International Journal of Pure and Applied Mathematics
Volume 118 No. 7 2018, 455-463
ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
Special Issue

ijpam.eu

Extraction of Facial Features for Depression Detection among Students

Dr. D. Venkataraman
Department of Computer Science and Engineering,
Amrita School of Engineering, Coimbatore,
Amrita Vishwa Vidyapeetham,
India.
d_venkat@cb.amrita.edu

Namboodiri Sandhya Parameswaran
Department of Computer Science and Engineering,
Amrita School of Engineering, Coimbatore,
Amrita Vishwa Vidyapeetham,
India.
cb.en.p2cvi16004@cb.students.amrita.edu

Abstract— Psychological health of college students prove a vital role on their overall academic performance. Neglecting this can result in several problems such as stress, anxiety, depression etc. These problems need to be detected and controlled at the initial stages itself for the better mental health of the student. Detecting depression in a vast no of college students is challenging task. Most of the students are totally unaware that they may be having depression. If at all they are aware of it, some students conceal their depression from everyone. So an automated system is required that will pick out the students who are dealing with depression. A system has been proposed here which captures frontal face videos of college students, extracts the facial features from each frame and analyses these facial features to detect signs of depression in them. This system will be trained with of frontal face images of happy, contempt and disgust faces. The presence of these features in the video frames will be analyzed to predict depression in the students.

Keywords— Keywords—Image processing, Feature Extraction, Facial Features, Depression Detection.

I. INTRODUCTION

Students are said to be in sate of depression when they find themselves feeling sad or hopeless almost every day for two or more weeks in a row.[10] Such students then start retrieving from their social life. They stop doing some of the usual activities they usually like doing. Depression in college students can also be due to the total lifestyle change when they enter college. Going off to college is a striking transition for students who are emotionally dependent on their parents. Some students find it very stressful to handle the intense academic pressure. This stress in the long run leads them to a state of depression.

A study has been undertaken to associate internet usage of college students with symptoms of depression [10]. This study revealed that students with depression accessed the internet more often as compared to normal students. Due to the rise of social networking sites, teenagers are constantly in a state questioning themselves whether they are social accepted by their peers. Social media sites create a sense of wrong belief in students that their peers are leading a more successful and happier life compared to theirs. Spending more time on social media may also make students victims of cyber-bullying which can cause depression in them. In addition to this certain online games like the "The Blue Whale game" have been targeting students who are in an emotionally vulnerable state or in a state of disturbed mental condition.

In many situations students 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 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.

II. RELATED WORK

A. Facial Expressions Related to Depression

Many studies have been conduced to identify the precise facial expressions that are related to depression. A study has been conducted for finding out Action Units (AU) related to different emotions exhibited by depressed patients [1]. The presence of AU12 which is associated with emotion smile was low in highly depressed patients. The presence of AU14 related to emotion contempt and AU10 related to emotion disgust was also present along with AU12. The video data for this study was collected through clinical interviews of depressed patients as well as non-depressed patients. The results showed that AU14 related to emotion contempt proved most accurate for depression detection.

Features related to eye movement to understand the eye activity of the depressed and features related to head pose movement to understand the head movement behavior of the depressed has been done in [2]. The classification of the features related to eye activity showed higher significance in detecting severe depression. Detection of depression from facial features can be done by measuring 'Multi-Scale Entropy' (MSE) on the patient interview video. [4] MSE helps to find out the variations that occur across a single pixel in the video. The entropy levels of highly expressive, non-depressed patients were high. The entropy level was low for depressed patients who were less expressive of their emotions.

The students suffering from depression would show less attentiveness in classrooms. If the students' emotions are mapped to the activities done in classroom, their emotional state can be found out whether they are depressed or not, and based on this the teacher can help the student by paying more attention to that particular student. [11] If different faces in the same scene show the same positive or negative sentiment, it would help to understand the whole situation of the scene, whether subjects in the scene are happy or whether something wrong is happening in the scene [12].

B. Data collection for Depression Analysis

Some of the datasets available for depression detection include – BlackDog (Black Dog Institute depression dataset), Pitt ('University of Pittsburgh depression dataset), AVEC (Audio/Visual Emotion Challenge depression dataset) etc. [2] Analysis of the combination of these three datasets has been done for detection of depression features. The eye activity modality showed better performance for both individual and combined datasets. This indicated that more the variability in much better will be the testing results.

Another study presented a technique which uses analysis of facial geometry along with analysis of speech for depression detection [3]. This work says that the expressions associated with depression are found to be in lower frequencies in smaller duration videos. Therefore longer time videos need to be captured for effective depression detection. Datasets are also created by capturing videos of patients while answering clinical interviews. Interviews recorded were for both for depressed patients as well as non-depressed patients. Videos are also recorded from the diagnosis of depression till the patient has improved. [1][4]. Studies showed that there is a significant relation between facial features and vocal behavior of the depressed [5].

In certain studies, patients were given wearable devises to monitor their physical health, emotional behavior and social interaction for identifying depression [6]. Some researchers have collected datasets by showing individuals film-strips to capture the facial expressions of subjects watching them. Data is also collected by giving a task of recognizing negative and positive emotions from different facial images [7]. Rather than analyzing a video for depression detection frame by frame, better results have been got for detection of depression when the video is considered as a whole. [8] For this the patient's face region is first initialized manually. Then KLT (Kanade-Tomasi-Lucas) tracker is used to track the face throughout the video. The KLT tracker extracts curvature information from an image, i.e. for a sad expression the corners of the mouth would be angled down. Video based approach showed more accuracy as it generalizes the face region more accurately and so the minute movements within the face region are also considered for depression detection.

C. Feature Extraction Techniques

Features of video and audio data are extracted from the video using a 'Motion History Histogram (MHH)' represents which represents the characteristics of minute changes that happen in face and vocal expressions of the depressed [9]. A method of modeling the face recognition procedure with the help of Gabor Wavelet has also been proposed [13]. Here the work is done for recognizing the faces invariant to Pose and Orientation. Gabor filters are used for facial feature extraction. Gabor filters have got certain basic invariance properties and thus they can be used for face recognition invariant to orientation and pose of the face. As features, the mean and standard deviation is extracted of 40 Gabor filters [15]. These are then classified with the help of SVM classifier. This system uses the RBF kernel and claims to have a good recognition rate and also mentions that it outperforms other face recognition techniques.

The main facial regions selected are eye, nose and mouth regions that can be extracted by applying Haar feature based Adaboost algorithm. For a large database, this method claims to reduce the face recognition processing time. Facial Action Units are also being detected, where combination of different facial action units can different complex emotions for expression analysis [14] The Viola-Jones face detection algorithm is the most robust face detection algorithm available at present because the algorithm works with very high True-Positive rates and under different illumination conditions. It is also very efficient for different face part detection like eyes, nose mouth etc.

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 counselling 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 SVM 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 modelled in the following way.

A. Architectural Diagram:

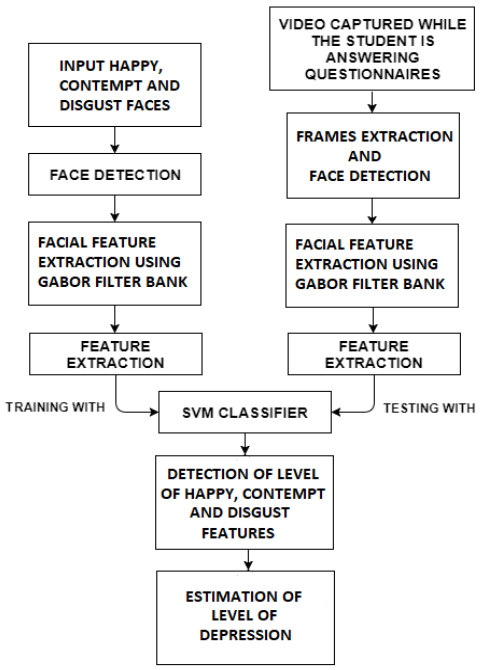


Fig. 1: Architectural diagram for the proposed 'Depression Detection' system.

B. Description of Proposed Architectural Diagram

The Algorithm depicted in the architectural diagram works in three phases as below:

1) Training Phase

a) Input Happy Faces

The basic idea here is to find out the amount of negativity in the input videos. As mentioned in previous work, the depression related features exhibited low presence of smile (AU12). And if smile was present, they also showed the presence of contempt (AU14) and disgust (AU10) along with smile (AU12). For this we can find the amount of presence of happy, contempt and disgust features within the videos, i.e. we are checking whether the student is smiling more often or less smiling with disgust or contempt or not at all smiling. For this

we need to find out the happy, contempt and disgust features from different faces. The input to the training phase is thus a dataset of happy, contempt and disgust faces.

b) Face Detection

For feature extraction first the face has to be detected. Face detection can be done using the Viola Jones face detection algorithm. The Viola-Jones face detection algorithm is the most robust face detection algorithm available at present because the algorithm works with very high True-Positive rates and under different illumination conditions. It is also very efficient for different face part detection like eyes, nose mouth etc. [22][23] This algorithm makes use of certain features called Haar features which are convolved throughout the images in different pixel sizes. More than 160,000+ features get extracted which is a very big number to work with. So the Adaboost classifier is used which narrows the features down to several thousand features (up to 25000) that are most useful. It eliminates all the irrelevant features and keeps only the relevant features required for face detection.

c) Facial Feature Extraction

Features of faces can be found out by applying Gabor filters. A Gabor filter bank of 40 filters, with 5 different scales and 8 different orientations can be used to extract the happy features. The orientations include: 0, 23, 45, 68, 90, 113, 135, 158, the scale values can vary starting from 4 till the required level which can be found out based on the requirement.

d) Training with SVM

Once the features are extracted, the features can be classified using Support Vector Machine classifier. Support Vector machine is a supervised machine learning algorithm that is mainly used for classification purpose as well as for regression purpose. SVM classification deals with the formation of a line or a hyper-plane that effectively classifies different classes. The SVM classifier can be trained with the happy, contempt and disgust features. In the testing phase, the features of the faces of the test dataset can be extracted and tested with this trained SVM to recognize a happy, contempt or a disgust face.

2) Testing Phase

a) Dataset Creation

Here the video of the student is captured while the student is answering a depression detection questionnaire as well as a non-stress related questionnaire. The need for two different types of questionnaire is that, the student may not answer the depression detection questionnaire correctly and put forth a different facial expression in order to conceal his state of mind. So another non-stress related questionnaire can be provided to record the subject's actual normal facial expression.

b) Extraction of Frames and Face Detection

Once the video is captured it is converted into frames and stored in a different folder. A simple one minute video when converted into frames creates around 1000 to 2000 images of the same person. Thus for two different videos (captured while the student is answering two different types of questionnaires) we can create a large set of input images for testing. Then the face is detected using the Viola Jones face detection algorithm. These faces are saved to create a dataset of test faces.

c) Test Features Extraction

Once the face is detected, the facial features can be detected for every frame of the video using Gabor filters as in the training phase. Then the feature set of the test face images can be formed. The system can test these features for the presence of happy, contempt or disgust facial features. Based on the level of the presence of these features in the entire video, the image can be predicted as 'Depressed or Non-Depressed' face.

3) Classification Phase

The SVM classifier is trained with happy, contempt and disgust features extracted from the input happy face dataset in the training phase. The features extracted from each frame of the video dataset can be tested with this trained SVM to check if these features are present in each frame. The level of these features can give an indication about the level of depression as below:

- If the happy features are present in less no of frames, it indicates that the student is not happy. The student is not expressive about his/her emotional state. If the video contains contempt or disgust features also, it may clearly indicate that the person is having a negative state of mind. This indicates the total negativity of the whole face. Thus the student may be classified as having high depression
- If the happy features are in moderate level and the video does not indicate any contempt or disgust features are high, the student may be classified as having mild depression
- If the presence of happy features is high and the presence of contempt and disgust features are low, the measure of depression can be considered as low.

The estimation of the depression level of a face can be tabulated as below:

TABLE 1: DEPRESSION LEVEL ESTIMATION TABLE

Happy features	**Contempt, Disgust features**	**Depression Level**
High	High	High Depression (Concealed Depression)
High	Low	Not depressed (Happy)
Moderate	High	Mildly depressed
Moderate	Low	Sad
Low	High	High depression

Low	Low	Sad

Once the facial expression is identified as that of depression, we can find out the level of depression, and appropriate counseling can be suggested.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

1) Training Phase:

a) Input Dataset

The input dataset has been created by collecting happy, contempt and disgust face images of people from the JAFFE database and from the internet. There are 30 faces of each emotion and so there are overall 90 images in the input dataset. The dataset of happy faces is created by collecting all the happy faces from the JAFFE database as below.

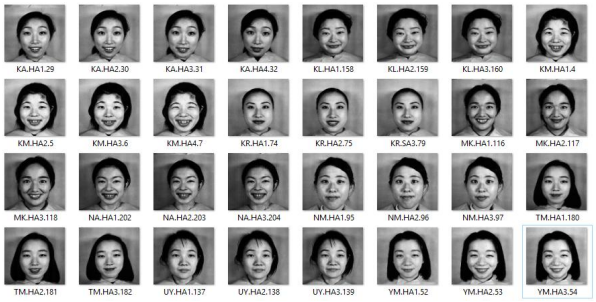


Fig. 13: Happy Faces from JAFFE database.

The contempt and the disgust face images are collected from the internet as below:

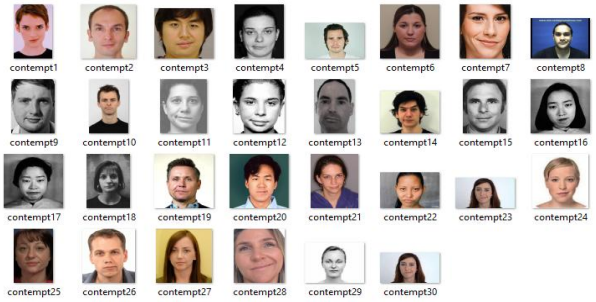


Fig. 14: Contempt Face images.

Fig. 14: Disgust Face images.

b) Face Detection and Facial Feature Extraction

For all the faces in the input dataset, the face detection is done using Viola Jones face detection algorithm. Then for each face, the happy facial expressions of the face are extracted using Gabor filter bank of filters. A bank of 40 Gabor filters are formed with different scales and orientations as below:

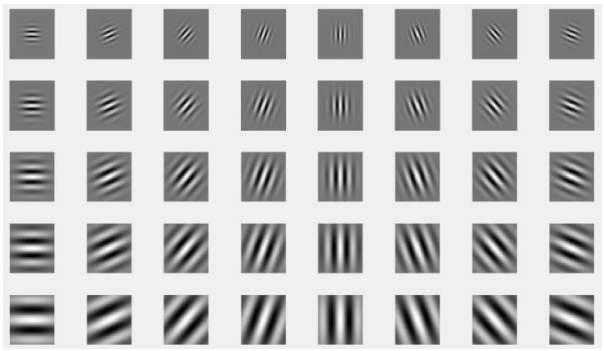


Fig. 14: Gabor Filter bank

Each filter is a combination of scale and orientation. The 5 different scales of Gabor filters are used here ranging from 4, 4.1, 4.2, 4.3, and 4.4. The orientation used here is 0, 23, 45, 68, 90, 113, 135 and 158. Combining them we get a total of 5x8 = 40 Gabor filters.

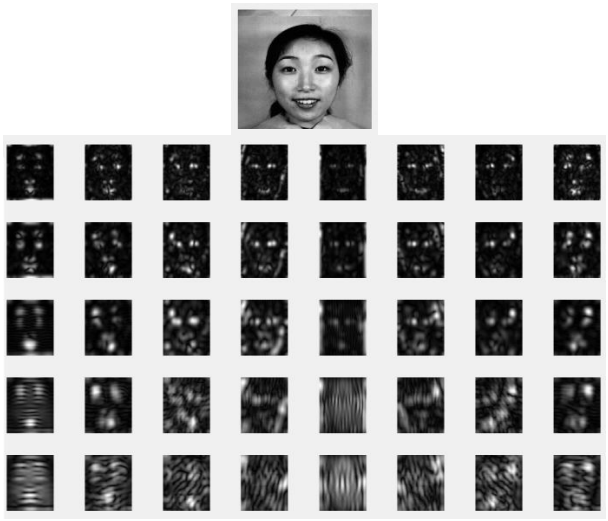


Fig. 15: Magnitude of images formed when filter bank applied is on happy face.

When these Gabor filters are applied to an image in creates the magnitude images as in figure 15. Each image shows the magnitude output of the Gabor filter applied to the image in a particular scale and orientation. Therefore overall we get 40 magnitude images of the same input image.

Fig. 16: Happy facial features extracted using Gabor filter bank.

Combining all the magnitude images we get above image. This image shows the happy features of the input image. Based on these features we can classify the test images.

c) Feature Set Creation

For every image feature vector is created as in figure 17. The feature set is created using the mean and standard deviation of the Gabor filtered image. The Gabor filtered image got is down-sampled to 160x160 dimension image for easier computation. The mean of the entire magnitude image is found out and then subtracted. It is then divided by the image's standard deviation to get a 64000x1 column vector. This is the feature vector of the sample image.

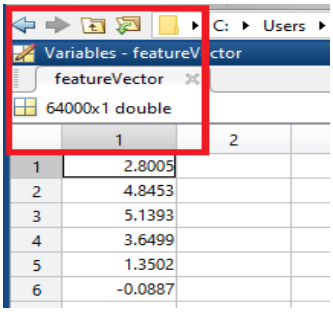


Fig. 17: Feature vector created for a single image.

The feature vectors for all the 30 images of the input happy faces are found out and concatenated to form a Feature-Set as in figure 18(a). Then PCA is applied for dimensionality reduction to form a 30x30 feature matrix and the class value '1' is given to the set of happy images as in figure 18(b).

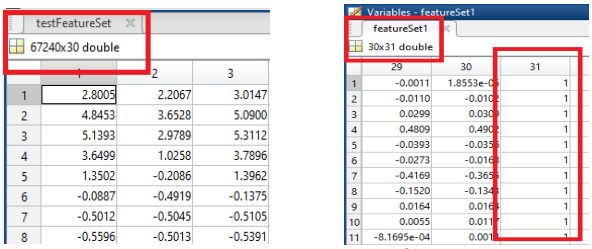


Fig. 18 (a): Feature Set created for all happy face images in dataset;
(b) PCA applied feature set with class value '1' for happy face images

Similarly, feature set for all the 90 images are formed and the class values '2' and '3' are given for contempt and disgust face images respectively.

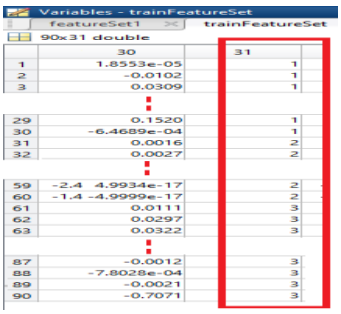


Fig. 19: Class values given to images happy-1, contempt-2 and disgust-3

d) *Creation of dataset for testing:*

A video is capture while the person is answering a depression related questionnaire. A simple web camera can be used here to capture the video. The image frames then are extracted from the video. A simple 1 minute video can be extracted into almost around 1000 + image frames. These

image frames are saved together in a folder to form the testing image database.

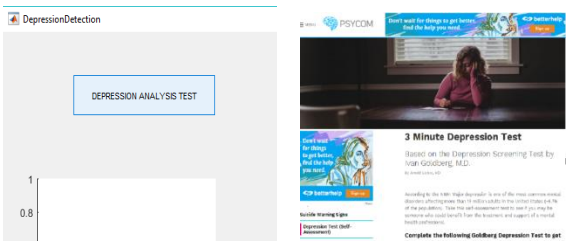


Fig. 20: Student is provided a link for a depression detection questionnaire.

Fig. 21: Video is captured and saved while answering questionnaires.

Once the video is captured, it is converted into frames. Each frame is saved to form the input image database for the testing phase.

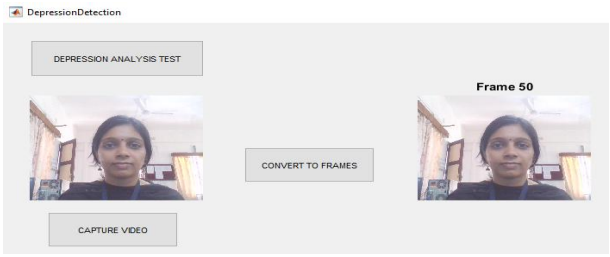


Fig. 22: Video is converted into frames

Then the face is extracted from each image using the Viola Jones face detection algorithm. The Viola Jones face detection algorithm is the most robust face detection algorithm available in the current scenario.

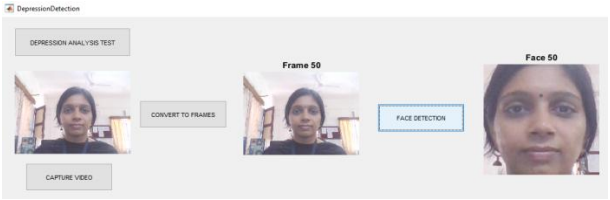
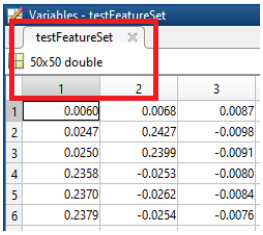


Fig. 23: From each frame face is detected and cropped.

The Cropped faces are saved in a different folder. These faces together form the dataset of faces for testing phase. The Gabor features of these faces are extracted and concatenated to form a 'test feature-set' for testing phase.

First 50 frames are considered here for test images. The Gabor features for these 50 images is found out in the same way as in the training phase and the test feature set is created as below:



	1	2	3
1	0.0060	0.0068	0.0087
2	0.0247	0.2427	-0.0098
3	0.0250	0.2399	-0.0091
4	0.2358	-0.0253	-0.0080
5	0.2370	-0.0262	-0.0084
6	0.2379	-0.0254	-0.0076

Fig. 24: Test feature set for 50 test images

The test feature set will be tested with the SVM classifier. The classification of the facial features will be done on the basis of the classification depicted in Table 1. If the happy features are present in less no of frames and the contempt and disgust features are also present, the student may be classified as having high depression. If the happy features are in moderate level and the video has high level of contempt or disgust features, the student may be classified as having mild depression. If the presence of happy features is high and the presence of contempt and disgust features are low, the student may be classified as not depressed.

V. CONCLUSION AND FUTURE WORK

A detailed study has been done on facial expressions related to depression, and the methods of extracting the same. A system for depression detection for college students has been proposed. A dataset of happy faces is created by collecting the happy faces from the JAFFE database. Face detection for each face is done using Viola Jones face detection algorithm. Then the facial features for each face are found out using a Gabor filter bank of 40 filters. The feature vector for each image is formed concatenated to form a feature-set for training.

For the testing phase a GUI is created for capturing video of student while the student is answering depression detection questionnaires. The video is converted into frames and for each frame the face is detected, cropped and saved to create a dataset of test faces.

Future work includes training of the feature-set with SVM classifier with RBF kernel. Then while the test image is given the system can predict if the input face is depressed or not by finding out the measure of happy features that may be present in the video frames.

The absence of happy features will give the amount of negativity in the video. Then identification of the level of depression of each video can be done into three levels – high, moderate and mild, based on the depression level estimation depicted in table 1. Depending on the overall negativity of the video the required counseling can be recommended.

ACKNOWLEDGMENT

The authors would like to extend the heartfelt gratitude to the faculty-in-charge of Amrita-Cognizant Innovation Lab, Department of Computer science and Engineering, Amrita school of Engineering, Coimbatore for the support extended in carrying out this work.

REFERENCES

[1] Girard, Jeffrey M., Jeffrey F. Cohn, Mohammad H. Mahoor, Seyedmohammad 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 Shamim Nemati. "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, Margara Tejera 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 FabriceMeriaudeau. "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., RajindaSenaratne, 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.

[10] Katikalapudi, Raghavendra, Sriram Chellappan, Frances Montgomery, Donald Wunsch, and Karl Lutzen. "Associating internet usage with depressive behavior among college students." IEEE Technology and Society Magazine 31, no. 4 : 73-80. 2012

[11] Sahla, K. S., and T. Senthil Kumar. "Classroom Teaching Assessment Based on Student Emotions." In The International Symposium on Intelligent Systems Technologies and Applications, pp. 475-486. Springer International Publishing, 2016.

[12] Athira, S., R. Manjusha, and Latha Parameswaran. "Scene Understanding in Images." In The International Symposium on Intelligent Systems Technologies and Applications, pp. 261-271. Springer International Publishing, 2016.

[13] Karthika, R., and LathaParameswaran. "Study of Gabor wavelet for face recognition invariant to pose and orientation." In Proceedings of the International Conference on Soft Computing Systems, Springer, New Delhi, 2016.

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

[15] Haghighat, Mohammad, Saman Zonouz, and Mohamed Abdel-Mottaleb. "CloudID: Trustworthy cloud-based and cross-enterprise biometric identification." Expert Systems with Applications 42, Vol no. 21, pp. 7905-7916, 2015.