Multi Modal Approach For User Responsive Depression Detection

A Synopsis

submitted in partial fulfillment of the requirements of the degree of

Doctor of Philosophy

in

Electronics & Communication Engineering

Submitted By:

Jiger Prakashchandra Acharya

[189999915002]

Supervisor:

Dr. Milind S.Shah

Professor, E.C. Department

Shantilal Shah Engineering College, Bhavnagar

submitted to

Gujarat Technological University

Ahmedabad, Gujarat, India.

Content

List of Figure

List of Table

I. ABSTRACT

During last one decade the mental health is becoming a major burning issue to society as number of various socio-economical, personnel and societal issues related to mental health are increased in exponential manner. Depression is mood disorders which result in severe disabling conditions which affect person's ability to cope with routine life and real-life challenges which are dynamic in nature. It may occur when person remain more than two weeks in negative state of mind continuously. Depression has observable behavioral symptoms related to affective and psychomotor domains which can be identified by human and/or computer. Classical approaches majorly depend on person's behavioral analysis and family observations during clinical interviews, which are effective if it can be precisely defined and assessed but persons have tendency to conceal it so they are less effective while proposed method with Deep Learning techniques perform the said task with accuracy of 79% and open a new era for mental health care domain as depression detection tool.

II. BRIEF ON THE STATE OF THE ART OF THE RESEARCH TOPIC

Depression will be leading cause of disability globally in all age groups which severely affects the global health index and social impact. It is an ongoing problem which may resolve with proper care but may reoccur via various risk factors. The medical community does not fully understand the causes of depression but mostly known various risk factors like environmental, psychological, social, financial factors, drug addictions change in genetic features, brain's neurotransmitter levels, and accidental events. Early detection plays a crucial role as by positive counseling, various physical and mental exercises like meditation and proper medication may resolve depression else situations become worst up to suicidal tendency. Human are expressive through verbal and nonverbal communication channels both channels are important to express feelings [4], [5]. Person can remain silent to suppress verbal communication and by fake expression or conceal expression one can hide real scenario still researcher believes that facial expression is crucial input to present emotional status of person so psychologists try to map facial expressions to emotional states. When Person is feeling depression both verbal and nonverbal channel give strong indicators [6]. Number of researchers has contributed their work to identify depression which closely related with psychology, affective, computer and clinical domains so in real time it's a multi-challenging task. As domain start with Psychology, Computer Science and Clinical domain and last long to social domain the major key challenges are mentioned as under to map psychological domain to computer and validate it with behavioral / clinical predictions.

A Literature Review:

Human express their emotions through multiple channels which can be monitored for emotion analysis [4],[5]. There are many approaches like Clinical, Visual, Verbal, Colour psychology, Self Response clinical inventory, EEG etc. To determine the mental health of a person various approaches brief mentioned below.

(1)Clinical based approach: Person is assessed critically by medical mental health expert of psychology based on response of one- to-one interaction and real time observation by medical expert decision is made and psychological treatment and/or medication is applied [6]. Major problem with this approach is quantity of such assessment as person doesn't accept such mentalhealthissuesand becauseofsocialstigma,they havetendency to concealsuch issues so approximately 10%actually affected get benefits of this approach.

Self –Response Inventory (S.R.I): Quiz based approach in which person is required to respond quiz truthfully based on response each question generate number and based on additive score mental health is decided.

Hamilton Rating Scale for Depression (HRSD): popular as HAM-D which is MCQ type questionnaires which medical professional use for depression severity assessment [33].

Beck Depression Inventory (BDI): The BDI is 21 questions which are multiple option types self-report which calculate the severity of depression symptoms and feelings based on the response given [38].

Patient Health Questionnaires (PHQs): ThePHQ-9 has nine questions about person's mood and daily activities like appetite, watching television which helps doctor come to a depression diagnosis.ThePHQ-2asksthefirsttwoquestionsofthe PHQ-9.Which is a screening test guide doctor for further processing [35].

Zung Self–Rating Depression Scale: It is again short surveys which identify depression levels [37].

Centre for Epidemiologic Studies-Depression Scale (CES-D): is a self-reported scale with 20 questions designed for caregivers to measure how often they have depression symptoms. A higher score can help doctor to identify risk level of depression [38].

Table 1: Comparative study of existing questionnaires approach

Sr	Tool with timeline	Type	Questions	Classification
1	Beck Depression Inventory (BDI),1961	Self-report	21	0-13: Min 14-19:Mild 20-28:Moderate 29-63:Severe
2	Beck Depression Inventory-Short Form (BDI- Short),1978	Self-report	13	0-4: Minimal 5 -7:Mild 8 - 15:Moderate >15:Severe
3	Center for Epidemiologic Studies Depression Scale (CES-D),1977	Self-report	20	< 16: Minimal 16-26:Mild-Moderate >26:Severe

4	Center for Epidemiologic Studies Depression Scale-Short Form (CES-D-Short),1997	Self-report	10	< 10: Minimal 10-14:Mild > 14:Severe
5	Medical Outcome Study-Depression Scale(MOS),1992	Self-report	8	Low Score: Normal HighScore:Higher Depression
6	Self-Diagnostic Depression Scale Primary Care (SDDS-PC),1994	Self-report	7	0-2: Normal 3-4:Mild 5-7:Severe
7	Patient Health Qustionnaire-9, (PHQ-9),2001	Self-report	9	0-4: Minimal 5-9:Mild 10-14:Moderate 15-19: Mod-Severe 20-27: Severe
8	Quick Inventory of Depressive Symptomatology, (QIDS),2003	Clinician/Self	16	0-5: None 6 -10:Mild 11-15:Moderate 16-20: Severe
9	Depression Measure Inventory -10 (DMI-10),2004	Self-report	10	0-7: Normal 8-14: Mild Moderate 15-20: Moderate >20: Severe
10	Proposed	Self-report	21	0-4: Normal 5-8: Mild 9-14: Moderate >14: Severe

(2) Visual channel base approach: in computer vision, pattern matching and image community detect depression features from behavior visualchannels.Wageretal. revealed that depression can be judge by head, pose movement of person and tendency of social withdrawal [8], [9] which can be judge from manual and automatic analysis. Cohn et al. [10] prove that during depression verbal and nonverbal channels are affected significantly which can be identified during external observations.

(3) Audio based approach: Ooi et al [11] in 2013 present system which automatically measure audio parameters for prediction possibility of development of depression in near future which is very useful for identification of depression features in adolescents with use of a novel multichannel classification approach with classification accuracy of 73% in prospective screening and major work on verbal channel reported by [12].

(4) Colour psychology-based approach: reveals how various colours impact person's mood and so behaviors.In 2022 Yang et al [32] enlightens colour impact on person's emotional response based on age and cultural background etc. Major evidence in this emerging area is based on the feedback result of anecdotal, but domain expert has made a few important conclusions that how colour theory correlate with behaviors.In2020 Abu- Akel et al [31] surveyed the emotional associations of 4598 persons from 30 countries and conclude that people commonly associate certain colours with specific emotions as below:

Table 2: Colour association with Emotion and Mood

Colour	Percentage	Emotions	Mood Category
Black	51%	Sad	Negative
White	43%	Relief	Positive
Red	68%	Love	Positive
Blue	35%	Relief	Positive
Green	39%	Contentment	negative
Yellow	52%	Joy	Positive
Purple	25%	Pleasant	Positive
Brown	36%	disgust	Negative
Orange	44%	Joy	Positive
Pink	50%	Love	Positive

(5) Bio sensor approach: in which various biomedical instruments and sensors can be used to identify the person's behavioral analysis but user acceptability of such approach is very less.Such sensor like EEG psychological signals received from brainare measured in real time manner which cannot be suppress or conceal like audio, video or textual information's so the reliability of such system is very high but acceptability is very poor [23], [29], [30].

(6) Social media approach: in present era of social networking platform person spare more time on screen so social media like face book, twitter, snaps etc. are reflect person's mood and emotions so based on status, text message and uploaded images and associated tag person's mood can be categorized from text messages, pictures or emoji person's mood behavior can be examined [30].

(7) Multi-modal approach: Many contributors' fuses above single modalities called multi modal approach in which person's data from multiple channels are evaluated make emotion or mood detection more accurate and reliable compared to single modality [28],[29],[30],[41].

Table 3: Literature Review Analysis.

Author(s)	Key work	Methodology	Performance Results	Limitations
Browner, W.S. et al [39] Journal of General Internal Medicine, 12: 439-445. 1997	Case-Finding Instruments for expression:Two questions Are as Good as Many	Classical, standardize and clinical questionnaires approach Result Analysis of SevenDepressionCase findingInstrument on 536 subjects	(1) "During the past month, have you often been bothered by feeling down, depressed, or hopeless?" (2) "During the past month, have you often been bothered by little interest or pleasure in doing things?"	Result relies on response of 2 answers only May have Higher False Positive ratio
Williams JB. et al [35] Journal of General Internal Med. 2001	The PHQ-9: validity of a brief depression severity measure.	Classical, standardize and clinical questionnaires approach	PHQ-9 scores of 5, 10, 15, and 20 represented mild, moderate, moderately severe, and severe depression	Proven standardized approach May have Higher False Positive ratio
Trivedi MH,et al. [40] a psychometric evaluation in patients with chronic major depression. Biol Psychiatry. 2003	The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR)	QIDS, QIDS-C, QIDS-SR	16-item quick inventory of depressive symptom clinician rating and self-report for psychometric evaluation in patients with chronic major depression	Alarming rate of false detection

PHQ-9 MDD module of the full PHQ et al. Levis B et al [36] 2019	Accuracy of Patient Health Questionnaire-9 (PHQ-9) for screening to detect major depression: individual participant data	Diagnose grade severity of symptoms Scores each of the 9 DSM criteria of MDD on 0-27 severity score. How difficult have these problems made it for you to do your work, take care of things at home, or get along with other people? is good indicator	High score decreased functional status and increased symptom-related issues.	May high false-positive (50%) positive actually had major depression
Jeffrey F Cohn etal. [10] Image and vision computing, pp. 641-647, 2014	Nonverbal social withdrawal in depression: Evidence from manual and automatic analyses	Automated tool for FER	Depress person reduced affiliative behavior like motivation to co-operate, comfort and request etc.	Automatic facial expression analysis can reduce human participation and burden.
Sahla, K. S et al. [20] InternationalSymposium on Intelligent systems Technologies and Applications, pp. 475-486.Springer	Classroom Teaching Assessment Based on Student Emotions	Emotion analysis of students from visual channel	face detection, LBP encoding and mapping using deep CNN	Uni modal Visual feature approach

Tzirakis et al [22] IEEE Journal of Selected Topics in Signal Processing, vol. 11, no. 8, pp. 1301–1309	End to-End Multimodal Emotion Recognition Using D.N.N.	Deep Neural Network	Work fuses three modalities and their complex interactions of spoken text, visual and vocal features.	Complexity
Dr. D. Venkataraman et al [24] International Journal of Pure and Applied Mathematics Volume 118 No. 7 2018, 455-463	Extraction of Facial Features for Depression Detection among Students	Automated tool for Depression detection applied in classroom T/L environment.	Automated system pick students via captures frontal face videos, extracts facial features from frame analyses to detect signs of depression	trained with frontal face images of Happy, Contempt and Disgust faces
S. Sahay et al[23] Proceedings of Grand Challenge and Workshop on Human Multimodal Language (Challenge-HML)-2018	Understanding Affect from video	Multimodal relational tensor network for sentiment and emotion classification	language, audio and video domains fuse in multimodal techniques to fuse the modalities	Treat the segments of a video independently.
M. M. Hassan et al [27]Information Fusion, vol. 10.1016/j.inffus. 2018.10.009. 51, pp. 10–18,2019	Human emotion recognition	Deep belief network (DBN) architecture	DBN for depth level feature extraction from fused observations of EDA, PPG and zEMG sensors statistical features to prepare a feature-fusion vector for classification	Classification of Happy, Relaxed, Disgust, Sad,Neutral expression Sensor based operating complexity
Priyasad et al [28] IEEE International Conference on Acoustics, Speech and Signal	Attention driven fusion for multi-modal emotion recognition	Emotion recognitionusing acoustic and text	SincNet layer and DCNN for acoustic feature extraction and parallel DCNN and Bi-directional RNN branches with cross-attention	Emotion recognition using acoustic and text

Processing (ICASSP), 2020			for text processing.	
J. Njoku et al [29]The Journal of Korean Institute of Communications and Information Sciences, vol. 47, no. 1, pp. 79–87, Jan. 2022	Deep Learning Based Data Fusion Methods for Multimodal Emotion Recognition	AdvancedDeep-learning to identify emotions	Visual, Vocal and physiological EEG signals data fusion at early, late, hybrid fusion.	Create a emotional profile avoiding noise and natural ambiguity of human expressions. Complex operating(EEG) for detection
Komal et al [30] Reliability: Theory and Application "ISSN 1932-2321, Volume 17 Issue 1 PP-67 March 2022 SCOPUS	Effect Of Preprocessing In Human Emotion Analysis Using Social Media Status Dataset	Social media status dataset is first pre-processed using various methods, given for feature extraction and classification purpose.	Machine Learning approach, count vectors and TF-IDF techniques for extracting the different features	Complex operating for detection
Abu-Akel A et al [31] Psychol Sci. 2020;31(10):1245-1260.	Universal patterns in color-emotions associations are further shaped by linguistic and geographic proximity	Explore the association of emotion with colour in Affective communication.	2020 study surveyed the emotional associations of 4,598 people from 30 different countries found people commonly associate certain colors with specific emotions.	According to the study results black as 51%negative, Red as 68%positive emotions
Yang J, Shen X et al.[32] J Environ Public Health. 2022	Application of color psychology in community health environment design	Symbolic meaning of each colour is revealed	Conclude that can be used as therapy	Red spectrum partled comfort to anger while blue spectrum part described as calm evokes feelings of sadness.

Shanliang Yang et al [41] Science Direct 2024	Enhancing multimodal depression diagnosis through representation learning and knowledge transfer	Multimodal fusion visual, audio, text signals and Knowledge transfer from pretrained language models (SentiLARE)and multimodal emotion datasets (CMU-MOSEI)	Auto encoder-based representation learning with BiLSTM encoders/ encoders/decoders and fused feature for final classification. conditions	Integration of findings from longitudinal and epidemiological studies provides bidirectional relationship between depression and chronic health

Table 4: Authors contribution of domain related research

Sr	Authors	Type	Dataset
1	Becker et al. (1994)	Audio/Video	Dementia Bank Database Reddit Self-reported
2	Valstar et al. (2013)	Audio/Video	AVEC 2013
3	Pradhan et al. (2014)	Text data	SemEval-2014 Task 7
4	Lieberman et al. (2013)	Text data	Crisis Text Line
5	Valstar et al. (2014)	Audio/Video	AVEC 2014
6	Gratch J et al. (2014)	Audio/Video	The Distress Analysis Interview Corpus
7	Valstar et al. (2016)	Audio/Video	AVEC 2016
8	Videbech (2016)	Audio/Video /Report	The Danish Depression Database
9	Yates et al. (2017)	Text data	Depression Diagnosis (RSDD) dataset
10	Garcia-Ceja et al. (2018)	Audio/Video	Depression
11	Komal Anadkat et al. (2022)	Audio / Video / Social media text/ EEG	FER2013, Ravdess, Kaggle's social media dataset, and EEG dataset

Major domain related challenges are as below:

- Psychology versus Computer Science: Psychology and Computer science have complexity and dependency in their specific domain.
- Natural versus posed expression: There are vast differences in actual and acted expressions. The expression intensity, reality, and movement of muscles at macro and micro expressions play a major difference in expression execution.
- Expression versus Emotion: Expression is movement of muscles while Emotions are related to heart. Expression can fabricate or acted but emotions are pure.
- Static versus Dynamic analysis: Static image analysis is still image based while dynamic or video-based analysis has many frames which required being process.
- Transitions among expressions and fusion expression: When transition of expressions occurs fixed out one expression is quite difficult and separate expression individually from fusion expression are quite challenging.
- Variationsin intensity level of facial expression: The intensity difference between expressions is subjective which is based on age, race, sex and ethnicity so take a decision globally quite challenging.
- Deliberate versus spontaneous expression: some time execution of spontaneous expression differs from deliberate expression.
- Headorientationandscenecomplexity:Headismajorregionwhereexpressionexecutionhappe nssoextract Region of Interest (ROI).
- Neural Network limitation to train and execute Neural Network model we required large and variety amount of data which is a big challenge.
- Spatial temporal setting during data collection: as in Image processing spatial temporal data play a crucial role and as it is very sensitive to light carefully data collection required in real time environment so standard metrics and requirement of wild implementation play a big role.
- Emotional classes versus data classes in data labeling: Mapping of emotions classes and standard labeled data are big challenges.
- Over fitting versus generalization: Balancing these two events is a quite difficult for any Neural Network.
- Diagnostic criteria, clinical instruments, and performance metrics which have wide variety globally.
- Clinical specificity as many psychiatric and medical disorders-occur.
- Temporal dynamics important in psychomotor retardation.
- Diversified large samples cultures, ethnicities, ages, and genders etc.

B. Research Gaps:

Non availability of user friendly, real-time, multi-modal depression detection tool.

(1) Negligence of mental health – care so quality and quantity of mental health care service provider
(2) Because of social stigma actually suffering person also not accept the ground truth and not ready to seek such professional help
(3) In Current scenario though various approaches for depression detection exists user-friendliness is major concern.
(4) Key challenges include limited labeled data, high complex sensor-based methodology for depression detection.

III. Definitions of the Problem Statement

"To prepare real-time mood monitoring and prediction tool which can further classify person's present mood in positive and negative mood from behavior analysis"

The increasing issues related tomental health highly demand real-time, automated systems that can accurately detect depression in user friendly manner.

IV. Research Objectives and Scope of Work

Objectives: The primary objective of the research is to design and implement real time,Multi Modal User Responsive tool for depression detection for said task we have selected three modalities called Visual features, Self Response Inventory and Colour psychology

To achieve this, the study is guided by the following specific objectives:

Objective 1: Drive deeper into existing depression detection system.

Objective 2: Developed a user-friendly real-time multi-modal depression detection system.

Objective 3: To validate the proposed models through performance parameters.

Objective 4: Apply the tool in wild conditions to confirm its compatibility, reliability and effectiveness.

Objective 5: Validate Automatic mood detection tool: Self Response Inventory

To detect Depression or depressive disorder person behavior analysis is required to carry out for that from existing approaches we have selected three simple and user-friendly options called Visual features from (static picture and/or video), Self Response Inventory (SRI) in which unique questionnaire related to physical and mental health related questions along with colour psychology are asked and if they are correctly responded we can judge the mood of person.

Scope of work:

1. Design a user-friendly depression detection system.

2. Develop a real-time mood monitoring and prediction tool which can further classify person's present mood in positive and negative mood based on behavior analysis.

3. Validate such tool on performance matrix and parameters.

4. Validate such tool in realtime wild conditions.

5. Apply the tool in wild conditions to confirm its compatibility, reliability and effectiveness.

V. Original Contribution by Thesis

We have decided to use three channels for emotion recognition

1) Visual features through static and video
2) Self Response Inventory (S.R.I.) a unique questionnaires for mood detection
3) Colour psychology which is decided in above S.R.I.

Depending on results of above modalities we can decide emotions/mood of a person.

Visual as major channel for detection:

Face is the strong channel of emotions indication. People reflect their feelings knowingly or unknowingly by their facial expression so face is considered as prime part of body for expression so based on this fact Ekman [4] categorize standard expression of six fundamental feelings along with head orientation, posture and movement majorly use to recognize depression [8],[9].Girard et al [10] explore strong co-relation between visual direction and depression. Table 5 list major visual features which are crucial for behavior analysis.

Table 5: List of major Visual Features for Behavior Detection

Features	Association with Facial Features
1	Eyelid opening / squinting movement
2	Outward appearance event with inconstancy with force
3	Head appearance with direction and deployment
4	Face movement
5	Pupil enlargement / inclination
6	Iris development
7	Grin force and term
8	Eye stare
9	Dismissal articulation/negative event
10	Mouth corner appearance

VI. Methodology of Research, Results and Comparisons

Real-time depression detection methodology

This research develops a depression detection system which incorporates multiple data as inputs from image, video (stored and live) and response provided by user in quiz assessment. as shown in Fig. 1. Data acquisition, data processing, system construction, and model performance evaluation are described as follows

Model take input data from various channels like camera for static and video images input, Stored image and/or video upload options for testing is also available and third input is of quiz in which user response are process and based on the outcomes of all modalities person's mood and its severity identified.

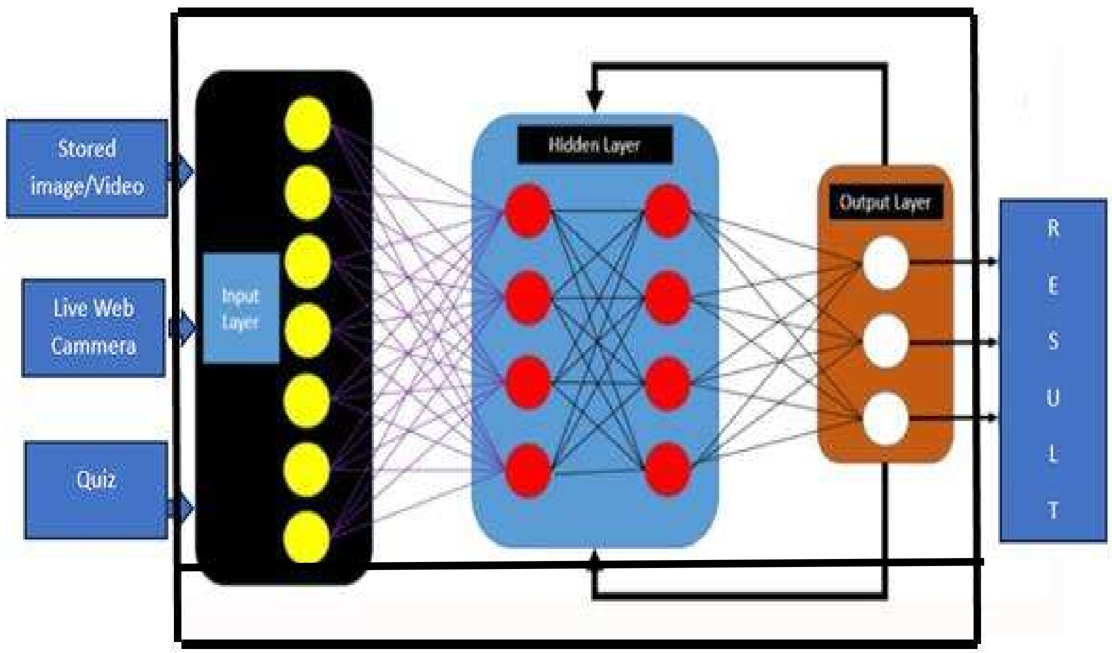

Figure 1: Depression detection Deep neural network model

Model Architecture:

Deep Learning model of CNN (Convolution Neural Network) is used for creating, training and analyze expression from static image and video channel major components of architecture are as below:

1. Convolution Layers: Multiple Convolution layers to extract features from images using multiple filters (32, 64, 128, and 256) of kernel sizes (3 x 3) which extract features from images.

2. Rectified Linear Unit (ReLU): function used for non-linearity introduction.
3. Max Pooling layers: perform reduction of spatial dimension reduction so computational complexity also reduced.
4. Dropout: Layer drops randomly some nodes in training phase so model does not have tendency of Over fitting.
5. Flatten Layer: Converts the 2D feature maps into a 1D vector for input to dense layers.
6. Dense Layers: These are fully connected layers for seven expression classifications

Above layers include Image Pre-processing, Data Augmentation, data generation, training and Validation.

Database:

The FER2013is open dataset (Facial Expression Recognition 2013) [13] which consist of facial images along with their emotion categories of person. The database has facial images of size 48×48 pixels of monochrome images with seven different categories of emotions Angry, Disgust, Fear, Happy, Sad, Surprise, and Neutral. The dataset having 28709 images in the training set and 3589 images in testing set.

Data acquisition:

Depression detection model obtained data from three sources which are static/dynamic images capture from camera both live and stored and quiz responses.

Data channel -1: Input image from Camera in real time video

Image and/or Video from camera is captured and continuously images are identified in categorical expression and positive and negative expression is determined. As shown in above figure face images are classified broadly into positive and negative groups. Happy is positive bias images where Sad, Fear, Angry, Surprise and Disgust are negative biased images. Neutral has no biasing as per its name.

Data channel -2: Stored image and video

Stored Image and/or Video from camera captured and continuously images are identified in categorical expression and positive and negative expression is determined.

Data channel -3: Quiz response

SRI quiz has unique general questionnaires associate with physical and mental health based on the answer selected by user it categorizes person’s mood identified from quiz response based on past 2- or 3-days experience.

Table 6: Proposed Self Response Inventory (S.R.I.)

Sr	Question (Select option based on past 2-3 days experience)	Outcome	Yes	No	Not Decided
1	Have you taken enough sleep?	Rest			
	Evaluates your physical well-being and rest. Adequate sleep is essential for health and productivity.				
2	Have you taken enough healthy food (in bf / lunch /dinner)?	Food			
	Assesses whether you're meeting your nutritional needs for energy and overall health.				
3	Have you done any exercise/yoga/workout?	Exercise			
	Reflects your physical activity levels, which are important for maintaining health, fitness, and mental well-being.				
4	Are you enjoying your daily routine?	Enjoyment			
	Indicates satisfaction and contentment with your daily life and habits.				
5	Can you select your clothes by yourself?	Decision			
	Measures independence and autonomy in decision-making and self-care.				
6	Can you prepare / suggest foods? (For bf / lunch /dinner)?	Selection/Creativity			
	Evaluates self-sufficiency in meal planning and preparation.				
7	Have you interacted with your family members to share your thoughts or feelings?	Sharing feelings			
	Assesses emotional connectivity and communication with loved ones.				
8	Are you comfortable to communicate with your neighbors / colleges or friends?	Communication			
	Measures your social comfort and ability to engage with people in various settings.				
9	Do you spare some time to do your liking things?	Hobby /liking			
	Reflects your balance between responsibilities and personal enjoyment or hobbies.				
10	Can you read newspaper/books/articles easily?	Concentration			
	Evaluates cognitive function, literacy, and engagement with information.				
11	Can you watch TV/mobile/Movie easily?	Concentration			
	Measures ease of leisure activity and mental relaxation.				
12	Are you satisfied with your individual achievement/contribution to family/Society?	Self-esteem			
	Reflects self-assessment of your accomplishments and your sense of purpose and fulfillment of your efforts and knowledge/Skills and attitude.				

13	Do you have interest in things which appeal you a lot previously?	Active Participation			
	Examines personal interests and desires related to romantic/pleasure connections.				
14	Do you feel your life going smoothly?	Satisfaction			
	Reflects your perception of life's trajectory and personal well-being.				
15	Do you feel tomorrow is better than present?	Optimistic			
	Indicates optimism and hope for the better future.				
16	Do you feel you are unique and no one can replace yourself?	Self-esteem			
	Reflects self-worth and a sense of individuality.				
17	Do you feel stress of your study/work /responsibilities?	Stress			
	Whether you experience feelings of pressure, anxiety, or overwhelm due to your academic, professional, or personal duties.				
18	Do you feel nothing going on track and after a lot of efforts from your side u can't do anything?	Out of control situation			
	if you ever feel like, despite putting in a lot of hard work and effort, things still aren't progressing as they should. It reflects a sense of frustration or helplessness,				
19	Do you feel frustrated or aloneness?	Helplessness			
	Reflects feelings of frustration, helplessness, or burnout. The outcome of this question could reveal several underlying emotional or psychological states, such as: Frustration or Disappointment, Helplessness or Powerlessness, Burnout, Lack of Progress or Direction, Self-Doubt.				
20	Selection of Colour base on your mood from these : Black/Red/None	Colour			
	Mood association with colour.				
21	Have you come across with negative thoughts /suicidal thoughts?	Negative			
	Negativity				

Scores obtained from responses - quiz score is converted from numeric to mood outcome label in classes as follows:

Table 7: Classification of S.R.I. score

Sr	Scorerange	Level	Depression	Actions
1	0-4 points	Level 0	No depression	Just Calm…continue routine Stay healthy physically and mentally
2	5-8points	Level 1	Mild depression	Repeat test, Take your mental health consciously
3	9-14 points	Level 2	Moderate depression	Repeat test, still result is same advisable to consult medical professional
4	more than 15 points	Level 3	Severe depression	Immediately consult medical professional

Figure 2: Prototypic Facial Expression intra class variability

Figure 3: Testing Facial Expression intra class variability

Figure 4: Testing Facial Expression interclass variability

As shown in above figures face images are classified broadly into positive and negative groups. Happy is positive bias images where Sad, Fear, Angry, Surprise and Disgust are negative biased images. Neutral has no biasing as per its name.

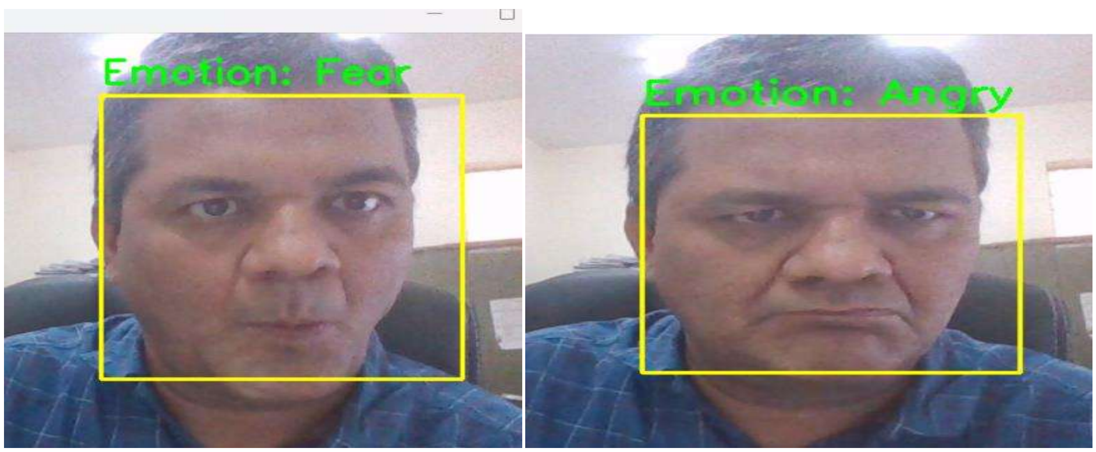

Figure 5: Testing Result of Fear and Angry expression by Proposed System

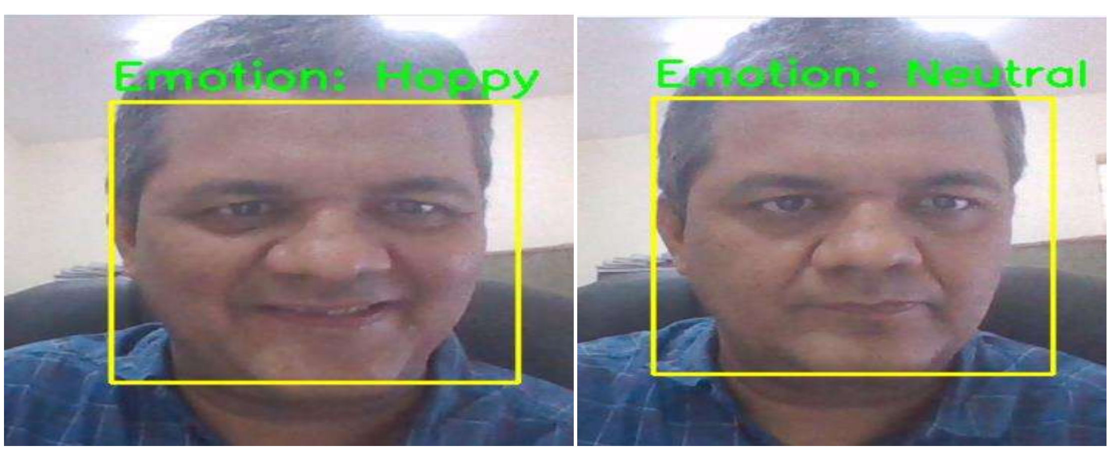

Figure 6: Testing Result of Happy and Neutral expression by Proposed System

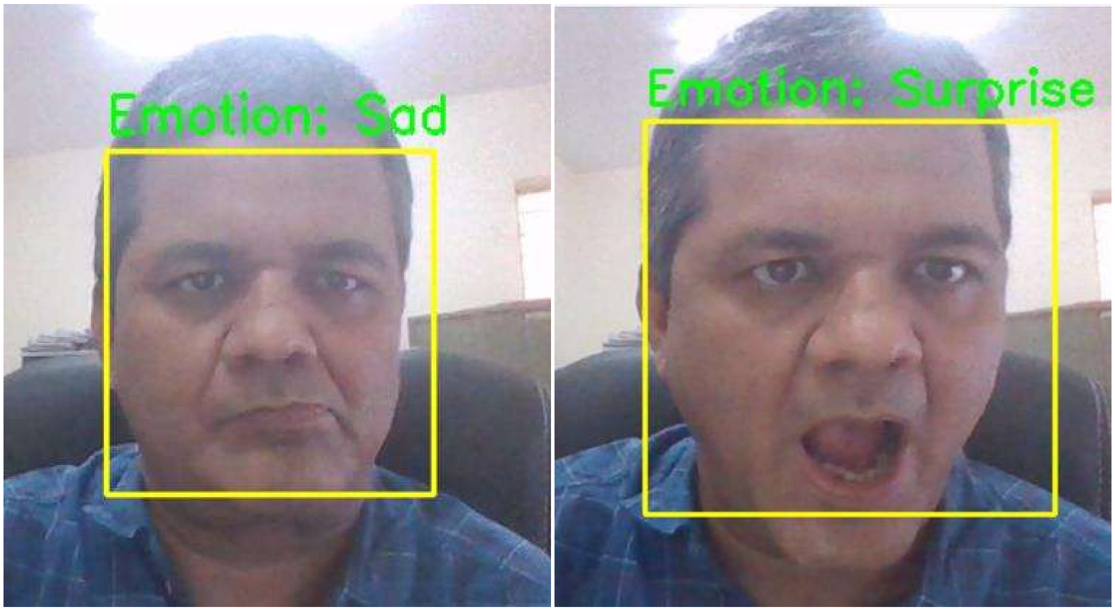

Figure 7: Testing Result of Sad and Surprise expression by Proposed System

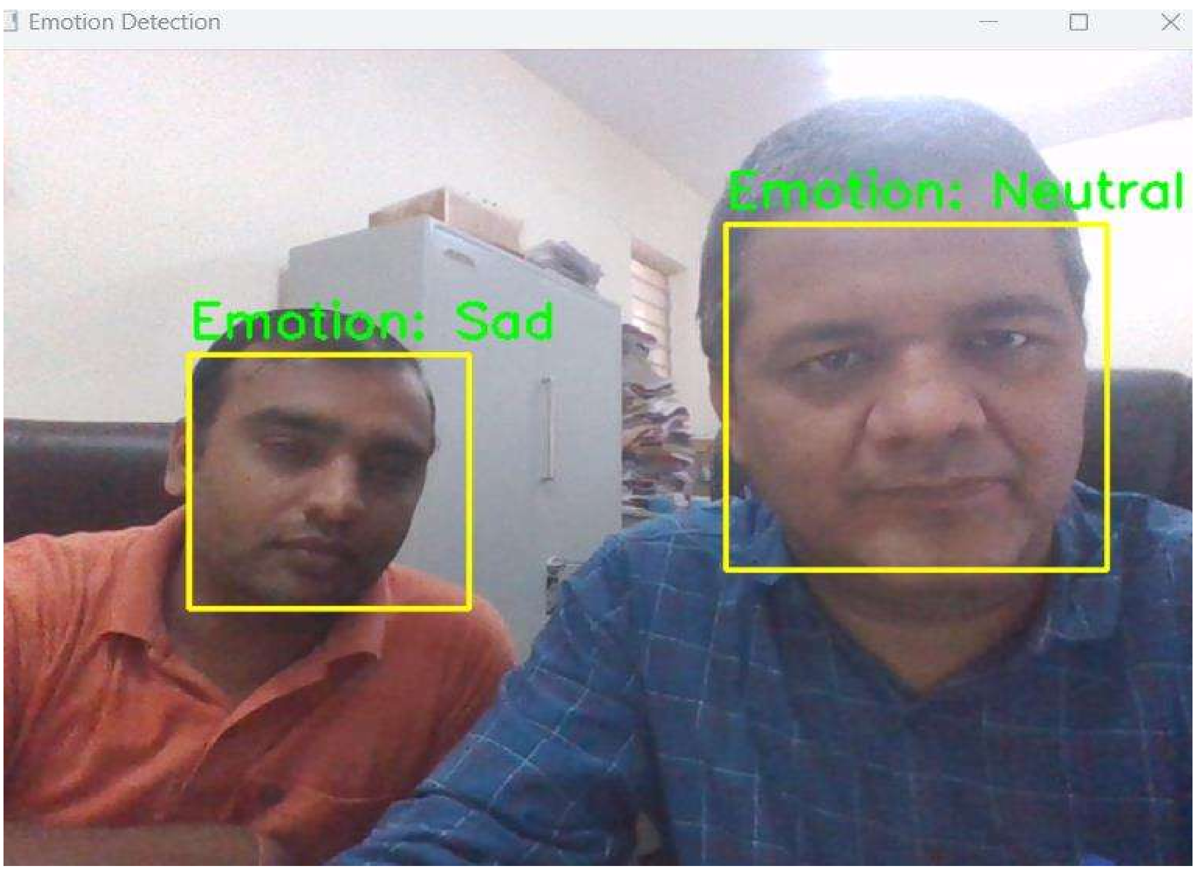

Figure 8: Testing Result of Sad and Neutral expression by Proposed System

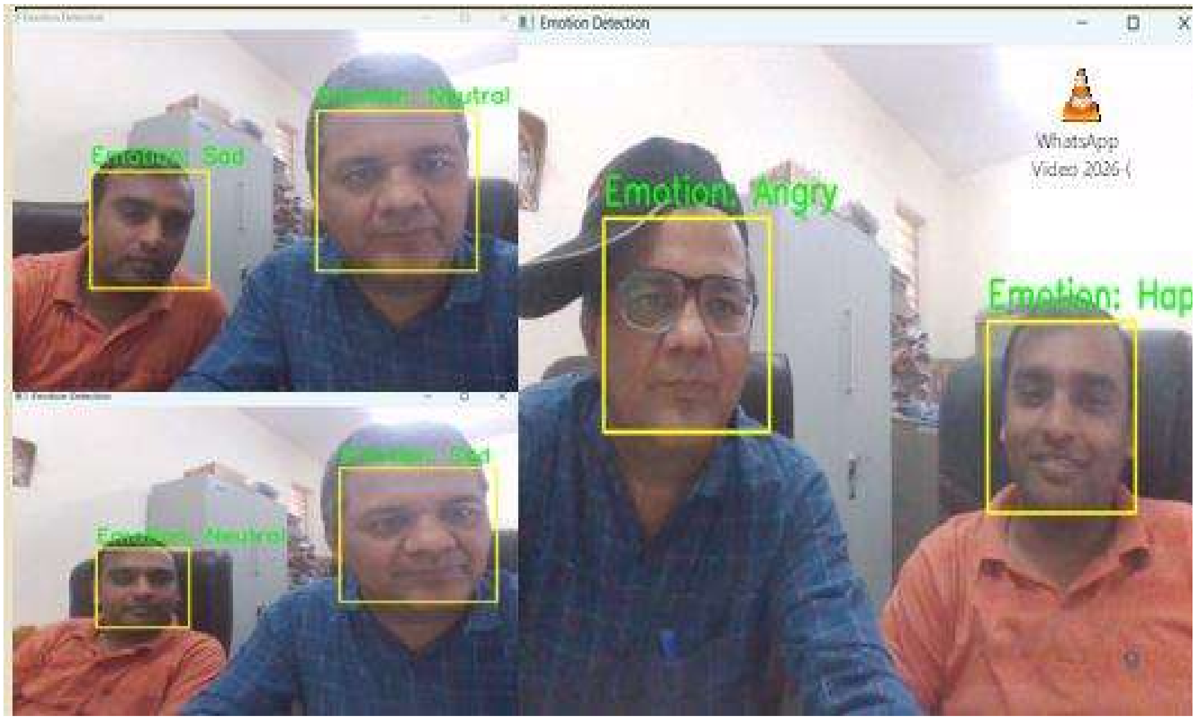

Figure 9: (from Left to Right)
Testing Result Sad and Neutral expression by Proposed System
Testing Result Neutral and Sad expression by Proposed System
Testing Result Angry and Happy expression by Proposed System

Figure 10: (from Left to Right)
Testing Result Sad and Neutral expression by Proposed System
Testing Result Angry and Neutral expression by Proposed System
Testing Result Happy expression by Proposed System

Figure 11: Testing Result of Sad and Happy expression by Proposed System

4. Results Analysis

Table 8: Model performance evaluation

(Epoch wise Training loss, Training accuracy, Validation loss and validation accuracy)

Epoch	Training Loss	Training Accuracy	Validation Loss	Validation Accuracy
1	2.2673	0.2043	2.0278	0.2493
2	1.8677	0.2697	1.6808	0.3757
3	1.6180	0.3736	1.6879	0.3589
4	1.4435	0.4470	1.2770	0.5007
5	1.3103	0.4982	1.2653	0.5237
6	1.2351	0.5255	1.2923	0.5056
7	1.2070	0.5357	1.2488	0.5189
8	1.1475	0.5628	1.0932	0.5803
9	1.1403	0.5671	1.1497	0.5628
10	1.0962	0.5850	1.0529	0.5992
11	1.0821	0.5887	1.0385	0.6027
12	1.0651	0.5986	1.0442	0.5936
13	1.0482	0.6021	0.9976	0.6180
14	1.0344	0.6097	1.0922	0.6041
15	1.0152	0.6157	1.1200	0.5887

16	0.9960	0.6218	0.9868	0.6250
17	0.9973	0.6255	1.0060	0.6271
18	0.9561	0.6398	0.9771	0.6411
19	0.9605	0.6380	1.0074	0.6082
20	0.9468	0.6449	1.0069	0.6341
21	0.9391	0.6445	0.9674	0.6487
22	0.9247	0.6490	0.9729	0.6522
23	0.8960	0.6622	0.9405	0.6508
24	0.9082	0.6603	0.9881	0.6187
25	0.8918	0.6643	0.9616	0.6341
26	0.8988	0.6594	0.9764	0.6529
27	0.8824	0.6713	0.9893	0.6243
28	0.8660	0.6786	0.9972	0.6397
29	0.8636	0.6774	0.9671	0.6508
30	0.8598	0.6847	0.9640	0.6459
31	0.8403	0.6897	1.0170	0.6501
32	0.8273	0.6951	0.9798	0.6543
33	0.8183	0.6951	0.9737	0.6278
34	0.8195	0.6944	0.9706	0.6655
35	0.8011	0.6966	1.0212	0.6404
36	0.8174	0.6997	0.9421	0.6397
37	0.7934	0.7029	1.0191	0.6404
38	0.7851	0.7029	1.0593	0.6397
39	0.7841	0.7100	0.9531	0.6494
40	0.7779	0.7063	0.9565	0.6522
41	0.7669	0.7110	0.9169	0.6745
42	0.7594	0.7183	0.9593	0.6536
43	0.7406	0.7243	0.9267	0.6767
44	0.7310	0.7318	0.9420	0.6725
45	0.7191	0.7329	0.9537	0.6613

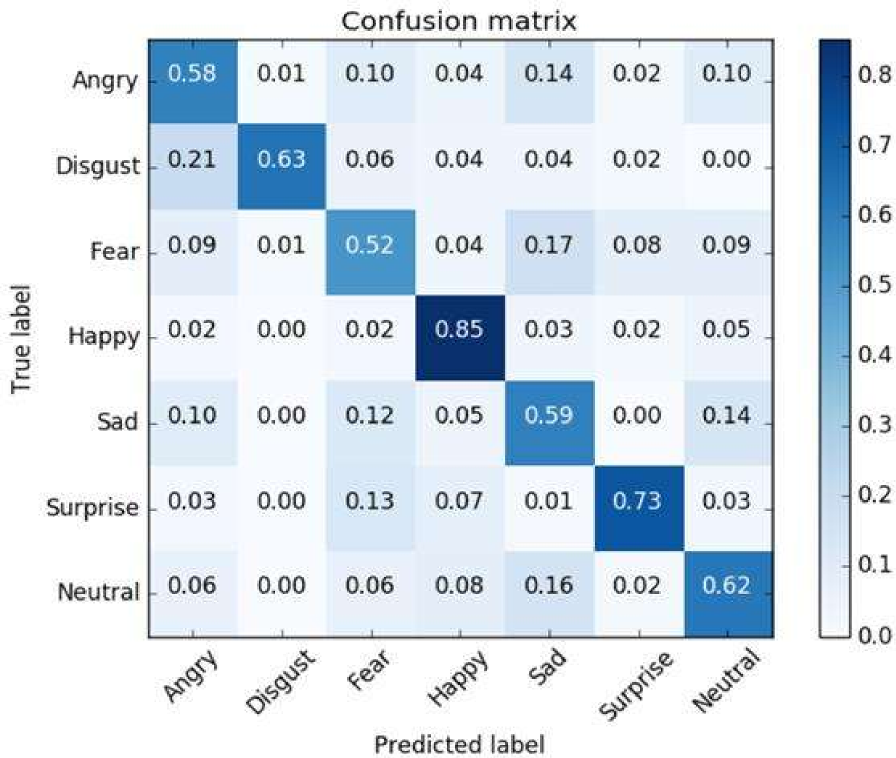

Figure 12: Confusion Matrix of proposed System-Visual Features

Confusion matrix of model represents:

True label (rows): The actual emotions present in the dataset (e.g., Angry, Disgust, etc.).
Predicted label (columns): The emotionspredicted by the classification model.Each cell contains the proportion of samples predicted for a given emotion compared to the true label.

Key observations:

1. Diagonal values0.58 for angry, 0.85 for Happy represents correct predictions. Higher values along the diagonal indicate better performance for that specific emotion. The model predicts "Happy" correctly with 85%and “Neutral" correctly 62% of the time so reaming emotions are negative or converge to negative emotions.

2. Off-diagonal values: represent misclassifications where the model incorrectly predicts a different emotion.eg. "Angry" is misclassified as "Fear" 10% of the time; "Neutral" is sometimes predicted as "Sad" (16%) etc.

3. Best and worst performance: The model performs best for "Happy" (85% accuracy on true labels) and struggles with "Fear" (52%) and confuses it with "Sad"(17%) and "Neutral" (9%).

4. Heat map intensity: Darker blue cells represent higher values, indicating stronger agreement.

To calculate precision, recall, and F1-score for each expression calculated with standard formulas. Each value is derived directly from the matrix:

1. TP: Diagonal value for each class (e.g., 0.58 for Angry).
2. FP: Sum of the column excluding the diagonal value.
3. FN: Sum of the row excluding the diagonal value.

Table 9: Model performance evaluation parameters (Expression class wise Precision, Recall, F1-Score)

Class	Precision	Recall	F1-Score
Angry	0.5321	0.5859	0.5577
Disgust	0.9692	0.6300	0.7636
Fear	0.5149	0.5200	0.5174
Happy	0.7265	0.8586	0.7870
Sad	0.5175	0.5900	0.5514
Surprise	0.8202	0.7300	0.7725
Neutral	0.6019	0.6200	0.6108

Observations:

1. Best Precision: Disgust (96.92%), meaning the model rarely predicts Disgust incorrectly.
2. Best Recall: Happy (85.86%), meaning the model captures most Happy instances accurately.
3. Best F1-Score: Happy (78.70%), showing a balance between precision and recall.

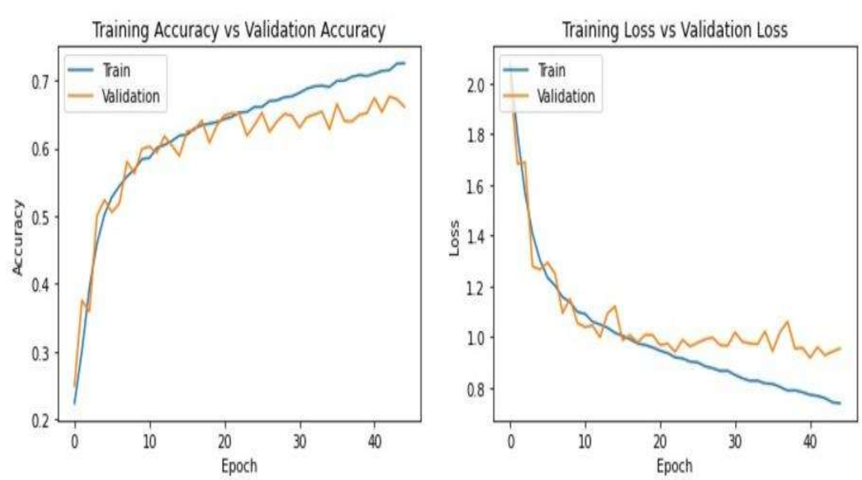

Figure 13: Accuracy vs Epoch & Loss vs Epoch confusion Matrix

Table10: Comparison of proposed method with existing approach performance

Ref No.	Year	Modality						fusion	Dataset	Accuracy
		Img	Audio	Text	EEG	SRI	colour			
[16]	2015	√	√					Decision level	Fer+tfd	47.67
[17]	2017	√	√					Decision level	Recola	76.00
[18]	2018	√	√	√				Decision level	Cmu-mosei	49.10
[19]	2018	√	√	√				Feature level	Iemocap	77.60
[20]	2020				√			Feature level	Deap	89.53
[21]	2020	√	√	√				Hybrid	Iemocap	73.98
[22]	2021	√	√		√			FL,DL,HL	Ravdess, Eeg	78.75
[23]	2022	√	√	√	√			Weighted decision level	Real-world collected dataset	67.00
Proposed Approach	**2025**	√				√	√	**Weighted decision level**	**Real-world collected dataset**	**79.00**

VII. Achievements with respect to objectives

Objective 1: Drive deeper into existing depression detection system.

Achievement: Paper - Publication -1 “Depression Detection System: A Systematic Review” is the outcome of this objective in which domain related research are examined.

Objective 2: Developed a user-friendly real-time, multi-modal depression detection system.

Achievement: Paper - Publication -2 “User Responsive Automatic Method for Real Time Depression Detection Using Deep Neural Network” is the outcome of this objective

The proposed depression detection system is user friendly, real-time, multi-model, user-responsive, reliableand confidential really a boon in health care domain in near future for pre-screening of depression symptoms compared to existing practices and its complex operating mechanism.

Tool effectively identified depression with response from Self Response Inventory (S.R.I.) quiz, Visual features and colour selection with almost 79% accuracy reliable and confidential result can be produced if tool is operated unbiased.

Objective 3: To validate the proposed models through performance parameters.

Achievement: Validation including confusion matrix, class-wise precision/recall/F1 values confirmed the robustness of model. Paper publication -2 is the outcome of objective 2 and 3.

Objective 4: Apply the tool in wild conditions to confirm its operability, reliability and effectiveness.

Achievement: The tool is performing effectively on real time camera taken picture frame and video sequences, demonstrating strong generalization and reliability in real-world scenarios.

Objective 5: Validate Automatic mood detection tool: Self Response Inventory

Achievement: Paper publication -3 is the outcome of this objective which effectively validate on 150 participants.

VIII. CONCLUSION

Depression is among leading cause of disability in each age group globally which affects health care system severely. Millions of affected persons require psychological and/or clinical treatments. It more affected to woman as before and after pregnancy changes in body. Teenagers are more emotional vulnerable and more connection to internet, social media, virtual world make situations worst. If negative state of mind / mental health issues/ depression can be detected in initial stage it can cure with little care else situation may become worst.

This research is dedicated to such pre diagnostic self-automated tool which can produced reliable result by performing fusion of available classical methods. This research is a novel approach which analyzes classical Self Response Inventory (SRI) and Visual features of person without intervention of other person so one can be more confidential and reliable result can be produced if tool is operated in an honest and truthful manner multi modal approaches give significant result such automated tool really a boon in health care domain in near future for pre-screening of depression symptoms.

To optimize the model's performance could be improved for emotions like Fear, Sad, and Neutral where misclassifications are significant further

(1) The participants' number could be increased performance of model will also improve.

(2) Quality of camera plays a major role and high-quality images contains visual features can retrieve specific information which further strengthen the visual behavioral analysis.

(3) Fusion of more modalities approaches like verbal features, social media text, EEG etc. could be improve the outcome.

The proposed system / tool effectively meets the established research objectives, creating a confidential,user-friendly,self -operated and reliable solution for depression detection.

IX. Publications:

(1) P. Acharya J, S.Shah M. Depression Detection System: A Systematic Review. J Neonatal Surg [Internet]. 2025 Jun. 21 [cited 2026 May 1];14(32S):1486-92. (Scopus)

Available from: https://www.jneonatalsurg .com/index.php/jns/article/view/7595/6605

(2) Acharya JP, Shah D. User Responsive Automatic Method for Real Time Depression Detection Using Deep Neural Network. J Neonatal Surg [Internet]. 2025 Jul. 8 [cited 2026 May 1];14(32S):4286-97. (Scopus)

Available from: https://www.jneonatalsurg.com/index.php/jns/article/view/8117/9194

(3)Acharya JP, Shah D.Automatic mood detection tool: Self Response Inventory.International Research Journal of Engineering and Technology.2026 Feb;13(02) : 821- 826

Available from: https://www.irjet.net/volume13-issue2, sr no:120

References

[1] 2021 Global Burden of Disease (GBD) [online database]. Seattle: Institute for Health Metrics and Evaluation; 2024 (https://vizhub.healthdata.org/gbd-results/, accessed 13 August 2025).

[2] Woody CA, Ferrari AJ, Siskind DJ, Whiteford HA, Harris MG. A systematic review and meta -regression of the prevalence and incidence of perinatal depression. J Affect Disord.2017;219:86–92.

[3] Evans-Lacko S, Aguilar-Gaxiola S, Al-Hamzawi A, et al. Socio-economic variations in the mental health treatment gap for people with anxiety, mood, and substance use disorders: results from the WHO World Mental Health (WMH) surveys. Psychol Med. 2018;48(9):1560-1571.

[4] P. Ekman, Handbook of Cognition and Emotion, JohnWiley& Sons Ltd, 2005, pp. 45-60

[5] H. Ellgring, R. J. Eiser, K. R. Scherer, "Non-verbal Communication in Depression," Cambridge University Press, 2007.

[6] American Psychiatric Association (5th edition, 2013). Diagnostic and statistical manual of mental disorders. Washington: DC.

[7] 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.

[8] Alghowinem S, Goecke R, Wagner M, Parker G, Breakspear M (2013). Eye movement analysis for depression detection. In IEEE International Conference on Image Processing, 4220–4224.

[9] Alghowinem S, Goecke R, Wagner M, Parker G, Breakspear M: Head pose and movement analysis as an indicator of depression. In Humaine Association Conference on

Affective Computing and Intelligent Interaction 2013:283–288.

[10] Jeffrey M Girard, Jeffrey F Cohn, Mohammad H Mahoor, S Mohammad Mavadati, Zakia Hammal, and Dean P Rosenwald, "Nonverbal social withdrawal in depression: Evidence from manual and automatic analyses," Image and vision computing, pp. 641-647, 2014

[11] Ooi KEB,Lech M,AllenNB(2013). Multi-channel weighted speech classification system for prediction of major depression in adolescents. IEEE Transactions on Biomedical Engineering, 60:497–506.

[12] Cummins N, SethuV, EppsJ, Krajewski J (2014). Probabilistic acoustic volume analysis for speech affected by depression. In Annual Conference of the International Speech Communication Association,1238–1242.

[13] Williamson JR, Street W, Quatieri TF, Helfer BS, Ciccarelli G, Mehta DD: Vocal and Facial Biomarkers of Depression Based on Motor Incoordination and Timing. In Proceedings of the 4th International Workshop on Audio/Visual Emotion Challenge 2014:65–72.

[14] Espinosa HP, Escalante HJ, Villasen˜orpineda L, Montes-y g´omez M, Pinto-Avedan˜oD, Reyes-Meza V: Fusing affective dimensions and audio-visual features from segmented video for depression recognition. In ACM International Workshop on Audio/Visual Emotion Challenge 2014.

[15] Gupta R, Malandrakis N, Xiao B, Guha T, Van Segbroeck M, Black MP, Potamianos A, Narayanan SS: Multimodal prediction of affective dimensions and depression in human - computer interactions. In International Audio/Visual Emotion Challenge and Workshop 2014

[16] Senoussaoui M, Sarria-Paja M, Santos JaF, Falk TH: Model Fusion for Multimodal Depression Classification and Level Detection. In Proceedings of the 4th International Workshop on Audio/Visual Emotion Challenge 2014:57–63.

[17] Pampouchidou, Anastasia, Kostas Marias, ManolisTsiknakis, P.Simos, Fan Yang, and Fabrice Meriaudeau (2015). 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.

[18] Z. Hammal, J. F. Cohn, C. Heike, and M. L. Speltz. What Can Head and Facial Movements Convey about Positive and Negative Affect? In Affective Computing and Intelligent int.2015.

[19] S. E. Kahou, X. Bouthillier, P. Lamblin, C. Gulcehre, V. Michalski, K. Konda, S. Jean, P.Froumenty, Y. Dauphin, N. Boulanger-Lewandowski, R. Chandias Ferrari, M. Mirza, D. Warde-Farley, A. Courville, P. Vincent, R. Memisevic, C. Pal, and Y. Bengio, “Emonets: Multimodal deep learning approaches for emotion recognition in video,” Journal on Multimodal User Interfaces, vol. 10, no. 2, pp. 99–111, 2015.

[20] 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.

[21] Morales, M., Scherer, S., & Levitan, R, "A Cross-modal Review of Indicators for Depression Detection Systems," Proceedings of the Fourth Workshop on Computational Linguistics and Clinical Psychology –- From Linguistic Signal to Clinical Reality, 2017

[22] P. Tzirakis, G. Trigeorgis, M. A. Nicolaou, B. W. Schuller, and S. Zafeiriou, "End to-End Multimodal Emotion Recognition Using Deep Neural Networks," IEEE Journal of Selected Topics in Signal Processing, vol. 11, no. 8, pp. 1301–1309, Dec. 2017, doi:10.1109 /jstsp.2017.

[23] S. Sahay, S. H. Kumar, R. Xia, J. Huang, and L. Nachman, "Multimodal relational tensor network for sentiment and emotion classification," Proceedings of Grand Challenge and Workshop on Human Multimodal Language (Challenge-HML), 2018.

[24] Dr. D. Venkataraman and Namboodiri Sandhya Parameswaran: "Extraction of Facial Features for Depression Detection among Students" International Journal of Pure and Applied Mathematics Volume 118 No. 7 2018, 455-463.

[25]D. Hazarika, S. Poria, A. Zadeh, E. Cambria, L.-P. Morency, and R. Zimmermann,"Conversational memory network for emotion recognition in dyadic dialogue videos," Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers), 2018.

[26] A. Pampouchidou, P.G. Simos, K. Marias, F. Meriaudeau, F. Yang, M. Pediaditis, M. Tsiknakis,"Automatic Assessment of Depression Based on Visual Cues: A Systematic Review," IEEE Transactions on Affective Computing, vol. 10, no. 4, pp. 445-470, 2019.

[27] M. M. Hassan, Md. G. R. Alam, Md. Z. Uddin, S. Huda, A. Almogren, and G. Fortino,"Human emotion recognition using deep belief network architecture," Information Fusion, vol.10.1016/j.inffus.2018.10.009. 51, pp. 10–18, Nov. 2019

[28] D. Priyasad, T. Fernando, S. Denman, S. Sridharan, and C. Fookes, "Attention driven fusion for multi-modal emotion recognition," ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2020.

[29] J. Njoku, Angela C. Caliwag, W. Lim, S. Kim, H.-J. Hwang, and J.-W. Jeong, "Deep Learning Based Data Fusion Methods for Multimodal Emotion Recognition," The Journal of Korean Institute of Communications and Information Sciences, vol. 47, no. 1, pp. 79–87, Jan. 2022, doi: 10.7840/kics.2022.47.1.79.

[30] Komal D. Anadkat,Dr. Hiteishi M. Diwanji& Dr. Shahid Modasiya "Effect of Preprocessing in Human Emotion Analysis Using Social Media Status Dataset" Reliability: Theory and Application "ISSN 1932-2321, Volume 17 Issue 1 PP-67 March 2022 (SCOPUS).

[31] Jonauskaite, D., Abu-Akel, A., Dael, N., Oberfeld, D., Abdel-Khalek, A. M., Al-Rasheed, A. S., Antonietti, J.-P., Bogushevskaya, V., Chamseddine, A., Chkonia, E., Corona, V., Fonseca-Pedrero, E., Griber, Y. A., Grimshaw, G., Hasan, A. A., Havelka, J., Hirnstein, M., Karlsson, B. S. A., Laurent, E., … Mohr, C. (2020). Universal Patterns in Color-Emotion Associations Are Further Shaped by Linguistic and Geographic Proximity. Psychological

Science, 31(10), 1245-1260.

[32] Yang J, Shen X. The Application of Color Psychology in Community Health Environment Design. J Environ Public Health. 2022 Aug 25; 2022:7259595. doi: 10.1155/2022/7259595. Retraction in: J Environ Public Health. 2023 Oct 18;2023:9786523. doi: 10.1155/2023/9786523. PMID: 36060888; PMCID: PMC9436572

[33] Hamilton M. A rating scale for depression (1960).Journal of neurology, neurosurgery, and psychiatry, 23 (1):56.

[34]BeckAT,SteerRA,BallR,RanieriWF. (1996). Comparison of beck depression inventories-ia and-iii n psychiatric out patients. Journal of personality assessment, 67(3):588–597.

[35] Patient Health Questionnaire-9, Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (1999). Patient Health Questionnaire-9 (PHQ-9) [Database record]. APA SysTest.

[36] Levis B, Benedetti A, Thombs B D. Accuracy of Patient Health Questionnaire-9 (PHQ-9) for screening to detect major depression: individual participant data meta-analysis BMJ 2019; 365: l1476 doi:10.1136/bmj. l1476

[37] Zung,WW(1965)Aself-ratingdepressionscale.ArchGenPsychiatry12,63-70.

[38] Radloff, L.S. (1977). The CES-D Scale: a self-report depression scale for research in the general population. Applied Psychological Measurement, 1:385-401.

[39] Whooley, M.A., Avins, A.L., Miranda, J. et al. Case-finding instruments for depression. J GEN INTERN MED 12, 439–445 (1997). https://doi.org/10.1046/j.1525-1497.1997.00076.x

[40]Rush A, Trivedi M, Ibrahim H et al.The 16-Item quick inventory of depressive symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression, Biological Psychiatry, 54, 573-583

[41] Yang, S., Cui, L., Wang, L., Wang, T., & You, J. (2024). Enhancing multimodal depression diagnosis through representation learning and knowledge transfer. Heliyon, 10(4), e25959. https://doi.org/10.1016/j.heliyon.2024.e25959

[42] Internet resources like: doc player.net,journalofbigdata., springeropen.com, deepai.org,arxiv.org, pubmed.ncbi.nlm.nih.gov. etc.

[43] https://www.verywellmind.com/colour-psychology-2795824.

[44] R. Verma, “FER2013,” Kaggle,26-May-2018.