Classification, Dataset, etc

I. INTRODUCTION

Human face and human behavioral pattern play an important role in person identification. Visual information is a key source for such identifications. Data analytics can be used for a wide variety of applications like motion detection, human activity prediction, person identification, abnormal activity recognition, vehicle counting, people counting at crowded places, etc. In many situations humans who are depressed are totally ignorant of their disturbed mental condition. They are unable to identify the cause of constant unhappiness in them and eventually such students fall into a state of mind where they start having suicidal tendencies. In some cases, students do know that they are suffering from depression, but they are hesitant to seek any kind of help from anyone mainly due to the wrongly conceived notion of 'humiliation' associated with depression. It is better to identify the signs of depression at initial stages of depression. Depression if identified in the initial stages, just a simple one hour talk with a counselor may be of immense help for the student. This may totally change the negative state of mind of that student into a positive one. Such a student can be given good counseling of how to deal with mental stress and can be guided to follow the right path to success. The most important form of non-verbal communications is facial expressions of a person. Many studies have been done for finding out the facial expressions that are related to depression. The current work is mainly undertaken to find out the presence of depression in college students by studying their facial features. This system mainly uses different image processing techniques for face detection, feature extraction and classification of these features as depressed or non-depressed. The

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Impact Factor: 6.252

Volume 2, Issue 7, May 2022

system will be trained with features of depression. Then videos of different students with frontal face will be captured using a web camera. Then the facial features of these faces will be extracted for prediction of depression. Based on the level of depression features the student will be classified as depressed or non-depressed. Feature selection methods should be classified to different aspect categories like wrapper, static and hybrid techniques. In filter based approaches, the selection of features can't reliant of any machine learning algorithm. In this, features are preferred on the base of their numerical weight. In the dynamic approach, first different subsets of features are identified then are evaluated using one of the classifiers. The hybrid approach is the combination different feature extraction as well features selection methods; it also used various machine learning algorithms, and In-univariate filter approach features are evaluated with respect to relevance. Multivariate method considers correlation between features and avoids redundant features. We are proposing a filter approach which selects relevant features and avoids redundant ones.

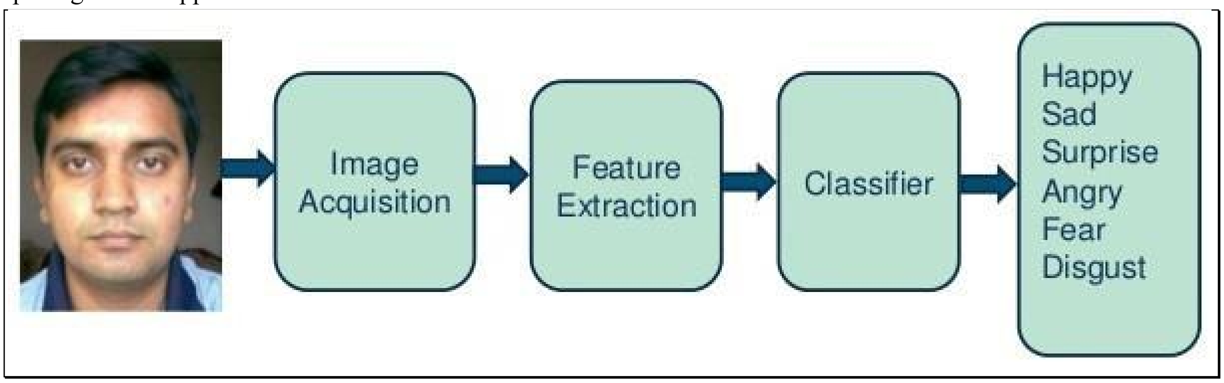


Fig.1: System Overview

II. RELATED WORK

According to the World Health Organization (WHO), more than 350 million people of all ages suffer from depression disorder globally (Reddy, 2012). Depression (depressive disorder or clinical depression) is one of the most severe but prevalent mental disorders globally. Depression can induce severe impairments that interfere with or limit one's ability to conduct major life activities for at least 2 weeks. During at least 2 weeks, there is either a depressed mood or a loss of interest or pleasure, as well as at least four other symptoms that reflect a change in functioning, such as problems with sleep, eating, energy, concentration, self-image, or recurrent thoughts of death or suicide. Depression can occur at any age, and cases in children and adolescents have been reported1. Because of its harmfulness and recent prevalence, depression has drawn increasing attention from many related communities. Although it is severe, depression is curable through medication, psychotherapy, and other clinical methods (National Collaborating Centre for Mental Health, 2010). The earlier that treatment can begin, the more effective it is. Thus, the early detection of depression is critical to controlling it at an initial stage and reducing the social and economic burden related to this disease. Traditional diagnosis approaches for depression are mostly based on patients self-reporting in clinic interviews, behaviors reported by relatives or friends, and questionnaires, such as the Patient Health Questionnaire (PHQ-9) (Kroenke and Spitzer, 2002) and the Beck Depression Inventory (BDI-II) (McPherson and Martin, 2010). However, all of them utilize subjective ratings, and their results tend to be inconsistent at different times or in different environments. During the diagnosis, several clinical experts must be involved to obtain a relatively objective assessment. As the number of depressed patients increases, early-stage diagnosis and re-assessments for tracking treatment effects are often limited and time consuming. Therefore, machine learning-based automatic potential depression risk detection or depression recognition is expected to facilitate objective and fast diagnosis to ensure excellent clinical care quality and fundamentally reduce potential harm in real life.

Under the influence of depression, behavior disorder-based signals for depression recognition are increasingly extensive, such as voices (Ooi et al., 2013; Yang et al., 2013; Nicholas et al., 2015; Jiang et al., 2017), facial expressions (Schwartz et al., 1976; Babette et al., 2005), gestures (Alghowinem et al., 2018), gaits (Michalak et al., 2009; Demakakos et al., 2015), and eye movements (Winograd-Gurvich et al., 2006; Alghowinem et al., 2013; Carvalho et al., 2015). This work focuses on using facial expressions to recognize patients with potential depression risk. The research on depression based on facial expression essentially utilize video or images (Gupta et al., 2014; Alghowinem, 2015; Pampouchidou et al., 2015, 2016a; Bhatia et al., 2017). To be more precise, the interests are localized on images, facial landmark points (Stratou et al., 2014; Morency et al., 2015; Nasir et al., 2016; Pampouchidou et al., 2016b), and/or facial action units (AUs) (Cohn et al., 2009; McIntyre et al., 2009; Williamson et al., 2014). However, the methods that adopt image analysis (the essence of the video-based method are still images analysis where videos are converted into images) are affected by environmental factors and instrument parameters, such as illumination, angle, skin color, and resolution power. If these factors are not addressed appropriately, the recognition performance will be affected. Several researchers (Gong et al., 2009; Zhao et al., 2010) proposed using in-depth information captured from 3D sensors, which is relatively illumination, angle, and skin color invariant. However, 3D points of information can lose the texture features of facial expression. Therefore, the fusion of 2D with 3D data can make up for each other to address these issues.

Depression recognition typically comprises two steps: feature extraction and recognition (depression or not/ depression severity). The quality of feature extraction directly affects the result of recognition. Conventional feature extraction methods for depression facial expression utilize geometric features, appearance features, and dynamics. These methods extract the displacement of facial edges, corners, coordinates (McIntyre et al., 2010; Bhatia et al., 2017), mean squared distance of all mouth landmarks to the mouth centroid (Gupta et al., 2014), and displacement from the mid-horizontal axis to depict the changes and intensity of basic expressions (Bhatia, 2016). The local binary pattern (LBP), LBP-TOP (Joshi et al., 2012), local Gabor binary pattern (LGBP-TOP) (Sidorov and Minker, 2014), local curvelet binary pattern (LCBP-TOP) (Pampouchidou et al., 2015), and LPQ from three orthogonal planes (LPQ-TOP) (Wen et al., 2015) extracted describe the texture changes in the facial region. Histogram of optical flow (Gupta et al., 2014), motion history histogram (MHH) (Meng et al., 2013), and space-time interest points (STIPs) (He et al., 2015) are extracted to describe the facial motion. These results indicated that depressed people display a lower performance when responding to positive and negative emotional content. Nevertheless, all those mentioned approaches are hand-crafted feature descriptors designed based on tremendous professional knowledge, and image processing is also complicated for hand-crafted features. However, our cognition of depression remains insufficient. Such features probably yield segmented representations of facial expressions and are insufficiently discriminative. Simultaneously, the dynamics are extracted from a video, which involves the effect of the environmental factors mentioned above. On the other hand, the time window is used to extract motion features (Pampouchidou et al., 2016a; He et al., 2018). The reported window lengths are 60 frames, 20 frames, 5 frames, or even 300 frames. However, the optimal window length cannot be determined because there are significant variations over time in the facial expression according to the particular person and experimental device.

In recent years, deep learning techniques have prevailed in audio- and video-based applications, especially in visual information processing (Girshick et al., 2014). The purpose of this study is to identify the patients at risk of depression. The selected subjects are outpatients, and the evaluated depression degree is moderate. Many samples with depression risk and the normal control group had no noticeable expression changes in some stimulation tasks. Therefore, we chose the generative model deep belief network (DBN). The DBN-based deep learning method can hierarchically learn good representation from original data; thus, the learned facial features should be more discriminative than hand-crafted features for depression recognition. Long short-term memory (LSTM) is an effective and scalable model for learning problems related to sequential data and can capture long- term temporal dependencies. Facial expression is a dynamic process of continuous change, and it is a time-series signal on a timeline. Then facial expression motion is captured by LSTM used on the entire timeline.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 7, May 2022
Impact Factor: 6.252

This paper builds on our previous work (Guo et al., 2019) by adding 2D static image information and 3D facial point motion information to identify depression, and it is a further improvement and summary of the original work. We build two different deep networks, respectively, one of which extracts static appearance feature using 2D images based on DBN, and the other learns the facial motion via 3D facial landmark points and facial AUs using DBN-LSTM. The two kinds of deep networks are then integrated by joint fine-tuning, which can further improve the overall performance. Therefore, our main contributions in this paper can be summarized as follows:

- We designed a reasonable and effective experimental paradigm, collected diversified data and three kinds of samples (normal population, outpatients, and inpatients) combined with specialized hospitals, and constructed a large-scale dataset for depression analysis.
- The two deep networks proposed can extract appearance features from 2D images and motion features from 3D facial landmark points. The integrated networks can achieve the fusion of static and dynamic features, which can improve recognition performance.
- We have proved qualitatively and quantitatively that depressive prone groups show significant differences from healthy groups under positive and negative stimuli.

III. PROPOSED SYSTEM

Depression detection from images alone, mainly depends on a clear and proper definition of a depressed face. The facial expression of a depressed face is slightly different from that of sadness. A depressed face expression has the same characteristics of a sad expression, such as the upward slanted eyebrows etc. but the main difference is that there is no major frown involved. Also a sad face may have eyes lowered looking downward showing the helpless, dejected mood. On the contrary a depressed person can put forth a face devoid of depression. This depicts a case of concealed expression of depression, i.e. the depressed face may not be a sad face, and instead the person may put forth a happy face to conceal depression.

Based the above analysis, it can be said that, most of the work done for the recognition of the emotion 'depression' from faces is done on databases that included only adult patients. But as mentioned above college students are more vulnerable to depression, and detection of depression at the college level can give more time to control it and better counseling can be suggested. Thus a system can be proposed that collects the video of college students for analysis of depression.

The current study proposes a system that will help in detecting depression in college students. The system will be trained with features of happy, contempt and disgust faces. Then in the testing phase videos of college students will be collected while they are answering different questionnaires. The students' facial features will be extracted and normalized for effective detection of features throughout the video. Then the extraction of facial features would be done for the test dataset, classifying them by NB classifier for depression detection.

Depression detection will be done by overall presence of happy, contempt and disgust features throughout the video frames. If the presence of happy features is low, and based on the presence of contempt or disgust features, the student will be classified as having mild, moderate or high depression. The level of depression will be found out by amount of negativity in the video. If the level of happy features is less, the amount of negativity will be more. The student will be classified as highly depressed, if the amount of negativity is high; mildly depressed if the negativity level is moderate and not depressed if the amount of negativity is much less. The architectural diagram of the proposed automated system can be modeled in the following way.

DOI: 10.48175/IJARSCT-4364

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Impact Factor: 6.252
Volume 2, Issue 7, May 2022

A. SYSTEM ARCHITECTURE

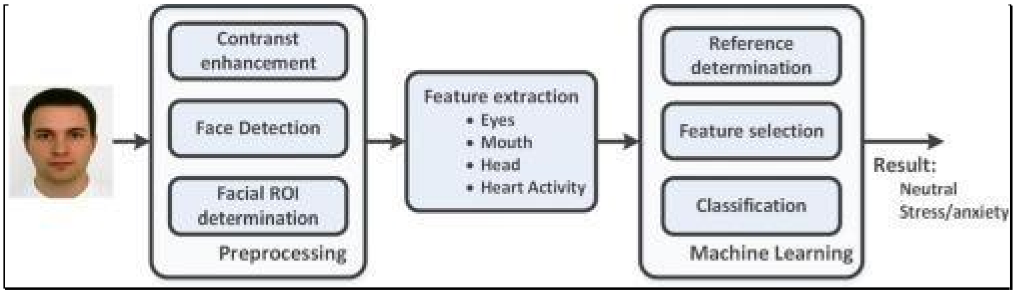
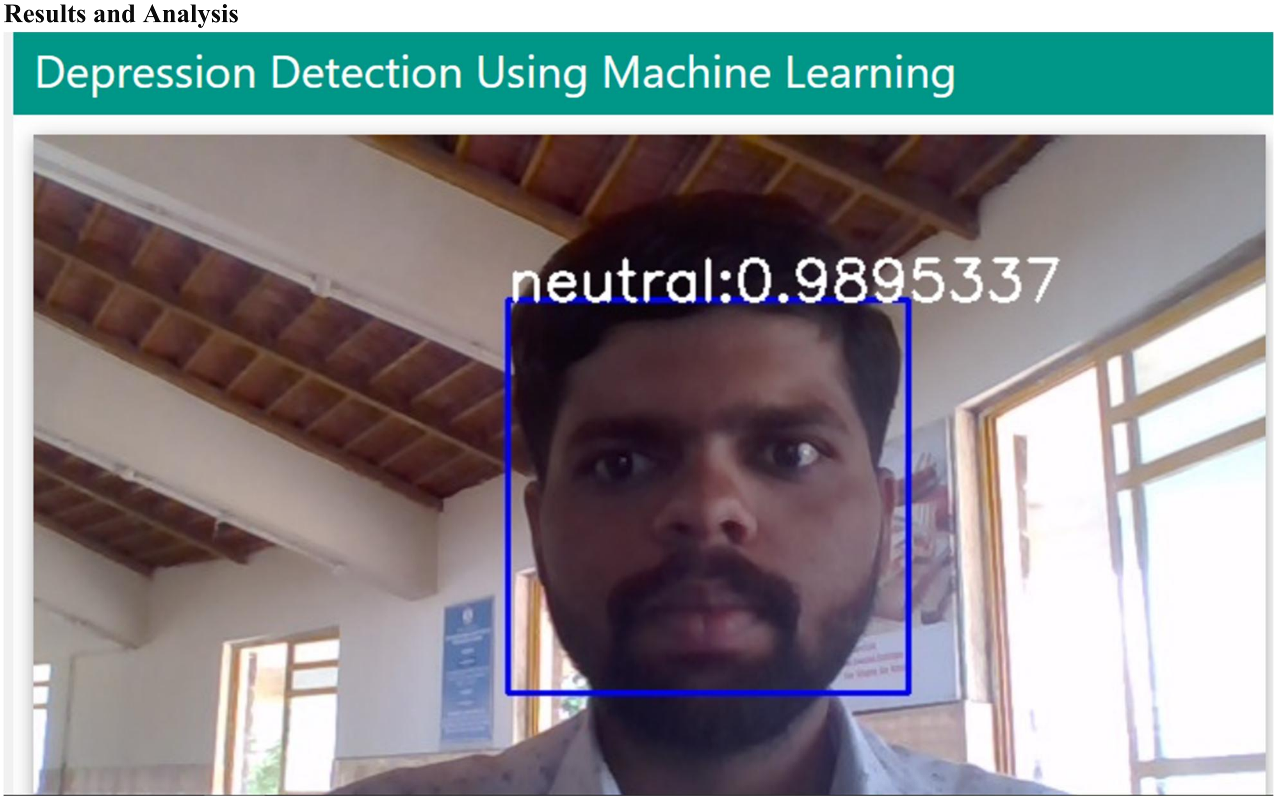


Fig.1: System Architecture

In this project to address the problem of stress detection three modules have been mainly defined in order to measure the differences of stressed and non-stressed users on social media platforms: System Framework, Social Interactions and Attributes categorization. In the proposed systemic approach, we formulate the task as a classification problem to detect two types of detection of psychological disorders in social networks using the machine learning framework: i.e. first is Stress and second is Depression.

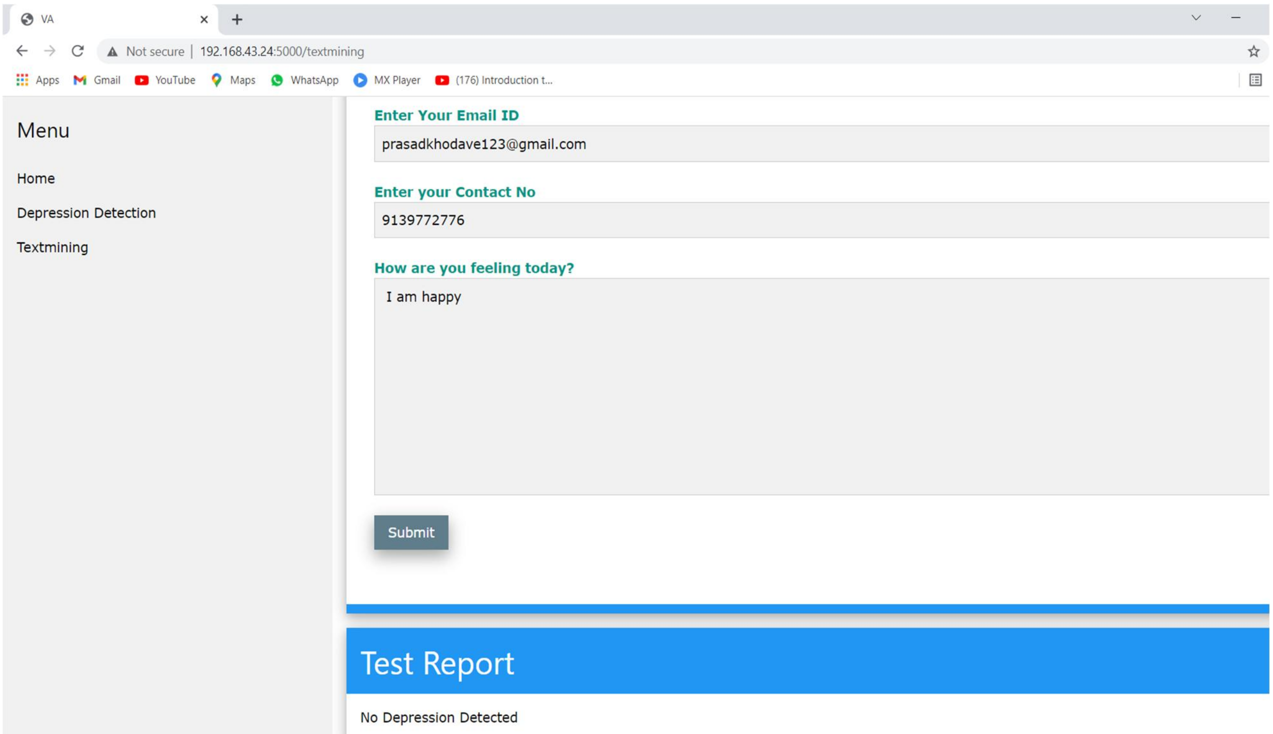
IV. RESULTS AND DISCUSSION

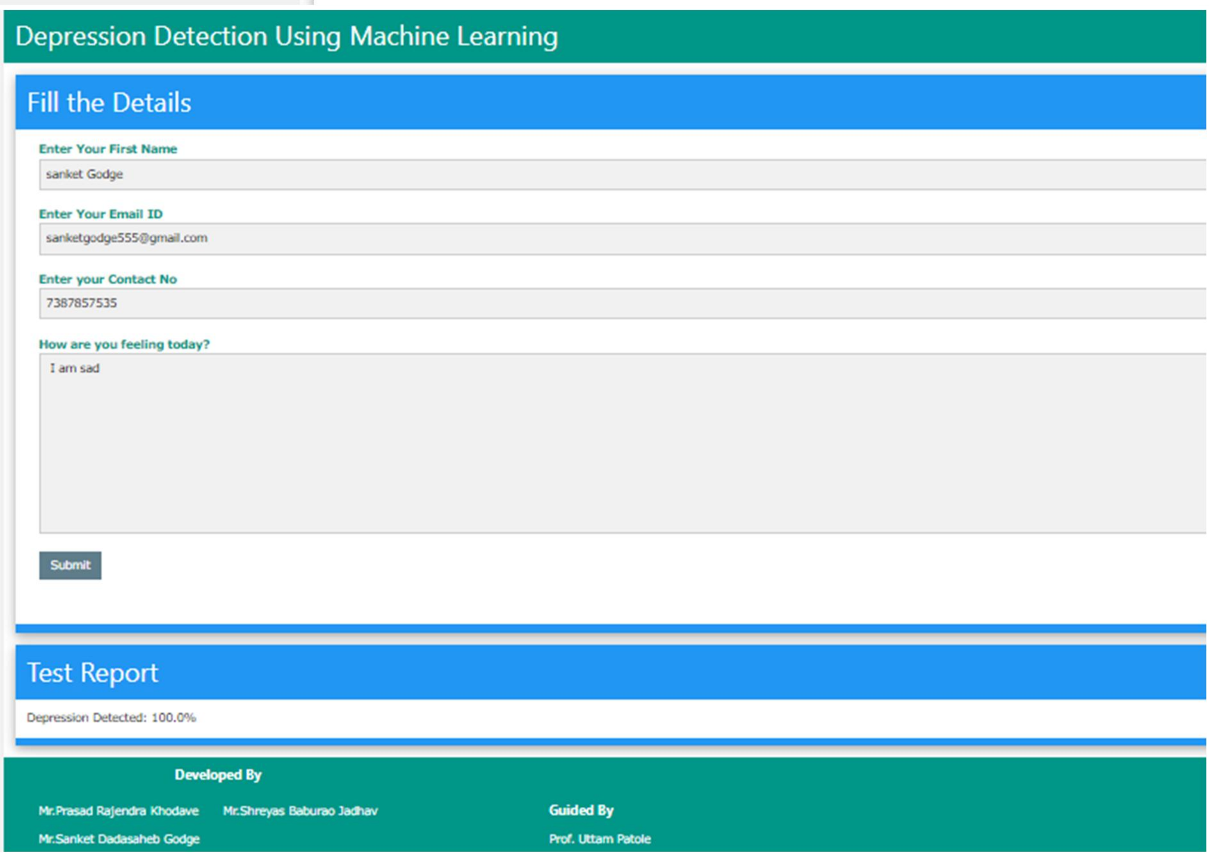
Results and Analysis



DOI: 10.48175/IJARSCT-4364

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
IJARSCT
Impact Factor: 6.252
Volume 2, Issue 7, May 2022





The system can provide the below outcomes once successfully executed training as well as testing phase

- System gets background knowledge as a training set
- System can able to find weight vector and define the similarity with respective domain.
- Reduce false positive ratio.
- Each cluster has categorized into multiple similar clusters, under the one master cluster.
- Finally, similarity score will classify each bucket into the respective domain.

V. CONCLUSION AND FUTURE WORK

An automated Facial Expression Recognition System has a wide range of applications in psychological research and human-computer 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 three 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. Finally, the last module, a text analysis is used that is used to recognize depression that further helps to differentiate between sadness and depression.

ACKNOWLEDGMENT

We would prefer to give thanks the researchers likewise publishers for creating their resources available. We are conjointly grateful to guide, reviewer for their valuable suggestions and also thank the college authorities for providing the required infrastructure and support.

REFERENCES

[1] Girard, Jeffrey M., Jeffrey F. Cohn, Mohammad H. Mahoor, Seyed mohammad Mavadati, and Dean P. Rosenwald. "Social risk and depression: Evidence from manual and automatic facial expression analysis." In Automatic Face and Gesture Recognition (FG), 10th IEEE International Conference and Workshops on, pp. 1- 8. IEEE, 2013

[2] Alghowinem, Sharifa, Roland Goecke, Jeffrey F. Cohn, Michael Wagner, Gordon Parker, and Michael Breakspear. "Cross-cultural detection of depression from nonverbal behaviour". In Automatic Face and Gesture Recognition (FG), 11th IEEE International Conference and Workshops on, vol. 1, pp. 1-8. IEEE, 2015

[3] Pampouchidou, A., O. Simantiraki, C-M. Vazakopoulou, C. Chatzaki, M. Pediaditis, A. Maridaki, K. Marias et al. "Facial geometry and speech analysis for depression detection." In Engineering in Medicine and Biology Society (EMBC), 39th Annual International Conference of the IEEE, pp. 1433-1436. IEEE, 2017

[4] Harati, Sahar, Andrea Crowell, Helen Mayberg, Jun Kong, and ShamimNemati. "Discriminating clinical phases of recovery from major depressive disorder using the dynamics of facial expression."In Engineering in Medicine and Biology Society (EMBC), 38th Annual International Conference of the, pp. 2254- 2257. IEEE, 2016

[5] Cohn, Jeffrey F., Tomas Simon Kruez, Iain Matthews, Ying Yang, Minh Hoai Nguyen, MargaraTejera Padilla, Feng Zhou, and Fernando De la Torre. "Detecting depression from facial actions and vocal prosody."In Affective Computing and Intelligent Interaction and Workshops. ACII 2009. 3rd International Conference on, pp. 1-7. IEEE, 2009

[6] Tasnim, Mashrura, RifatShahriyar, NowshinNahar, and Hossain Mahmud. "Intelligent depression detection and support system: Statistical analysis, psychological review and design implication." In e-Health Networking, Applications and Services (Healthcom), 18th International Conference on, pp. 1-6. IEEE, 2016

[7] Pampouchidou, Anastasia, Kostas Marias, ManolisTsiknakis, P. Simos, Fan Yang, and Fabrice Meriaudeau. "Designing a framework for assisting depression severity assessment from facial image analysis."In Signal and Image Processing Applications (ICSIPA), International Conference on, pp. 578-583. IEEE, 2015

[8] Maddage, Namunu C., Rajinda Senaratne, Lu-Shih Alex Low, Margaret Lech, and Nicholas Allen. "Video-based detection of the clinical depression in adolescents."In Engineering in Medicine and Biology Society, (EMBC).

Annual International Conference of the IEEE, pp. 3723-3726. IEEE, 2009
[9] Meng, Hongying, Di Huang, Heng Wang, Hongyu Yang, Mohammed AI-Shuraifi, and Yunhong Wang. "Depression recognition based on dynamic facial and vocal expression features using partial least square regression." In Proceedings of the 3rd ACM international workshop on Audio/visual emotion challenge, pp. 21-30.ACM, 2013.