International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 9, Issue 11, pp 129-133, November-2025

https://doi.org/10.47001/IRJIET/2025.911016

MindMirror - Mirror that Reads Emotions

Sudhanva Kale

Student, Electronics and Telecommunication, College of Engineering, Pune, Maharashtra, India

***Abstract* - Automatic depression assessment based on visual and vocal cues is a rapidly growing research domain. The present exhaustive review of existing approaches as reported in over sixty publications during the last ten years focuses on image processing and machine learning algorithms. Visual manifestations of depression, various procedures used for data collection, and existing datasets are summarized. The review outlines methods and algorithms for visual feature extraction, dimensionality reduction, decision methods for classification and regression approaches, as well as different fusion strategies.**

A quantitative meta-analysis of reported results, relying on performance metrics robust to chance, is included, identifying general trends and key unresolved issues to be considered in future studies of automatic depression assessment utilizing visual and vocal cues alone or in combination with cues. The proposed work also carried out to predict the depression level according to current input of videos using deep learning as well as NLP.

Keywords: Image Processing, Machine Learning, Classification Rule, Convolution Neural Networks, NLP etc.

I. INTRODUCTION

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 users/peoples fall into a state of mind where they start having suicidal tendencies. In some cases, peoples 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 people. 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, NLP for speech identification, 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/peoples 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.

Automatic detection of depression has attracted increasing attention from researchers in psychology, computer science, linguistics, and related disciplines. As a result, promising depression detection systems have been reported. In this proposed work these efforts by presenting the first cross-modal review of depression detection systems and discuss best practices and most promising approaches to this task.

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:

- Stress
- Depression

An innovative solution to monitor and detect potential users with emotional disorders, according to the classification of sentences with depressed or stressed content.

A machine learning is used for representation at the character level and for the recognition of the extent of the disturbance.

II. LITERATURE SURVEY

1. Artificial Intelligent system for automatic depression level analysis through visual and vocal expression.

Author: Asim Jan, Hongying Meng, Yona Falinie Binti A. Gaus,and Fan Zhang.

Description: Automatic depression assessment based on visual and vocal cues is a rapidly growing research domain. The present exhaustive review of existing approaches as reported in over sixty publications during the last ten years focuses on

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 9, Issue 11, pp 129-133, November-2025

https://doi.org/10.47001/IRJIET/2025.911016

image processing and machine learning algorithms. Visual manifestations of depression, various procedures used for data collection, and existing datasets are summarized.

The review outlines methods and algorithms for visual feature extraction, dimensionality reduction, decision methods for classification and regression approaches, as well as different fusion strategies. A quantitative meta-analysis of reported results, relying on performance metrics robust to chance, is included, identifying general trends and key unresolved issues to be considered in future studies of automatic depression assessment utilizing visual and vocal cues alone or in combination with cues. The proposed work also carried out to predict the depression level according to current input of videos using deep learning as well as NLP.

In this paper FDHH generated from patterns but in our system we avoid Histogram.

2. Objective methods for reliable detection of concealed depression.

Author: Cynthia Solomon, Michel F. Valstar, Richard K. Morriss & john crowe.

Description: This work aims to not only determine audio features that differ between healthy and depressed people, but also to investigate how they change when people with depression try to conceal their true emotions. Based on a set of optimized features, our goal was automatic depression recognition which will still be able to correctly classify a person as depressed even if they are trying to hide their depression. Healthy individuals who alter their behavior to appear depressed were not of interest in this study.

We presented our results of a study that looked into the automatic detection of depression using audio features in a human–computer interaction setting. This paper used for voice extraction but we are not using its method of determining normal and concealed behavior. We directly focus on depression level.

3. Approaches to understanding and addressing treatment - resistant depression: A scoping review.

Author: E. Jenkins and E. M. Goldner

Description: The objective of this study is to synthesize extant literature on approaches currently being applied to understand and address this condition. It is hoped that the findings can be used to inform practitioners and guide future research. A scoping review of the scientific literature was conducted with findings categorized and charted by underlying research paradigm. Currently, the vast majority of research stems from a biological paradigm (81%). Research on treatment-resistant depression would benefit from a broadened field of study. Given that multiple etiological mechanisms likely contribute to treatment-resistant depression and current efforts at prevention and treatment have substantial room for improvement, an expanded research agenda could more effectively address this significant public health issue.

This scoping review highlights the need to expand the scope of research being conducted in order to decrease the substantial burden associated with TRD and improve health outcomes for those experiencing this debilitating illness.

4. Facial expression to emotional stimuli in non-psychotic disorders: A systematic review and meta-analysis.

Author: Cynthia Solomon, Michel F. Valstar, Richard K. Morriss & john crowe.

Description: This article provides a systematic review and meta-analysis of the literature on automatic emotional facial expression in people with non-psychotic disorders compared to healthy comparison groups. Studies in the review used an emotionally salient visual induction method, and reported on automatic facial expression in response to congruent stimuli. In depression, decreases in facial expression are mainly evident for positive affect. In eating disorders, a meta-analysis showed decreased facial expressivity in response to positive and negative stimuli. Studies in autism partially support generally decreased facial expressivity in this group. In conclusion, this review has shown that facial expression of emotion is altered in people with mental health problems with broad similarities across certain clinical groups. This paper focuses on non-psychotic disorder but our system analyses the psychotic depression.

III. SYSTEM DESIGN

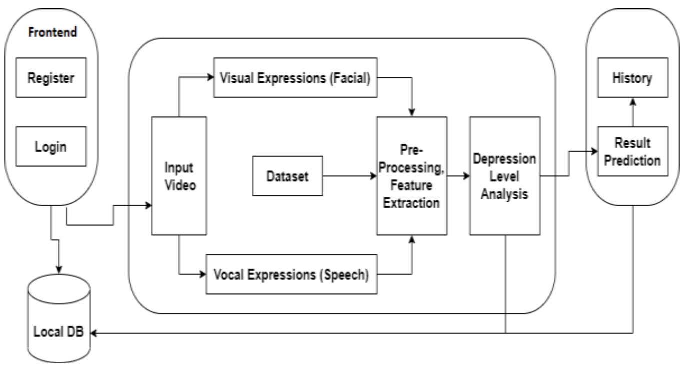

Figure 1: System Architecture

This is our system architecture diagram, in which user will register and the details get store in the local database, the user will login and provide a video, which is then separated in two parts a visual and audio.

In image processing CNN will be used for detecting face and to find emotions on face like happy, sad, angry, fear, neutral, disgust. In this process, images are taken from video and different frames are generated.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 9, Issue 11, pp 129-133, November-2025

https://doi.org/10.47001/IRJIET/2025.911016

In audio extraction, when we upload a video then audio is separated from video and wave generation takes place. For speech recognition we use bag of words logic, we have some positive word dataset as well as some negative word dataset, with the help of this we can find the probability of how many words person speaks positive or negative, at last we apply Naive Bayes algorithm for final result. After getting final result the user has a facility to see his report.

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.

1. System Framework:

In this framework we propose a novel hybrid model - a factor graph model combined with Convolution Neural Network to leverage contents and social interaction information for stress detection. Experimental results show that the proposed model can improve the detection performance by 6-9% in F1-score. By further analyzing the social interaction data, we also discover several intriguing phenomena, i.e. the number of social structures of sparse connections (i.e. with no delta connections) of stressed users is around 14% higher than that of non-stressed users, indicating that the social structure of stressed users' friends tends to be less connected and less complicated than that of non-stressed users.

2. Social Interactions:

We analyze the correlation of users' stress states and their social interactions on the networks, and address the problem from the standpoints of: (1) social interaction content, by investigating the content differences between stressed and non-stressed users' social interactions; and (2) social interaction structure, by investigating the structure differences in terms of structural diversity, social influence, and strong/weak tie. Our investigation unveils some intriguing social phenomena. For example, we find that the number of social structures of sparse connection (i.e. with no delta connections4) of stressed users is around 14% higher than that of non-stressed users, indicating that the social structure of stressed users' friends tends to be less connected and complicated, compared to that of non-stressed users.

3. Attributes categorization:

We first define two sets of attributes to measure the differences of the stressed and non-stressed users on social media platforms: namely image and speech-level attributes from a user's single user and user level attributes summarized from a user's weekly activities.

IV. ALGORITHMS USED

1. CNN Algorithm:

CNN is one of the main categories to do image recognition, image classification. Object detection, face recognition, emotion recognition etc., are some of the areas where CNN are widely used. CNN image classification takes an input image, process it and classify it under certain categories (happy, sad, angry, fear, neutral, disgust). CNN is a neural network that has one or more convolutional layers.

- Step 1: Dataset containing images along with reference emotions is fed into the system. The name of dataset is Face Emotion Recognition (FER) which is an open – source data set that was made publicly available on a Kaggle.
- Step 2: Now import the required libraries and build the model
- Step 3: The convolutional neural network is used which extracts image features f pixel by pixel
- Step 4: Matrix factorization is performed on the extracted pixels. The matrix is of m x n.
- Step 5: Max pooling is performed on this matrix where maximum value is selected and again fixed into matrix.
- Step 6: Normalization is performed where every negative value is converted to zero.
- Step 7: To convert values to zero rectified linear units are used where each value is filtered and negative value is set to zero.
- Step 8: The hidden layers take the input values from the visible layers and assign the weights after calculating maximum probability.

2. Text Mining:

Text Mining is the process of deriving meaningful information from natural language text. As Text Mining refers to the process of deriving high quality information from the text. The overall goal is, essentially to turn text into data for analysis, via application of Natural Language Processing.

3. Natural Language Processing (NLP):

Natural language processing (NLP) is a field of artificial intelligence in which computers analyze, understand and derive meaning from human language in a smart and useful way. By utilizing NLP, we can organize and structure knowledge to perform tasks such as automatic summarization, translation, named entity recognition, relationship extraction, sentiment analysis, speech recognition, and topic

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 9, Issue 11, pp 129-133, November-2025

https://doi.org/10.47001/IRJIET/2025.911016

segmentation. NLP primarily acts as an important aspect called as speech reorganization in which system analyze primary source of audio data in the form of spoken words. In NLP, syntactic analysis is used to assess how the natural language aligns with the grammatical rules. Here are some syntax techniques that can be used:

- **Tokenization:** Tokenization is an essential task in natural language processing used to break up a string of words into semantically useful units called tokens. Generally, word tokens are separated by blank spaces, and sentence tokens by stops.
- **Part-of-speech tagging:** It involves identifying the part of speech for every word. It signifies the word is noun, pronoun, adjective, verb, adverb, preposition or conjunction.
- **Bag of Words:** It splits each string into words and listing it into vocabulary and convert every word of data into its root word.

4. Naive Bayes Classifier:

It is probability based algorithm mostly used in text classification. Naive Bayes model is easy to build and particularly useful for very large data sets. Along with simplicity, Naive Bayes is known to outperform even highly sophisticated classification methods. Naive Bayes algorithm observes each feature independently even if they are related.

V. IMPLEMENTATION DETAILS

Emotion and Face Detection:

Facial expression for emotion detection has always been an easy task for humans, but achieving the same task with a computer algorithm is quite challenging. The current approaches primarily focus on facial investigation keeping background intact and hence built up a lot of unnecessary and misleading features that confuse CNN training process. The current manuscript focuses on five essential facial expression classes reported, which are displeasure/anger, sad/unhappy, smiling/happy, feared, and surprised/astonished. This is divided into 3 parts:

Facial Detection — Ability to detect the location of face in any input image or frame. The output is the bounding box coordinates of the detected faces

Facial Recognition — Compare multiple faces together to identify which faces belong to the same person. This is done by comparing face embedding vectors

Emotion Detection — Classifying the emotion on the face as happy, angry, sad, neutral, surprise, disgust or fear Humans are used to taking in non-verbal cues from facial emotions. Now computers are also getting better to reading emotions. So how do we detect emotions in an image? We have used an open source data set — Face Emotion Recognition (FER) from Kaggle and built a CNN to detect emotions. The emotions can be classified into 7 classes — happy, sad, fear, disgust, angry, neutral and surprise.

Speech to Text Conversion

Speech recognition is an important feature in several applications used such as home automation, artificial intelligence etc. In this process recorded audio was given to google which creates converted wave file which is in the text format from that audio file.

VI. CONCLUSION

We propose a unified deep learning based framework for Depression Detection. In conclusion, we presented a novel approach to optimize word-embedding for classification tasks. We performed a comparative evaluation on some of the widely used deep learning models for depression detection from tweets on the user level. We performed our experiments on publicly available datasets. Our experiments showed that our CNN-based models perform better than CNN-based models. Models with optimized embedding have managed to maintain performance with the generalization ability.

We presented our results of a study that looked into the automatic detection of depression using audio and video features in a human–computer interaction setting. In particular, we set out to discover how hard it would be to fool or cheat such an automated system. In our study maximum matched healthy and depressed participants, we found that depressed participants seemed to follow the predicted pattern of lower energy levels in speech. Many of the features that have before been used in emotion recognition were also found to be significant in depression recognition. However, not all features that were significant in differentiating depressed and healthy participants were the same as with those used in emotion recognition.

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

ISSN (online): 2581-3048
Volume 9, Issue 11, pp 129-133, November-2025
https://doi.org/10.47001/IRJIET/2025.911016

from nonverbal behavior." *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, Rifat Shahriyar, Nowshin Nahar, 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.

AUTHOR'S BIOGRAPHY

Sudhanva Kale, Student, Electronics and Telecommunication Engineering, College of Engineering, Pune, Maharashtra, India.

Citation of this Article:

Sudhanva Kale. (2025). MindMirror - Mirror that Reads Emotions. *International Research Journal of Innovations in Engineering and Technology - IRJIET*, 9(11), 129-133. Article DOI https://doi.org/10.47001/IRJIET/2025.911016

****