Archives available at journals.mriindia.com

International Journal of Advanced Scientific Research and Engineering Trends

ISSN: 2456-0774

Volume 10 Issue 05, 2026

A Review of the Integration of Machine Learning Techniques for the Detection of Depression

[1]N. N. Patil, [2]K. G. Kharade, [3]V. S. Kumbhar

[1] MCA Department, KIT'S IMER, Kolhapur, Maharashtra, India

[2,3]Department of Computer Science, Shivaji University, Kolhapur, Maharashtra, India.

Email: [1]navni48@gmail.com, [2]kgk_csd@unishivaji.ac.in, [3]vsk_csd@unishivaji.ac.in

Peer Review Information

Submission: 15 April 2026

Revision: 27 April 2026

Acceptance: 09 May 2026

Keywords

Artificial Intelligence, Depression Detection, Deep Learning, Machine Learning, Mental Health Computing, Multimodal Analysis

Abstract

Depression is a widespread mental health disorder that significantly impacts emotional well-being, behavior, and daily functioning. It is characterized by persistent sadness, loss of interest, cognitive impairment, and reduced productivity, making it a major public health concern. Early identification of depression is critical for preventing severe psychological, social, and economic consequences, including increased suicide risk and long-term disability. Traditional diagnostic approaches rely primarily on clinical interviews and self-report questionnaires, which are often subjective, time-consuming, and dependent on expert interpretation. Recent advances in machine learning (ML) have enabled the development of automated systems capable of detecting depressive symptoms through the analysis of diverse data sources, including textual content, speech signals, facial expressions, physiological measurements, and behavioral patterns. This paper presents a comprehensive review of machine learning techniques applied to depression detection. It examines commonly used data modalities, feature extraction strategies, traditional machine learning models, deep learning architectures, multimodal fusion approaches, and evaluation methodologies. A comparative analysis of existing research studies is provided to highlight performance trends and methodological differences. The review indicates that deep learning models and multimodal frameworks generally achieve superior detection performance; however, ethical, privacy, generalizability, and interpretability challenges remain significant barriers to real-world deployment.

Introduction

Depression is one of the most prevalent mental health disorders worldwide and is recognized as a leading cause of global disease burden [1]. It affects individuals across different age groups, cultures, and socioeconomic backgrounds, often resulting in impaired social functioning and reduced quality of life. Despite its high prevalence, depression frequently remains underdiagnosed due to social stigma, limited access to mental health professionals, and reliance on subjective self-assessment tools.

Early detection plays a crucial role in improving treatment outcomes and reducing long-term consequences. Conventional diagnostic procedures typically involve clinical interviews and standardized questionnaires such as the Beck Depression Inventory (BDI) and the Patient Health Questionnaire (PHQ-9) [2], [3]. Although effective, these methods depend

heavily on patient honesty and clinician expertise, which may introduce bias.
Recent advancements in machine learning (ML) and artificial intelligence (AI) provide promising alternatives by enabling objective, scalable, and data-driven mental health assessment systems [12], [14]. This review aims to systematically analyze existing ML-based depression detection approaches, summarize current trends, and identify key challenges and future research opportunities relevant to IEEE conference audiences.

Data Modalities

Machine learning systems for depression detection utilize a variety of data modalities, each capturing different aspects of human behavior and emotional state. Textual data obtained from social media platforms, online forums, and clinical transcripts has been extensively studied due to its accessibility and strong correlation with emotional expression [4], [14]. Linguistic patterns, sentiment polarity, and semantic content provide valuable indicators of depressive symptoms.
Speech and audio signals represent another important modality. Variations in pitch, tone, speech rate, and pause duration have been shown to reflect psychological conditions [6]. Visual data, including facial expressions and body movements, captures non-verbal cues such as reduced facial activity and limited eye contact [7].
Physiological and behavioral data collected from wearable devices and smartphones enable continuous and passive monitoring of sleep patterns, activity levels, and social interaction, offering insights into long-term behavioral changes [11].

Feature Extraction Techniques

Feature extraction is a critical step in transforming raw data into meaningful representations suitable for machine learning models. Traditional approaches rely on handcrafted features derived from domain expertise, such as Linguistic Inquiry and Word Count (LIWC) features for text analysis [5], Mel-frequency cepstral coefficients (MFCCs) for speech signals [6], and facial landmarks for video-based analysis [7]. These features are often interpretable but may not fully capture complex patterns.
Deep learning-based feature extraction methods automatically learn hierarchical representations directly from raw data. Convolutional Neural Networks (CNNs) extract spatial features from images and spectrograms, while Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks model temporal dependencies in sequential data [12]. Transformer architectures further enhance feature learning by capturing long-range contextual relationships in textual data [9], [10].

Machine Learning Models

Traditional machine learning models such as Support Vector Machines (SVM) [15], Logistic Regression, Decision Trees, Random Forests, and K-Nearest Neighbors (KNN) have been widely used for depression detection, particularly when datasets are limited in size [4], [6]. These models are computationally efficient and relatively interpretable, making them suitable for baseline analysis.
Deep learning models have demonstrated superior performance in recent studies due to their ability to handle high-dimensional and unstructured data [12]. CNNs are commonly applied to visual and speech-based tasks, while RNNs and LSTMs are effective for modeling temporal patterns [8]. Transformer-based models such as BERT have achieved state-of-the-art results in text-based depression detection by capturing contextual semantics [9].

Comparative Analysis

Comparative analysis of existing studies provides valuable insights into the effectiveness of different machine learning approaches. Table I summarizes representative research works, highlighting the datasets used, data modalities, models employed, and reported performance metrics [4], [6]–[9], [11].
Results indicate that deep learning models generally outperform traditional approaches, particularly when multimodal data is integrated [8], [11]. Multimodal fusion techniques combine complementary information from multiple sources, leading to improved robustness and accuracy. However, performance variations across datasets suggest challenges related to generalization, dataset bias, and reproducibility.

Challenges And Ethical Considerations

Despite promising results, several challenges hinder the practical deployment of ML-based depression detection systems. Limited availability of large, labeled datasets restricts model training and validation [7]. Demographic and cultural bias in training data may lead to unfair predictions, while deep learning models often lack interpretability, reducing trust among clinicians [13].
Ethical considerations play a critical role in mental health applications. Privacy protection, informed consent, and data security must be prioritized when handling sensitive personal

data [14]. ML-based systems should be designed as decision-support tools to assist clinicians rather than replace professional diagnosis.

Conclusion

This paper presented a comprehensive review of machine learning techniques for depression detection, covering data modalities, feature extraction methods, model architectures, and evaluation strategies. The review highlights the growing importance of deep learning and multimodal approaches in achieving improved detection performance [8], [11]. Future research should focus on developing explainable and fair models [13], improving cross-domain generalization, and integrating ML-based systems into clinical workflows. Addressing ethical and practical challenges will be essential for the responsible adoption of automated depression detection technologies.

Table 1: Summary of Machine Learning Techniques for Depression Detection

Author	**Dataset**	**Modality**	**Model**	**Performance**
Coppersmith [4]	Twitter	Text	SVM	72% Acc
Cummins [6]	AVEC	Speech	GMM+SVM	F1=0.68
Valstar [7]	DAIC-WOZ	Audio/Video	RF	75% Acc
Al Hanai [8]	DAIC-WOZ	Audio+Text	CNN+LSTM	AUC=0.77
Huang [11]	AVEC	Multimodal	CNN+Attention	82% Acc
Devlin [9]	Reddit	Text	BERT	F1=0.84

References

World Health Organization, Depression and Other Common Mental Disorders: Global Health Estimates, Geneva, Switzerland: WHO Press, 2017.

A. T. Beck, C. H. Ward, M. Mendelson, J. Mock, and J. Erbaugh, “An inventory for measuring depression,” Archives of General Psychiatry, vol. 4, no. 6, pp. 561–571, 1961.

K. Kroenke, R. L. Spitzer, and J. B. W. Williams, “The PHQ-9: Validity of a brief depression severity measure,” Journal of General Internal Medicine, vol. 16, no. 9, pp. 606–613, 2001.

G. Coppersmith, C. Harman, and M. Dredze, “Measuring post traumatic stress disorder in Twitter,” in Proc. 8th Int. AAAI Conf. Weblogs and Social Media (ICWSM), 2014, pp. 579–582.

J. W. Pennebaker, R. L. Boyd, K. Jordan, and K. Blackburn, “The development and psychometric properties of LIWC2015,” Behavior Research Methods, vol. 47, no. 1, pp. 1–18, 2015.

N. Cummins, J. Epps, M. Breakspear, and R. Goecke, “An investigation of depressed speech detection: Features and normalization,” IEEE Transactions on Affective Computing, vol. 6, no. 3, pp. 231–244, Jul.–Sep. 2015.

M. Valstar[11], “AVEC 2014: 3D dimensional affect and depression recognition challenge,” in Proc. ACM Int. Conf. Multimodal Interaction (ICMI), 2014, pp. 3–10.

A. Al Hanai, M. M. Haque, and M. S. Ghanem, “Detecting depression with audio/text sequence modeling of interviews,” in Proc. Conf. Empirical Methods in Natural Language Processing (EMNLP), 2018, pp. 1716–1726.

J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of deep bidirectional transformers for language understanding,” in Proc. North American Chapter of the Association for Computational Linguistics (NAACL-HLT), 2019, pp. 4171–4186.

Z. Yang, D. Yang, C. Dyer, X. He, A. Smola, and E. Hovy, “Hierarchical attention networks for document classification,” in Proc. NAACL-HLT, 2016, pp. 1480–1489.

Y. Huang, J. Liu, and X. Zhang, “Multimodal depression detection based on audio, visual, and textual information,” IEEE Access, vol. 8, pp. 173196–173207, 2020.

I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.

T. Van Steenkiste, A. Rademaker, J. J. T. J. De Mey, and S. Verbeke, “Explainable artificial intelligence for mental health: A survey,” IEEE Journal of Biomedical and Health Informatics, vol. 24, no. 11, pp. 3287–3298, Nov. 2020.

R. A. Calvo, D. Milne, M. S. Hussain, and H. Christensen, “Natural language processing in mental health applications using non-clinical texts,” IEEE Intelligent Systems, vol. 34, no. 2, pp. 46–56, Mar.–Apr. 2019.

S. R. Gunn, “Support vector machines for classification and regression,” Univ. of Southampton, Southampton, U.K., Tech. Rep., 1998.