ISSN 2395-1621

Depression Detection Using Artificial Intelligence through Visual And Text Expressions

Assitant Prof. Prajwal Gaikwad, Maariyaah Shabbir Sayyad, Shekhar Bhorade, Stuti Thorat, Navin Kakani

AISSMSIOIT, Pune, Email:prajwaltrk@gmail.com AISSMSIOIT,Pune, Email: sayyadmaariyaah603@gmail.com AISSMAIOIT, Pune, Email:shekharbhorade1000@gmail.com AISSMSIOIT, Pune, Email:stutithorat@gmail.com AISSMSIOIT, Pune, Email:kakani.navin@gmail.com

ABSTRACT

Psychological health plays a very impor- tant role in every person's life. Neglecting this can result in several problems such as stress, depression and etc. These problems need to be detected and controlled at the early stages of life for the better mental health. Depression is considered to be one of the leading causes of mental ill health and it has been found to increase the risk of early deaths. Moreover it is a major cause of suicidal tendencies and this may to lead significant impairment in a person's daily life. Detecting depression is one of the most challenging task. Most of the people are totally unaware that they may have any depression caused due to some stress in the daily life. If at all people are aware of it then some people conceal their depression from everyone. So an automated system is required which will pick out people who are suffering from depression. A system has been proposed which will analyse facial features of the person from an image and will help in detecting signs of depression if present in them. This system will be trained with frontal face images of happy and disgust faces and will then classify them as neutral or negative based on the word-list to detect depression tendencies.

ARTICLE INFO

Article History

Received: 8th March 2020

Received in revised form : 8th March 2020

Accepted: 10th March 2020

Published online : 11th March 2020

I. INTRODUCTION

MENTAL health issues such as depression have been linked to deficits of cognitive control. It affects one in four citizens of working age, which can cause significant losses and burdens to the economic, social, educational, aswell as justice systems. Depression is defined as a common mental disorder that presents with depressed mood, loss of interest or pleasure, decreased energy, feelings of guilt or low self-worth, disturbed sleep or appetite, and poor concentration. Among all psychiatric disorders, major depressive disorder (MDD) commonly occurs and heavily threatens the mental health of human beings. 7.5 percent of all people with disabilities suffer from depression, making it the largest contributor, exceeding 300M people. Recent study indicates that having a low income shows an increased chance of having MDDs. It can also affect the major stages in life such as educational attainment and the timing of marriage. According to , majority of the people that obtain treatment for depression do not recover from it. The illness still remains with the person. This may be in the form of insomnia, excessive sleeping, fatigue, loss of energy, or digestive problems. Artificial intelligence and mathematical modeling techniques are being progressively introduced in mental health research to try and solve this matter. The mental health area can benefit from these techniques, as they understand the importance of obtaining detailed information to characterize the different psychiatric disorders [7]. Emotion analysis has shown to been an effective research approach for modeling depressive states. Recent artificial modeling and methods of automatic emotion analysis for depression related issues are extensive. They demonstrate that depression analysis is a task that can be tackled in the computer vision field, with machine-based automatic early detection and recognition of depression is expected to advance clinical care quality and fundamentally reduce its potential harm in real life.

A. Overview

Various procedures used for data collection, and existing datasets are summarized in this project.Summary outlines some methods and algorithms for visual feature extraction and dimensionality reduction and decision methods for clas- sification and regression approaches for depression detection. Quantitative analysis of reported re- sults relying on the perfor- mance metrics is included, identifying general trends and the key unresolved issues to be considered in future studies of the automatic depres- sion assessment utilizing visual cues alone or in combination method with the visual cues. The proposed work is carried out to predict the depression level according to given input of the face images using deep learning.

1) Motivation: Recent classification techniques which are present run the risk of confusing normal sadness with de- pression and may raise a false diagnos result. Depression assessment is generally considered as a complex process. Diagnosis of this depression is associated with a significant degree of uncertainty and the symptoms within the person's current psycho-social context needs to be evaluated based on the past history. The current work assessment is mainly undertaken to find out the presence of depression in the college students by studying their facial features. The proposed system mainly uses image processing techniques for face detection and also feature extraction.It also classifies the features as depressed or non-depressed one. The aim of this paper is to build an artificial intelligent system that can automatically predict the depression level from a user's visual and vocal expression. The system is understood to apply some basic concepts of how parts of the human brain works. This can be applied in robots or machines to provide human cognitive like capabilities, making intelligent human–machine applications. The main contribution of the proposed framework are as follows: 1) a framework architecture is proposed for automatic depression scale prediction that includes frame/segment level feature extraction, dynamic feature generation, feature dimen- sion reduction, and regression; 2) various features, including deep features, are extracted on the frame-level that captured the better facial expression information; 3) a new feature (FDHH) is generated by observing dynamic variation patterns across the frame-level features; 4) advanced regressive techniques are used for regression.

II. RELATED WORK

Recent years have witnessed an increase of research for clinical and mental heatlh analysis from facial and text expressions. There is a progress on emotion recognization from facial expression and text expression. We proposed a computational approach in which we can detect depression by face expression and also depression by text expression. We are going to detect face expression through video data and text expression through datasets of a persons chat history of any social media. To help automatically quantify emotional expression differences between patients with psychiatric disorders.

In depression analysis who is a pioneer in the affectibe computinh area, performed an experiment where he fused both computing area that is visual and vocal analysis. In that both computational area visual model is used to detect emotional expressions like happy, sad, depressed. If a person is sad the system will analyze the emotion and show the result as depressed if a person is happy then it show that person is not depressed. In this proposed system we are goin to use some emotions dataset in which there are thousands of emotions which will help us to detect each and every emotions. Asim Jan, Hongying Meng and Yona Falinie explored variations in the vocal prosody of the participants, and found moderate perdictibility of the depression scores based on the combination of F0 ans switching pauses. Yona Falinie analyzed both the manual and automatic facial expression dusring semisturctured clinical interviews of clinically depressed patients. They concluded that participants with high symtoms severity tend to expressmore emotions associated with contempt, and amile less. Asim Jan examined the effects caused by depression to younger patients of the both genders that is female and male.

The depression recognition subchallenge of AVEC2013 and AVEC2014; had proposed some good methods which achieved good results . From this,Asim Jan. were the winner of the depression subchallenge (DSC) for the AVEC2013 and AVEC2014 competitions. In 2013, they exploited the effects that reflected changes in coordination of vocal tract motion associated with MDD. Specifically, they investigated changes in correlation that occur at different time scales across dormant frequencies and also across channels of the delta-mel-cepstrum. In 2014, they looked at the change in motor control that can effect the mechanisms for controlling speech production and facial expression. They derived a multiscale correlation structure and timing feature from vocal data. Based on these two feature sets, they designed a novel Gaussian mixture model-based multivariate regression scheme. They referred this as a Gaussian staircase regression, that provided very good prediction on the standard Beck depression rating scale.

The above methods have achieved good performance. However, for the visual feature extraction, they used methods that only consider the texture, surface, and edge information. Recently, deep learning techniques have made significant progress on visual object recognition, using deep neural networks that simulate the humans vision-processing procedure that occurs in the mind. These neural networks can provide global visual features that describe the content of the facial expression. Recently, Chao et al. proposed using multitask learning based on audio and visual data. They used long short-term memory modules with features extracted from a pretrained CNN, where the CNN was trained on a small facial expression dataset FER2013 by Kaggle. The performance they achieved is better than most other competitors from the AVEC2014 competition, however, it is still far away from the state-of-the-art. A few drawbacks of their approach are the image size they adopted is very small, which would result in downsizing the AVEC images and reducing a significant amount of spatial information. This can have a negative impact as the expressions they wish to seek are very subtle, small and slow. They also reduce the color channels to grayscale, further removing useful information.

Improvement will be made in this proposed model from feature extraction though artificial intelligence, regression, with fusion, and build a complete system for automatic depression level prediction from both text and visual expression.

III. FRAMEWORKS

Human facial expression and textes in depression are theriotically different from thoes under normal mental scales. An attempt to find a solution for depression scale prediction is achiened by combining dynamic descriptions within naturalistics facial and text expression. A novel method is developed that comprehensively models the variations in visual and text cues, to automatically predict the BDI-II scale of depression. The proposed Framework is an exptention of the previous method of detection visual and vocal expression by replacing the hand crafted techniques with deep face representations as a base feature to the system.

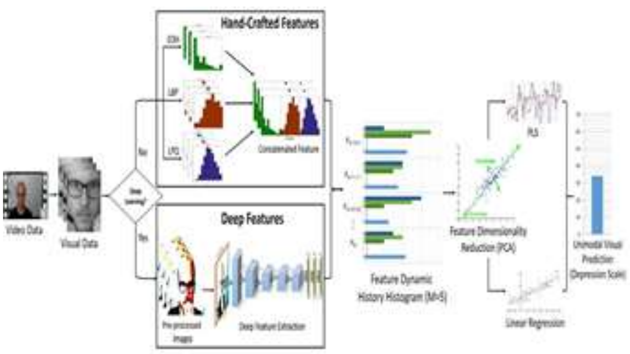

Fig. 1. System Architecture

B. Feature Exraction in Image Processing

Let us define the feature extraction first. Feature extraction means it involves reducing the number og resouces required to describe a large set of data. In our application there are some complex analysis in which we require feature extraction for complex analusis of face detection. feature extraction is a general term for methods of constructing combinations of the varianles to get around these problems while still describing the data with sufficient accuracy.

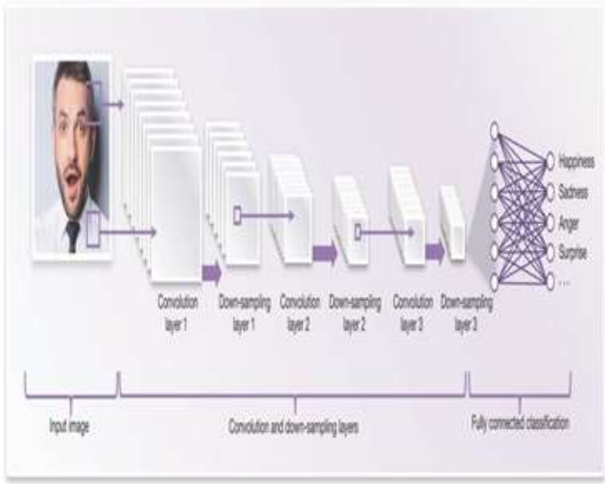

Fig. 2. Convolutional Neural Network

A. Convolutional Nueral Network

A convolutional neural network in artifical intelligence is a network which consist of an input and an output layers, as well as multiple hidden layers. The activation function is commonky a RELU layer, and is subsequently followed by additional convolution such as pooling layers, fully connected layers and normalizations layers reffered to as hidden layers because their inputs and outputs are masked by the activation function and final convolution.

In this proposed application when programming a CNN, the inputs is a tensor with shape(number of images)*(image width)*(image height)*(image depth). Then after passing through a convolutional layer, the images becomes abstracted to feature map, with shape(number of images)*(feature map width)*(feature map height)*(feature map channels). A convolutional layer within a neural network should have the following attributes.

In this system we have used constructed sets of application dependent features, typically built by an expert. One such process is called feature engineering. Alternativelty, general dimensionality reduction techniquies are used such as Isomap, Kernal PCA, Independent component analusis ans so on. In this model one very important area of the application is image processing, in which alogrithm is used to detect and isolate desired portions or shapes of the faces. Face has very complex features such as shape of the nose, color of the retina, shape of the cheeks and what not, for this type of complex analysis we are using feature extraction for image processing.

Many data analysis Software packages provide for feature extraction and dimension reduction. Common numerical programming environments such as MATLAB, SciLab, NumPy provides some of the simple feature extraction techniques. More specific algorithms are often available as publicly packages targetting specific software machine learning application that specialize in the feature extraction.

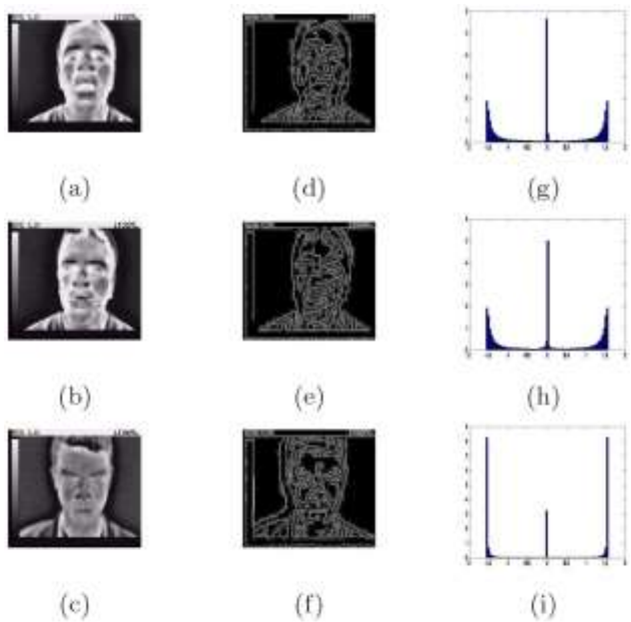

Fig. 3. Feature Extraction

C. Sentiment Analysis

Sentiment analysis is the process of using natural langurage processing, text analysis, and statistics to analyze customer sentiment. The best understanding of the sentiment of analysis is what people are saying, how they’ re saying it by the text expression. Suppose a person is chatting with another person or take an example of the hotel reviews in which their are some good reviews and more bad reviews than this sentiment analysis analyze the that which word that is good or bad is used more and result is shown in the form of ratings. In this model same way we are going to analyze the words such as positive, negative or neutral categories.

In addition to the defination problem, there are multipe layersof meaning in any human-generated sentence. People express opinions in complex ways; The onlly way to really understand these devices are through context knowing how a paragraph is started can strongly impact the sentiment of later internal sentences. Most of the current thinking in sentiment analysis happens in a categorical framework: sentiment is analyzed as belonging to a certain bucket, to a certain degree. For example, a given sentence may be 45 percent happy, 23 percent sad, 89 percent excited and 55 percent hopeful. these numbers dont add up to 100 percent they’ re individual indications of how ”X” a sentence’s sentiment is.

Create an instance of the SocialSentimentAnalysis algorithm. Call the algorithm on both of our sets of tweets and store the results. Convert the tweets into pandas dataframes. Show descriptive statistics.

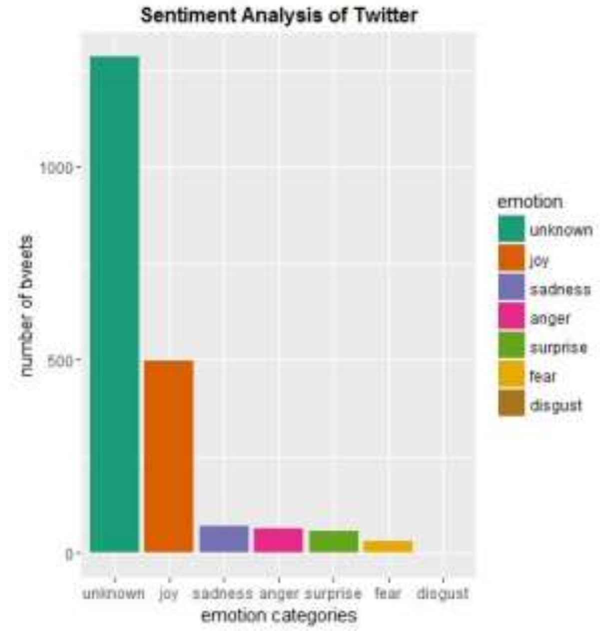

Fig. 4. Sentiment Analysis

IV. CONCLUSION

Depression is a serious issue prevailing nowadays. Many people, students and em- ployees of companies are going through this. So it has become necessary to detect it so that it could be cured at the early stages only.For this an artificial intelligent system was proposed for automatic depression scale prediction.The proposed system will provide accurate results for depression detection with the help of the depression scale. To improve the accuracy of automatic depression recognition from visual features , we proposed a new method based on deep learning , which we can use to overcome the difficulties caused by designing hand-crafted features for depression recognition.We conclude that ,powerful regression model can improve the accuracy of depression recognition.

There are three main contributions from this paper. First is the general framework that can be used for automatically predicting depression scales from facial and vocal expressions. The second contribution is the FDHH dynamic feature, that uses the idea of MHH on the deep learning image feature and hand-crafted feature space. The third one is the feature fusion of different descriptors from facial images. The overall results on the testing partition are better than the baseline results, and the previous state-of-the-art result set by Williamson et al. FDHH has proven it can work as a method to represent mathematical features, from deep features to common hand-crafted features, across a temporal domain. The proposed system has achieved remarkable performance on an application that has very subtle and slow changing facial expressions by focusing on the small changes of pattern within the deep/hand-crafted descriptors. In the case that a sample contains other parts of the body; has lengthier episodes; or reactions to stimuli, face detection and video segmentation can adapt the sample to be used in our system.

Further ideas can be investigated to improve the system performance. The performance may improve if additionalfacial expression images are added into the training process of the VGG-Face deep network. The raw data itself can be used to retrain a pretrained network, which can be trained as a regression model. For the vocal features, a combination ofdescriptors have been tested. However, other vocal descriptors should also be considered to be integrated in the system, or even adapting a separate deep network that can learn from the vocal data. Other fusion techniques can also be considered at feature and prediction level that would improve the performance further.

V. ACKNOWLEDGEMENT

It gives us great pleasure in presenting the preliminary project report on 'Depression Detection Using Arti cial Intelligence Through Visual Expressions'.With due respect and gratitude we would like to take this opportunity to thank our internal guide Prof.P.S.Gaikwad for giving us all the help and guidance we needed. We are really grateful for her kind support. She has always encouraged us and given us the motivation to move ahead. She has put in a lot of time and effort in this project along with us and given us a lot of con dence.We are also grateful to Dr.Zaware, Head of Computer Engineering Department, AISSMS' institute of information technology for his indispensable support.Also we wish to thank all theother people who have helped us in the successful completion of this project. We would also like to extend our sincere thanks to Principal Dr.P.B.Mane, for his dynamic and valuable guidance throughout the project and providing the necessar facilities that helped us to complete our dissertation work.We would like to thank my colleagues friends who have helped us directly or indirectly to complete this work.

VI. FUTURE SCOPE AND ENHANCEMENT

We perform the comparative evaluation on some of the widely used deep learning modules for the depression de- tection from to tweets on the user level. We perform are experiment on publically available dataset. Our exper- iments showed that our CNN based models perfom better then RNN based modules. Modules with optimized embeddings manage to maintain perfor- mance with the generalisation ability. Further ideas can be invetigated to improve system perfor- mance. The performancem may improve if additonal facial expression images are added into training processes of VGG- Face depp network. The new data itself can be used to retrain a pretrained network that can be trained as a regression model. Other Fusion techniques can also be considered as feature and prediction level that would improve the performance further.

REFERENCES

[1] H. Davies et al., "Facial expression to emotional stimuli in non-psychotic disorders: A systematic review and meta-analysis," Neurosci. Biobehav. Rev., vol. 64, pp. 252–271, May 2016.

[2] Depression and other common mental disorders: Global health estimates," World Health Org., Geneva, Switzerland, Tech. Rep. WHO/MSD/MER/2017.2, 2017

[3] D. D. Luxton, Artificial Intelligence in Behavioral and Mental Health Care. London, U.K.: Academic Press, 2015.

[4] M. Valstar et al., "AVEC 2016: Depression, mood, and emotion recogni- tion workshop and challenge," in Proc. 6th Int. Workshop Audio/Vis. Emotion Challenge (AVEC), Amsterdam, The Netherlands, 2016, pp.3–10.

[5] M. Nasir, A. Jati, P. G. Shivakumar, S. N. Chakravarthula, and P. Geor- giou, "Multimodal and multiresolution depression detection from speech and facial landmark features," in Proc. 6th Int. Workshop Audio/Vis. Emotion Challenge (AVEC), Amsterdam, The Netherlands, 2016, pp.43–50.

[6] L. Yammine, L. Frazier, N. S. Padhye, J. E. Sanner, and M. M. Burg, "Two-year prognosis after acute coronary syndrome in younger patients: Association with feeling depressed in the prior year, and BDIII score and Endothelin-1," J. Psychosom. Res., vol. 99, pp. 8–12, Aug. 2017.

[7] 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.

[8] Alghowinem, Sharifa, Roland Goecke, Jerey F. Cohn, MichaelWagner, 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.

[9] Harati, Sahar, Andrea Crowell, Helen Mayberg, Jun Kong, and Shamim Nemati. "Discriminating clinical phases of recovery from major depres- sive 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.

[10] L. Wen, X. Li, G. Guo, and Y. Zhu, "Automated depression diagnosis based on facial dynamic analysis and sparse coding," IEEE Trans. Inf. Forensics Security, vol. 10, no. 7, pp. 1432–1441, Jul. 2015.

[11] R.Weber, V. Barrielle, C. Soladie´, and R. Se´guier, "High-level geome- trybased features of video modality for emotion prediction," in Proc. 6th Int. Workshop Audio/Vis. Emotion Challenge (AVEC), Amsterdam, The Netherlands, 2016, pp. 51–58.

[12] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Las Vegas, NV, USA, 2016, pp. 770–778.

[13] H. Davies et al., "Facial expression to emotional stimuli in non-psychotic disorders: A systematic review and meta-analysis," Neurosci. Biobehav. Rev., vol. 64, pp. 252–271, May 2016

[14] A. Jan and H. Meng, "Automatic 3D facial expression recognition using geometric and textured feature fusion," in Proc. 11th IEEE Int. Conf. Workshops Autom. Face Gesture Recognit. (FG), vol. 5. May 2015, pp. 1–6.

[15] J. Han et al., "Learning computational models of video memorability from fMRI brain imaging," IEEE Trans. Cybern., vol. 45, no. 8, pp. 1692–1703, Aug. 2015.

