Depression Prediction System using Visual and Vocal Features

Dr. Jayshree Pansare[1], Jagruti Dambre[2]

Dept. of Computer Engineering, Modern Education Society's College of Engineering, Pune, Maharashtra, India.

---***---

Abstract - *An Depression is a significant emotional well-being illness of human which is quickly influencing carries on with around the world. The vocal and visual component gives valuable data to treat depression. It shows the indication takes after nonappearance of enthusiasm with works out, the steady notion of harshness. Huge sadness can realize an assurance of social and physical symptoms. It could recall changes in rest, longing for, essentialness level, focus, step by step lead or certainty. Mental illness is routinely related to contemplations of self destruction. As of late, profound educated applications focused on neural systems have demonstrated unrivaled execution close by made applications in different regions. Deep learned applications that settle the above issues that may absolutely survey the level of voice and face depression. In the proposed technique, Convolutionary Neural Networks (CNN) is first created for learning deep educated highlights and distinct crude waveforms for visual articulations. Second, The MFCC procedure is utilized for vocal information handling. It is most generally utilized in speaker acknowledgment for sound highlights. To measure the depression severity scale the PHQ-9 questionnaire test is utilized to foresee the nearness and seriousness of depression. This depression recognition procedure is dependable and effective.*

Key Words: **depression detection, visual expression, vocal expression, deep learning, etc**

1. INTRODUCTION

Depression is foreseen to transform into the fourth best mental issue, according to the World Health Organization (WHO). Depression is normally difficult to examine, as it happens from various perspectives. Physical and Psychological quality of people shows a basic activity on their life execution. Dismissing this can achieve a couple of issues, for instance, stress, anxiety, misery, etc. These issues should be identified and controlled at the underlying stages itself for the better psychological wellness of the individuals. The vast majority of the individuals are absolutely unconscious that they might be having misery. In the event that they know about it, a few people cover their downturn from everybody. So a mechanized framework is necessitated that will select the individual who are managing sadness. A structure will get the visual and vocal verbalizations and concentrates the facial features and vocal features from each casing, assessments these features to envision the signs of Depression. The proximity of these features in the video edges will be bankrupt down to anticipate pity in the person. People with misery may experience a nonattendance of energy for step by step works out, basic weight decrease or expansion, a dozing issue or over the top resting, nonappearance of imperativeness, feebleness to think and dreary thoughts of death or self destruction. Around 350 million people generally experience the evil impacts of despondency, which is generally 5% of the world's finished masses. Depression is a principle wellspring of impairment around the globe, and is a critical supporter of the overall load of disorder. Depression causes one going at ordinary intervals the world over.

Programmed depression assessment reliant on visual articulation is a rapidly creating investigation region. The current exhaustive review of existing strategies as itemized in excess of sixty circulations during the latest ten years revolves around picture taking care of and AI counts. There is no comprehensive standard that depicts which feature is commonly significant for sadness treatment. Already, deep learning has been adequately applied to separate the depression. Both models and examination have indicated that a great deal of significant data can be gotten from the sound and visual signs and pictures. Numerous variations of the deep learning framework exist, for example, single-layer learning models, probabilistic models, auto-encoders, and Convolutionary Neural systems. Convolutionary Neural Networks have been broadly used to accomplish bleeding edge efficiency in numerous methodologies. In any case, the surface arrangement situation has end up being genuinely productive.

The CNN-based methodology having informational index with naturally visible pictures and the best-distributed tiny pictures. Fittingly in this examination we investigate how the expectations for depression reality can benefit by the usage of CNN in surveying visual and vocal articulations. The proposed work did to foresee the depression state as indicated by current contribution of visual and vocal information.

The PHQ-9 is the 9-interrogation despondency level of Patient Health Questionnaire (PHQ). It is self-controlled variation screening gadget that assesses 12 mental and enthusiastic prosperity queries. The PHQ-9 requests rely upon suggestive models of misery and get some data about the patient's association with the latest fourteen days. Queries are about the level of excitement for completing things, feeling down or disheartened, issue with resting, essentialness levels, dietary examples, self-wisdom, ability to imagine, momentum of effort and insights of implosion. Responses go from "0" (Not using any and all means) to "3" (reliably). PHQ-9 is in like manner the most by and large

used sorrow examines in the United Kingdom's National Health Service, which anticipates that providers should use a depression determination appliance while handling depression.

Total Score	Depression Severity
1-4	Minimal Depression
5-9	Mild Depression
10-14	Moderate Depression
15-19	Moderately severe Depression
20-27	Major/Severe Depression

Fig -1: PHQ-9 (Depression Severity Level)

2. LITERATURE REVIEW

Emotional wellness issues, for example, depression have been connected to shortages of psychological control. A fake shrewd framework is used to predict the downturn. It can predict the degree of Beck depression inventory II (BDI-II) from vocal and visual enunciations. To begin with, different visual features are isolated from outward appearance pictures. Also, otherworldly low level descriptors and mel-frequency cepstral coefficients are expelled from sound areas to get the vocal verbalizations. Third, feature dynamic history histogram (FDHH) is proposed to get the short lived advancement on the part space. Finally, these visual enunciations and vocal verbalization features are consolidated using backslide methodologies for the forecast of the BDI-II scales. [1]

A structure with the capacity of serving a decision support method is relies upon novel highlights which are removed from outward appearance math and discourse explanation, by translating non-verbal presentation of depression. The system has been tried both in sex self-governing and sex based modes, and with different blend procedures. The estimations were surveyed for a few mixes of boundaries and request plans. Perfect system execution was acquire using a nearest neighbor classifier on the decision mix of facial mathematical features in the sexual direction free mode, and talk based features in the sex based mode.[2]

Depression can be recognized by electro-physiological signs. Be that as it may, very few assessments investigate how to consistently screen patient's electro-physiological signs through a logically invaluable way for a master, especially on the seeing of electroencephalogram (EEG) signals for depression end. Since a person's mental state and physiological state are changing after some time. Through the AI methodology, system can give a sound probability of distress under each lift as a customer's self-rating score from constant EEG information.[3]

Depression is a commonplace mental issue and one of the essential inadequacies around the globe. Lacking target burdensome issue assessment methods is the key clarification that numerous burdensome patients can't be managed suitably. For assessing the association among sadness and discourse, extricate highlights however many as could be allowed by past conditions to make a tremendous voice include set. By then use some segment assurance methods to discard immaterial, abundance and uproarious highlights to shape an insignificant subset. To evaluate practicality of this new subset, using a couple of standard classifiers and 10-fold cross-validation.[4]

To recognize depression Novel methodology is used for researching facial district visual-based nonverbal conduct. Dynamic portion descriptors are isolated from facial territory sub-volumes, and scanty coding is used to undeniably arrange the removed component descriptors for depression end. Discriminative planning and decision combination are applied to moreover improve the accuracy of visual-based finding.[5]

Depression is a run of the mill and passionate wellbeing issue, which impacts on the victim just as on their families, allies and the economy when all is said in done. Regardless of its high prevalence, current end relies just upon understanding self-report and clinical conclusion, inciting different enthusiastic inclinations. Target loaded with feeling identifying structure that bolsters clinicians in their decision and seeing of clinical trouble. To separate the introduction of eye advancement features removed from face accounts using Active Appearance Models for a combined portrayal task. Oddly, in spite of the way that the squinting rate was not through and through uncommon among disheartened and sound controls, it find that the ordinary division between the eyelids (illuminating) was on a very basic level smaller and the typical term of flashes basically longer in debilitated subjects, which might be an indication of shortcoming or eye to eye association shirking.[6]

A couple of investigators have accumulated datasets by exhibiting individuals film-strips to get the outward appearances of subjects watching them. Data is in like manner assembled by giving an errand of seeing negative and positive emotions from different facial pictures. [7]

The understudies encountering from depression would show less care in study corridors. In case the understudies' emotions are planned to the activities done in study corridor, their enthusiastic state can be checked whether they are debilitated or not, and reliant on this the teacher can help the understudy by giving more thought to that particular understudy.[8]

3. PROPOSED METHODOLOGY

The proposed technique is used to construct a system that can consequently foresee the depression state from

customer's visual and vocal articulations. In this examination, the system devices used the deep learned credits to take a gander at the degree of depression. Next, CNN is used to pick up the deep learned abilities clearly from the raw visual pictures. The MFCC strategy is used for vocal data preparing. MFCC and CNN based repetitive neural system is prepared to identify depression and to foresee the depression level investigation through PHQ.
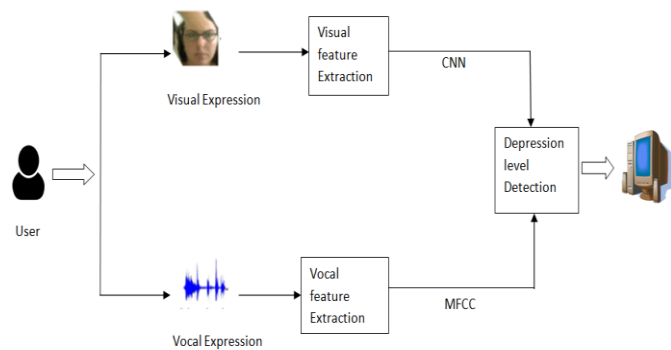

Fig -2: System Architecture

3.1 Convolutional Neural Network

Among various deep learning models CNN technique is used in highlight extraction including picture and example acknowledgment, discourse acknowledgment, trademark language planning and video assessment. CNN's information on highlight elements has demonstrated its benefits to find the significant element subtleties showed in visual and vocal signs. The profound system determines deep learned in visual highlights from outline level crude waveforms, while the other profound system model gains portrayals of articles straightforwardly from spectrogram pictures. The CNN configuration has been shown to be ground-breaking in performing incredible on various errands for instance, object affirmation, conduct acknowledgment, etc.

3.2 Mel Frequency Cepstral Coefficient

The MFCC strategy is used for vocal data handling. It is most commonly used in speaker acknowledgment for sound highlights. The vocal signs are first apportioned into outlines. By then cepstral include vector is delivered for each casing. The low-level highlights are expelled from the sound parts and normalized. Finally, the MFCC based monotonous neural framework is set up to recognize depression or to foresee the downturn reality.

4. RESULT AND DISCUSSIONS

Programmed recognizable proof of depression pulled in growing consideration from authorities in mind research, programming designing and related controls. In this manner, vocal and visual articulations are used to perceive depression severity. In this proposed work these endeavors by introducing audit of depression recognition frameworks

and talk about accepted procedures and most encouraging ways to deal with this task. Conduct of a discouraged individual shows relative change in the extents that talk structure, outward appearances and head advancement when stood out from a non-depressed person.
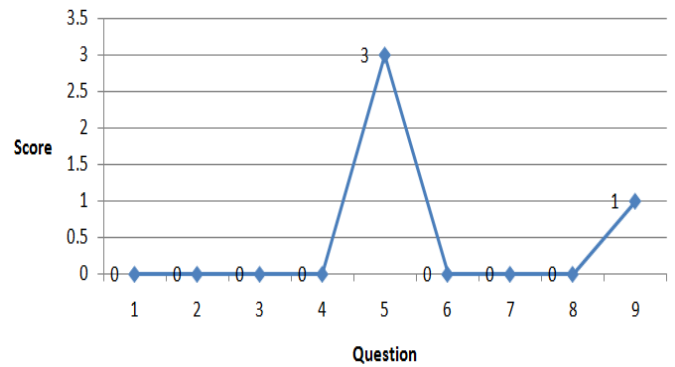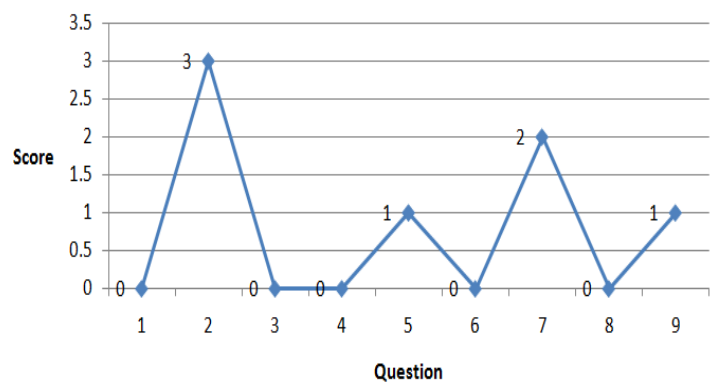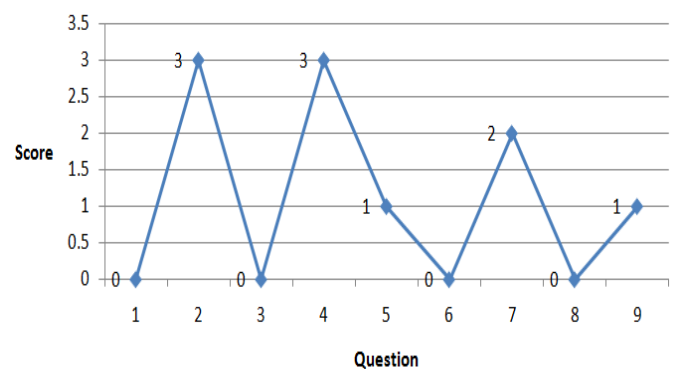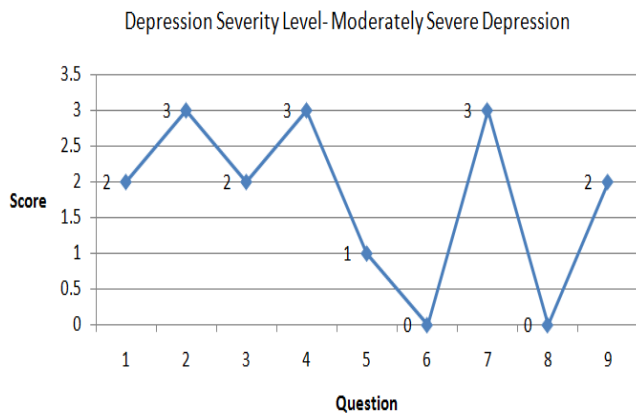

Chart -1
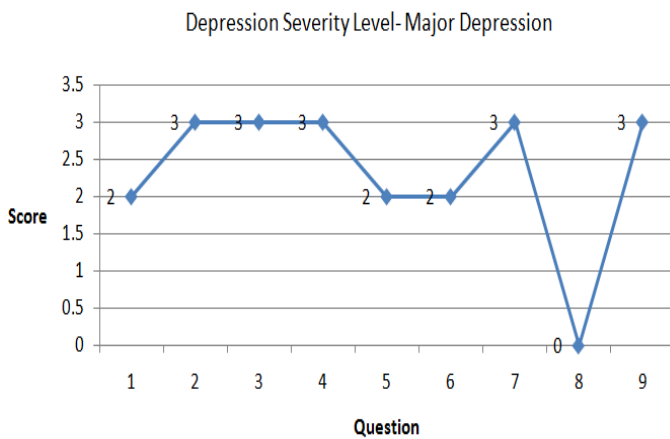

Chart -2

Chart -3

Chart -4

Chart -5

As per each chart we get the following scores for Depression Severity Level:

Table -I

Chart No.	Score Result	Depression Severity Level
1.	4	Minimal Depression
2.	7	Mild Depression
3.	10	Moderate Depression
4.	16	Moderately severe Depression
5.	21	Major/Severe Depression

5. CONCLUSIONS AND FUTURE WORK

Conduct of depressed individual shows relative change in regards to of discourse design and outward appearances (visual articulations), when contrasted and a non-discouraged person. The depression acknowledgment task through visual and vocal articulations utilizing CNN and Mel Frequency Cepstral Coefficient (MFCC) procedure gives the depression state desire. The new philosophy that fixated on deep learning and customary strategies that are utilized to address the issues of planning hand-made applications for consciousness of depression. The comparative evaluation on some of the generally used deep learning models for depression detection which improve depression recognition proficiency. For future work, look for more successful relapse models to additionally improve the exactness of depression recognition and the depression forecast is determine by utilizing text information alongside the accommodating solutions for conquer depression.

REFERENCES

[1] Asim Jan, Hongying Meng, Yona Falinie Binti A. Gaus, Fan Zhang, "Artificial Intelligent System for Automatic Depression Level Analysis Through Visual and Vocal Expressions", IEEE Transactions On Cognitive And Developmental Systems, vol. 10, no. 3, 2018.

[2] A. Pampouchidou, O. Simantiraki, C. M. Vazakopoulou, C. Chatzaki, M. Pediaditis, A. Maridaki, K. Marias, P. Simos, F. Yang, F. Meriaudeau, M. Tsiknakis, "Facial Geometry and Speech Analysis for Depression Detection", 978-1-5090-2809-2/17/31.00 IEEE ,2017.

[3] Shengjie Zhao, Qinglin Zhao, Xiaowei Zhang(B), Hong Peng, Zhijun Yao, Jian Shen, Yuan Yao, Hua Jiang, and Bin Hu, " Wearable EEG-Based Real-Time System for Depression Monitoring", Springer International Publishing AG 2017.

[4] Zhenyu Liu, Bin Hu*, Lihua Yan, Tianyang Wang, Fei Liu, Xiaoyu Li, Huanyu Kang,"Detection of Depression in Speech", IEEE 978-1-4799-9953-8/15/31.00 ©2015.

[5] Lingyun Wen, Xin Li, Guodong Guo, Yu Zhu, "Automated Depression Diagnosis based on Facial Dynamic Analysis and Sparse Coding",10.1109/TIFS. 2015.2414392, IEEE Transactions on Information Forensics and Security, 2015.

[6] Sharifa Alghowinem, Roland Goecke, Michael Wagner, Gordon Parker, Michael Breakspear, "Eye Movement Analysis For Depression Detection", IEEE 978-1-4799-2341-0/13/31.00 ©2013.

[7] Pampouchidou, Anastasia, Kostas Marias, Manolis Tsiknakis, 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] 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.

[9] 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.

[10] 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.

[11] 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.