Journal of Neonatal Surgery
ISSN(Online): 2226-0439
Vol. 14, Issue 32s (2025)
https://www.jneonatalsurg.com

Depression Detection System: A Systematic Review

Jiger P. Acharya[1], Dr. Milind S.Shah[2]

[1] Senior Lecturer Computer Dept., Government Polytechnic –Ahmadabad, Gujarat, India

Email ID: jigeracharya@gmail.com

[2] Professor Electronics & Communication Dept., S.S.E.C-Bhavnagar, Gujarat, India

Email ID: mssvgec@gmail.com

Cite this paper as: Jiger P. Acharya, Dr. Milind S.Shah, (2025) Depression Detection System: A Systematic Review. *Journal of Neonatal Surgery*, 14 (32s), 1486-1492.

ABSTRACT

Depression is mood disorders which result in severe disabling conditions affect person's ability to cope with routine life challenges. It may occur when person remain more than two weeks in negative state of mind continuously. Depending on severity depression is classified as mild, moderate and severe. The World health organization (WHO) list depression as major cause of suffering and disability worldwide more than 350 million people are affected and predict to be leading cause in 2020[1],[2]. Psychosocial and clinical treatments are available but persons have tendency to conceal it. Depression has observable behavioral symptoms related to affective and psychomotor domains which can be identified by human or machine. Classical approaches concerned to person's behavioral analysis and family observations during clinical interviews which are effective if it can be explicitly defined and precisely assessed while automatic system perform the said task effectively and open a new era for health care domain.

Keywords: *Audio Video Emotion Challenge (AVEC), Visual features, Vocal features, Diagnostic and Statistical Manual of mental disorders (DSM), Deep Neural Network (DNN) etc.*

1. INTRODUCTION

Depression is leading cause of disability globally in all age groups which severely affects the global health index. 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 factors, psychological, social factors, financial factors, drug addictions change in genetic features, brain's neurotransmitter levels, accidental events like head major head injury and many more. Early detection plays a crucial role as by positive counseling, various physical and mental exercises like meditation and proper medication may resolved depression else situations become worst up to suicidal tendency. Depression can be detected through various behavioral symptoms and during clinical interview by medical expert but the rate of the affected person reach at right time with right medical professional is very low so many researchers from physiological, pattern matching, computer vision have tried to detect depression features from behavior verbal and nonverbal channels with help of various of machine learning algorithms under Artificial Intelligence domain. Neural Network has shown shows great potential to resolve this burning issue of human health care domain to serve mankind.

2. METHODS

Human express their emotions through multiple channels which can be monitored for emotion analysis.[3],[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 as mentioned below.

In clinical 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 mental health issues and because of social stigma they have tendency to conceal such issues so approximately 10% actually affected get benefits of this approach.

In Visual channel in computer vision, pattern matching and image community detect depression features from behavior visual channels. Wager et al. [7] suggested 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. [9] prove that during depression verbal and nonverbal channels are affected significantly which can be identified during external observations.

In Audio based approach Ooi et al. [10] 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 [10] [11].

In Colour psychology-based approach reveals how various colours impact person's mood and so behaviors. It enlightens colour impact on person's emotional response based on age and cultural background etc. [12]. 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. In 2020 study that surveyed the emotional associations of 4598 persons from 30 countries and conclude that people commonly associate certain colours with specific emotions as below:

Table 1: Colour association with Emotion and Mood

Colour	Percentage	Emotions	Mood Category
White	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

Self –Response Inventory (S.R.I): Quiz based approach in which person is required to respond quiz based on the best knowledge of 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 [13].

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 [14].

Patient Health Questionnaires (PHQs): The PHQ-9 has nine questions about person's mood and daily activities like appetite, watching Television which helps Doctor come to a depression diagnosis [15]. The PHQ-2 asks the first two questions of the PHQ-9.Which is a screening test guide doctor for further processing.

Zung Self –Rating Depression Scale: It is again short surveys which identify depression levels [16].

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 [17].

In EEG based approach physiological signals receive from brain are measured which are real time in nature and cannot suppress or hide like facial expressions, audio and textual emotions so it is more authentic and precise only drawback is the cost and complexity associate with this approach [18].

In social media related approach data in which Text data in various levels and categorize with positive or negative broadly and much specific like emphatic, glad etc. databases for textual emotion analysis [18].

Many contributors used fusion of above in specific group called multi-modal data in which persons reflect their feelings through multiple channels [19], [20], [21]. Which are more accurate and reliable than focusing only on a single modal like audio, video, text, or EEG etc. [18],[22].Major research of said domain is as mentioned.

Table 2: Authors contribution of domain related research

Sr	Authors	Type	Dataset
1	Becker et al. (1994)	Audio/Video	Dementia Bank Database Reddit Self-reported
2	Valstaret 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	Valstaretal. (2014)	Audio/Video	AVEC 2014
6	Gratch J et al. (2014)	Audio/Video	The Distress Analysis Interview Corpus
7	Valstaretal. (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	Anadkat et al. (2022)	Audio/Video/Social media text/EEG	FER2013, Ravdess, Kaggle's social media dataset, and EEG dataset

3. RESULTS/KEY OBSERVATIONS:

Visual as major channel for detection:

Face is the major 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 [3] categorize standard expression of six fundamental feelings along with Head orientation, posture and movement majorly use to recognize depression [7],[8],[24], [25].Girard et al [20] explore strong co-relation between visual direction and depression.

Table 3: 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

Below table indicate performances for various approaches

Table 4: Comparison of proposed methods with existing approach performance evaluation

Ref No.	Year	Modality						fusion	Dataset	Accuracy
		Image	Audio	Text	EEG	SRI	Colour			
[29]	2015	√	√					Decision level	Fer+tfd	47.67
[30]	2017	√	√					Decision level	Recola	76.00
[31]	2018	√	√	√				Decision level	Cmu-mosei	49.10
[32]	2018	√	√	√				Feature level	Iemocap	77.60
[33]	2020				√			Feature level	Deap	89.53
[34]	2020	√	√	√				Hybrid	Iemocap	73.98
[35]	2022	√	√	√	√			Weighted decision	Combined customized	93.00
[35]	2022	√	√	√	√			Weighted decision level	Real-world collected dataset	67.00

4. DISCUSSION

Human are expressive through verbal and nonverbal communication channels both channels are important to express feelings [3], [4]. Person can remain silent to suppress verbal communication and by fake expression or conceal expression 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 [22],[23],[24],[25]. 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 [22].

- 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.
- Variations in 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.

1. Deliberate versus spontaneous expression: some time execution of spontaneous expression differs from deliberate expression.
2. Head orientation and scene complexity: Head is major region where expression execution happen so extract region

of Interest (ROI).

3. Neural Network limitation to train and execute Neural Network model we required large and variety amount of data which is a big challenge.
4. 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.
5. Emotional classes versus data classes in data labeling: Mapping of emotions classes and standard labeled data are big challenges.
6. Over fitting versus generalization: Balancing these two events is a quite difficult for any Neural Network.
7. Diagnostic criteria, clinical instruments, and performance metrics which have wide variety globally
8. Clinical specificity as many psychiatric and medical disorders-occur.
9. Temporal dynamics important in psychomotor retardation
10. Diversified large samples cultures, ethnicities, ages, and genders etc.

Depression is among leading cause of disability in each age group globally and affects health care system severely. Millions of affected and likely to be affected persons require psychologicaland/orclinicaltreatments.Itmoreaffectedtowomanasbeforeandafterpregnancy their hormones and body get changed. Teenagers are more emotional vulnerable and more connection to internet, social media, virtual world make situations worst. If depression can be detected in initial stage it can cure with little care else situation may become worst.

To all who have directly or indirectly motivated us to do little contribution to society by developing automated depression detection tool which effectively identified depression via a small set of quiz while attempting psychology related questionnaires person's mental health is estimated and while attempting quiz person's video is captured so visual and/ or vocal features are analyzed and matched with depression features if all result are inline Person is advice to take care of mental health visit for clinical advice for further treatment.

REFERENCES

[1] World Health Organization. The world health report (2003).shaping the future.

[2] Mathers C., BoermaJ, Fat D.(2008).The Global Burden of Disease. Update Geneva, Switzerland: WHO.

[3] P. Ekman, Handbook of Cognition and Emotion, John Wiley& Sons Ltd, 2005, pp. 45-60.

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

[5] YangY, FairbairnCE, CohnJF (2013).Detectingdepressionseverityfromvocal prosody. IEEE Transactions on Affective Computing,4:142–150.

[6] MH, Ibrahim HM, Carmody TJ, Arnow B, Klein DN, Markowitz JC, Ninan PT, K ornstein S, Manber R, et al. (2003). 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(5):573–583.

[7] Alghowinem S, Goecke R, Wagner M, Parker G, Breakspear M (2013). Head pose and movement analysis as an indicator of depression. In Humaine Association Conference on Affective Computing and Intelligent Interaction,283–288.

[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] Girard JM, Cohn JF, Mahoor MH, Mavadati SM, Hammal Z, Rosenwald DP (2014). Nonverbal social withdrawal in depression: Evidence from manual and automatic analyses.Image and Vision Computing, 32:641–647.

[10] Ooi KEB, Lech M, Allen NB (2013). Multichannel weighted speech classification system for prediction of major depression in adolescents. IEEE Transactions on Biomedical Engineering,60:497–506.

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

[12] Internet resource: https://www.verywellmind.com/colour-psychology-2795824

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

[14] Beck AT, Steer RA, Ball R, Ranieri WF. (1996). Comparison of beck depression inventories-ia and-ii in psychiatric outpatients. Journal of personality assessment, 67(3):588–597.

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

[16] Zung, WW (1965) A self-rating depression scale. Arch Gen Psychiatry 12, 63-70

[17] 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.

[18] Komal D. Anadkat & Dr. Hiteishi M. Diwanji "Effectof Activation Function in Speech Emotion Recognition On The Ravdess Dataset" Reliability: Theory and Application "ISSN 1932-2321, Volume 16 Issue 3 PP-63 September 2021 (SCOPUS Indexed).

[19] N.Cummins, J.Joshi, A.Dhall, V.Sethu, R.Goecke,J .Epps (2013, October). Diagnosis of depression by behavioural signals: a multimodal approach. AVEC'13: Proceeding of the 3rdACM international workshop on Audio/Visual Emotion Challenges.

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

[21] 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.

[22] 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.

[23] Kaya H, Salah AA (2013). Eyes whisper depression. In ACM International Workshop on Audio/Visual Emotion Challenge,1–4.

[24] Y. Zhou, S. Scherer, D. Devault, J. Gratch, G. Stratou, L.-P. Morency, and J. Cassell, "Multimodal Prediction of Psychological Disorders: Learning Verbal and Nonverbal Commonalities in Adjacency Pairs," 17thWorkshop on the Semantics and Pragmatics of Dialogue. SEMDIAL, pp. 160-169, 2013.

[25] SchererS, StratouG,LucasG,MahmoudM,BobergJ,GratchJ,Rizzo AS, Morency LP (2014). Automatic audiovisual behavior descriptors for psychological disorder analysis. Image andVisionComputing, 32:648–658.

[26] Lopez-Otero P, Docio-Fernandez L, Garcia-Mateo C (2014). A study of acoustic features for the classification of depressed speech. Proceedings of the International Convention on Information and Communication Technology, Electronics and Microelectronics,1331–1335.

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

[28] A Summarization of the Visual Depression Databases for Depression Detection978-1-7281-7596-6/20 ©2020 IEEE explore

[29] 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.

[30] 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.2017.2764438.

[31] 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.

[32] 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.

[33] S. Siriwardhana, A. Reis, R. Weerasekera, and S. Nanayakkara, "Jointly fine-tuning 'bert-like' self supervised models to improve multimodal speech emotion recognition," Interspeech 2020, 2020.

[34] 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.

[35] 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 Indexed).

[36] A. Ashraf, T. S. Gunawan, F. D. A. Rahman, M. Kartiwi, N. Ismail and Ulfiah, "A Summarization of the Visual Depression Databases for Depression Detection," 2020 6th International Conference on Wireless and Telematics (ICWT), Yogyakarta, Indonesia, 2020, pp. 1-6, doi: 10.1109/ICWT50448.2020.9243625.

[37] Internet resources like: doc player.net, journal of bigdata., springeropen.com, deepai.org, arxiv.org ,pubmed.ncbi.nlm.nih.gov