INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact factor: 6.078
(Volume 6, Issue 3)
Available online at: www.ijariit.com

Depression analysis using facial feature and local binary pattern

Shraddha Satam
shraddhasatam98@gmail.com
Usha Mittal Institute of Technology, Mumbai, Maharashtra

Shravani Satam
shravanisatam07@gmail.com
Usha Mittal Institute of Technology, Mumbai, Maharashtra

Bhavna Singh Parihar
bhavnap199@gmail.com
Usha Mittal Institute of Technology, Mumbai, Maharashtra

Kiran Dange
kiran.dange@umit.sndt.ac.in
Usha Mittal Institute of Technology, Mumbai, Maharashtra

ABSTRACT

Depression is one of the most serious problems faced by today's world. At the same time, it is the most difficult mental disorder to detect. Symptoms of depression varies from person to person. So, it's difficult for a doctor to diagnosis and treat it. And most of the methods to diagnosis of depression depend on interactive session with the patient or on the basis of behavioural observation. It is observed that most of the time patients don't give correct answers to the question asked or tries to behave normally in front of doctor, making it difficult for a doctor to diagnose depression. To overcome all these difficulties, we have developed automatic depression analysis method which uses dynamic facial landmark descriptor. This method gives the Depression score, on the bases of which we decide a person is depressed or not. As it does not involve any human interference such as clinical interview, doctor's interpretation, it gives accurate results.

Keywords— *Depression, Local Binary Pattern, Depression score, Histogram*

1. INTRODUCTION

Depression is considered one of the worst medical conditions. Symptoms of depression are hard to identify and may lead to serious mental disorders. According to World Health Organization (WHO), depression will become the 4th most serious mental health problems by 2020.In India around 200 million people suffer from depression, making India, 6TH most depressed country in the world. Due to lack of awareness about the symptoms related to depression and the social stigma associated with it, depression remains unrecognised and untreated. People don't even approach the doctor for the same. To overcome this, we have developed a technology which can detect depression using facial expression.

2. RELATED WORK

Various automatic depression recognition approaches such as Depression Recognition Sub-Challenge (DSC) of the Audio-Visual Emotion Challenge and Workshop (AVEC), AVEC2013, AVEC2014, AVEC2016, AVEC2017, automatic depression analysis using dynamic facial appearance descriptor and Dirichlet Process Fisher Vector Encoding have been developed [1]. This all methods have been developed using different facial landmark detection technique. All these approaches are based on the clinical interviews of depressed patients as well as non-depressed patients. In these present papers, we make use of the JAFEE database and system have been developed using OpenCV.JAFFE database contains images of 7 facial expressions posed by Japanese women. Based on only their facial expression, we have tried to decide the depression level.

2.1 Facial expressions related to depression

Facial expression plays an important role in depression detection. Many studies such as "Action Unit (AU)" were conducted to identify precise facial expressions that are related to depression. Here, the presence of AU14 related to emotion contempt and AU10 related to emotion disgust was also present along with AU12, where AU12 is related to smile. Study shows that AU14 is the most accurate to predict depression. Eye expression as well as head position play an important role in depression detection. Other facial features that plays an important role in depression detection are inner brow raiser outer brow raiser, brow lowered, upper lip raiser, check raiser, lip tightener, nose wrinkle, upper lip raiser, lip corner puller, dimpler, lip corner depressor.[1]

2.2 Data collection for Depression Analysis

Some of the data-sets available for depression detection include, Audio/Visual Emotion Challenge depression data-set (AVEC2013, AVEC2014, AVEC2016, AVEC2017), Black Dog Institute depression data-set, and University of Pittsburgh depression data-set. These all data-sets are based on audio-visual interviews taken.

(a) AVEC2013 Depression Database: In this dataset, interview is based on set of 21 questions asked to depressed and non-depressed person.The recorded behaviour contains sustained vowel phonation-speaking out loud while solving a task, counting numbers from 1to10,reading speech, singing, telling a story from the subject's own past experiences (the best event and a sad event from childhood), and telling an imagined story.[1]

(b) AVEC2014 Depression Database: The AVEC2014 corpus consists of recordings of human-computer interaction tasks. Participants read aloud an excerpt of the fable and respond to a number of questions.[1]

In both of the above data-set, Beck Depression Inventory–II(BDI–II) is used to detect depression. It takes into consideration age, gender, facial expression, past and present history of a person to determine depression level.[1]

(c) Black Dog Institute depression data-set, and University of Pittsburgh depression data-set: Data-set is formed while answering clinical interview. This interview questions are asked to depressed as well as non-depressed person. Videos of a person is recorded from the diagnosis of depression up to his recovery. Experimental study shows that there is a close relationship between facial features and vocal voice and it plays a significant role in depression prediction.

In this paper, we have used JAFEE database. It is based on facial expression. [13]

2.3 Feature Extraction Techniques

Features of video and audio data are extracted using method such as a Motion History Histogram (MHH), Local Binary Pattern (LBP), Medium Robust Local Binary Pattern in Three Orthogonal Plane (MRLBP-TOP) which represents the characteristics of minute changes that take place in face and vocal expressions of the depressed person.The main facial regions are selected and can be extracted by applying various methods such Adaboost algorithm, Haar Cascade Classifier, etc. Any minute changes in these features are clearly shown in this feature extraction technique. Out of all the methods mentioned above, MRLBP-TOP gives the accurate histogram of all the changes that are taking place in the facial region. The mean and standard deviation is extracted from the data that is collected from the histogram.[1]

3. PROPOSED SYSTEM

The proposed technology is a framework towards determining a clinical depression-specific score, called Depression score, based upon analysis of facial expressions. We propose a feature descriptor denoted as LBP, which can capture both the micro-structure and macro-structure of facial appearance and dynamics. Then aggregate the LBP over an image sequence and predict the depression score.

(a) Architectural Diagram
(b) Training Phase

3.1 Data Training

In this paper we have trained JAFFE dataset to find whether a person is depressed or not. We have made two different folders, one containing happy face and other folder containing mixed emotion since depression is a mixture of emotions. This help us to get more accurate result.[13]

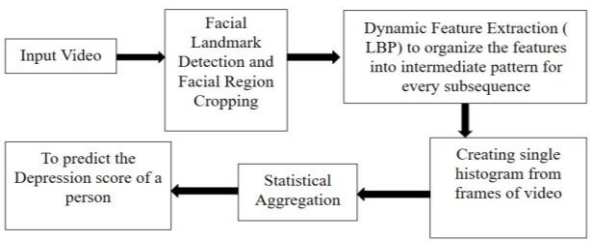

Fig. 1: Proposed system for diagnosis of depression

3.2 Converting into grayscale

As colour increases the complexity of the model, it is preferable to convert video into grayscale. Grayscale videos/images decreases the computational time, making it less time-consuming process. It also helps us identify important edges and other important features. It is also decrease coding complexity. So video is converted into grayscale.

Algorithm for converting input video into grayscale

- A video is taken as input
- It is converted into frames by frame rate defined by user.
- The frame extracted are then converted into grayscale and they are saved into output folder.

3.3 Face Detection and Facial Landmark Detection

From the grayscale video, face is detected using Haar Cascade Classifier. And after the detection of face, sixth-eight facial landmarks are detected. Facial landmark related to eyes, nose, mouth play an important role in depression detection.[1]

Algorithm for Face Detection

- Train data using Haar Cascade Classifier
- If camera is "ON" then give name to frame
- Convert image into grayscale and define the parameter for face detection
- Define colour and thickness
- Detect frame by comparing with trained data and create rectangle near face

Algorithm for Facial Landmark Detection

- Import libraries and train data
- Capture video and get frame from webcam
- Convert frame to grayscale and detect face region on frame
- Draw a rectangle near face
- Now extract 68 facial landmarks from face

3.4 Cropping of facial region

After the detection of sixty-eight facial landmarks, it is necessary to crop the facial region from the background. This is a very important step in depression detection as it helps us to get an accurate reading of any deviation taking part in the region of concrete without the interference of unwanted region. These reduces the error in depression detection.

3.5 Converting video into frames

Since we are using clinical video interviews for depression detection, it is necessary to convert videos into frames as it is impossible to plot histogram of a video. For these purposes it is important to define the number of frames formed per minute or second. In this paper we have set the number of frames formed per second is five.[1]

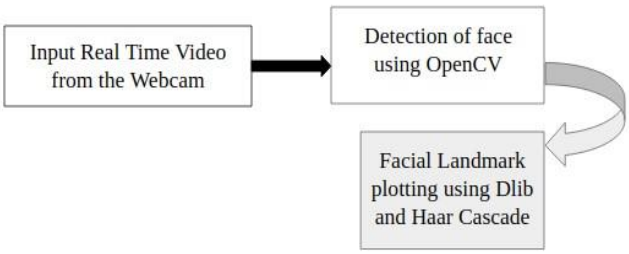

Fig. 2: Facial Landmark Detection from Input Video

3.6 Facial Feature Extraction

After cropping the face from background and converting video into frames, we require to plot a histogram of this various features. This can be done using Local Binary Pattern (LBP). LBP is used for describing two-dimensional surface textures, which in return help us to describe and plot a histogram of each feature in the frame. Here we have plot histogram of each frame of a video.[3]

$$LBP_{(P.R)} = \sum_{p=0}^{p-1} (g_p - g_c)2^p \quad (1)$$

$$s(x) = \begin{cases} 1, & if\ x \leq 0 \\ 0, & Otherwise \end{cases} \quad (2)$$

where P represent a number of sampling points on the circle of radius R; gp and gc are the pixel intensities of a pixel on the circle and the central pixel.

3.7 Combining histogram

Once the histogram is plotted for each frame of a video. All this histogram is combined and a single histogram.

All the above steps are done for practical example as well as for the training dataset.

(c) Testing Phase

In testing phase, histogram that we got from combining is compared with the histogram of trained data set. And by comparing those two histograms, we can predict whether a person is depressed or not.[8]

4. EXPERIMENTAL RESULTS

4.1 Experimental Requirement

Here we have used Ubuntu version 18.04 and Python version: 3.7. For face detection we have used OpenCV version 3.4.4 and for facial landmark detection, we have used dlib version 19.19.0 and numpy version 1.13.3. For implementation of histogram, we have used-Matplotlib version 3.1.2, Argparse version 0.3 and Imutlis version: 0.5.3.[9][10][11][12]

4.2 To predict the Depression score

Statistical Aggregation: Aggregate score of happy and sad person in training data set is 0.00325 and 0.0038 respectively. By comparing the practical values with this value, we can find out whether a person is depressed or not.

X	Happy	Depressior	Practical
12.5	0.02	0.022	0.016
25	0.006	0.006	0.005
37.5	0.007	0.009	0.007
50	0.002	0.003	0.003
62.5	0.007	0.008	0.008
75	0.002	0.002	0.002
87.5	0	0.001	0
100	0.001	0.002	0.002
112.5	0.002	0.002	0.003
125	0.006	0.006	0.008
137.5	0.001	0.002	0.002
150	0	0	0
162.5	0	0.001	0.001
175	0	0	0
187.5	0	0.001	0
200	0.001	0.001	0.001
212.5	0	0	0
225	0.001	0.001	0.003
237.5	0.002	0.002	0.005
250	0.007	0.007	0.013
Average	0.00325	0.0038	0.00395

Fig. 3. Aggregating all frames result

4.3 Results

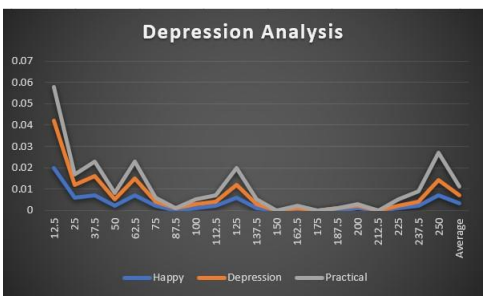

Fig. 4: Final Result

In the above example the aggregate score is 0.00395, which is above the depression score. So, the person is depressed.

5. CONCLUSION

Using JAFEE dataset and 68 facial landmarks we can get accurate results of whether a person is depressed or not. As we are not just concentrating on a few features but on 68 facial landmarks, it gives an idea of each feature in detail and reduces the error.

But this method too comes with some drawbacks such as it is increasing the computational speed as we require to plot histogram of each frame of a video and then combine it and create one histogram. At the same time, we require to do the same thing with the training dataset, which increase the complexity of the model. We even require large amount of dataset to get accurate results.

Mentioned above are some of the area which require improvement.

6. ACKNOWLEDGMENT

We would like to express our sincere gratitude to Prof.Dr. Nema Shikha, Head of Department (ENC) and our guide, Prof.Kiran Dange for her valuable guidance and continuous support during the project; her patience, motivation, enthusiasm, and immense knowledge. Her direction and mentoring helped us to work successfully on the project topic.

Our sincere gratitude to Dr. Sanjay Pawar, Principal for his valuable encouragement and insightful comments.

We would also like to thank to all the teaching and nonteaching staff for their valuable support. Last but not the least we would like to thank to our parents and friends.

7. REFERENCES

[1] L. He, D. Jiang and H. Sahli," Automatic Depression Analysis Using Dynamic Facial Appearance Descriptor and Dirichlet Process Fisher Encoding," in IEEE Transactions on Multimedia, vol. 21, no. 6, pp. 14761486, June 2019.

[2] K. Vikram and S. Padmavathi," Facial parts detection using Viola Jones algorithm," 2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, 2017, pp. 1-4.

[3] Guoying Zhao and M. Pietikainen," Local Binary Pattern Descriptors for Dynamic Texture Recognition," 18th International Conference on Pattern Recognition (ICPR'06), Hong Kong, 2006, pp. 211-214.

[4] "Ubiquitous Computing and Ambient Intelligence," Google Books. [Online]. Available: https://books.google.co.in/books?id=WtiGDQAAQBAJpg=PA343lpg= PA343dq=CHEHAR tool in face detectionf=false. [Accessed: 16-Oct2019].

[5] [2] R. Zhang, C. Mu, M. Xu, L. Xu and X. Xu," Facial Component Landmark Detection with Weakly-Supervised LR-CNN", IEEE Access, vol. 7, pp. 10263-10277, 2019. Available: 10.1109/access.2018.2890573.

[6] " Local binary patterns", En.Wikipedia.org, 2020. [Online]. Available: https://en.Wikipedia.org/wiki/Local$_b$inary$_p$atterns.[Accessed : 13 − Jan − 2020].

[7] " Look at the LBP histogram - AURlabCVsimulator", Aurlabcv simulator.readthedocs.io, 2020. [Online]. Available:

https://aurlabcvsimulator.readthedocs.io/en/latest/PlotLBPsHistogram/. [Accessed: 14- Jan- 2020].

[8] G. Zhao and M. Pietikainen," Dynamic Texture Recognition Using Local Binary Patterns with an Application to Facial Expressions", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, no. 6, pp. 915-928, 2007. Available: 10.1109/tpami.2007.1110 [Accessed 8 January 2020].

[9] Opencv.org. (2020). About. [online] Available at: https://opencv.org/about/ [Accessed 13 Jan. 2020].

[10] Matplotlib.org. (2020). Matplotlib: Python plotting Matplotlib 3.1.2 documentation. [online] Available at: https://matplotlib.org/ [Accessed 13 Jan. 2020].

[11] Docs.python.org. (2020). 16.4. argparse — Parser for command-line options, arguments and sub-commands — Python 3.4.10 documentation. [online] Available at: https://docs.python.org/3.4/library/argparse.html [Accessed 13 Jan. 2020].

[12] PyPI. (2020). imutils. [online] Available at: https://pypi.org/project/imutils/ [Accessed 13 Jan. 2020].

[13] ashishpatel26/Facial-Expression-Recognition-using-JAFFE", GitHub, 2020. [Online]. Available: https://github.com/ashishpatel26/Facial-ExpressionRecognition-using-JAFFE. [Accessed: 01-Mar- 2020].

