ICCK Transactions on Emerging Topics in Artificial Intelligence
http://dx.doi.org/10.62762/TETAI.2025.416797

REVIEW ARTICLE

Advances in Artificial Intelligence-Based Depression Diagnosis: A Systematic Review

Liangguo Wang[1] **and Jiaqian Wu**[2,*]

[1] Xinjiang Hetian College, Xinjiang 848000, China
[2] Beijing Anding Hospital, Capital Medical University, Beijing 100088, China

Abstract

This study systematically reviews the current status and recent advances in intelligent depression detection, aiming to provide insights for applying artificial intelligence in mental health. Using a systematic review approach, we analyze detection methods based on multiple data types including voice, facial expressions, body signals, and social media texts, while examining how algorithms have evolved from traditional machine learning to deep learning. Results show that AI technology has clear benefits in improving detection accuracy, reducing costs, and enabling early warning systems. Current research still faces important challenges in data collection, technical reliability, clinical use, and privacy concerns. Future work should focus on combining knowledge from different fields, implementing systems in clinical practice, and developing standards for wider adoption.

Keywords: depression, artificial intelligence, multimodal features, deep learning, intelligent detection.

Submitted: 22 May 2025
Accepted: 18 July 2025
Published: 27 August 2025

Vol. 2, **No.** 3, 2025.
10.62762/TETAI.2025.416797

***Corresponding author:**
Jiaqian Wu
1310908555@qq.com

1 Introduction

Depression represents one of the most prevalent mental health disorders worldwide, with far-reaching implications for global public health. In recent years, artificial intelligence (AI) has emerged as a promising approach for enhancing depression identification and diagnosis. Modern AI-based detection systems analyze multiple types of data simultaneously, examining vocal characteristics, facial expressions, and physiological measurements to create more objective diagnostic tools [1, 2].

The global impact of depression is substantial and growing. According to the World Health Organization (WHO), approximately 350 million people suffer from depression globally, with incidence rates showing a consistent increase [3]. This situation worsened during the COVID-19 pandemic, which triggered a 25% rise in depression cases worldwide during 2020 [4]. In China specifically, studies report a lifetime prevalence rate of 6.8%, leading to significant social and economic burdens [5]. Current global health assessments consistently rank depression among the leading contributors to worldwide disease burden [6].

Traditional depression diagnosis relies heavily on clinical judgment and standardized questionnaires. While these approaches remain essential, they face

Citation
Wang, L., & Wu, J. (2025). Advances in Artificial Intelligence-Based Depression Diagnosis: A Systematic Review. *ICCK Transactions on Emerging Topics in Artificial Intelligence*, 2(3), 148–156.

important limitations: (1) inconsistent diagnoses between different clinicians, (2) lengthy assessment procedures, and (3) difficulties implementing these methods at scale, especially in settings with limited resources [7]. AI technologies help address these challenges by reducing subjective elements in diagnosis and improving overall efficiency [8].

The application of AI in depression diagnosis has led to notable improvements in both early detection capabilities and diagnostic accuracy [9]. Recent technological progress has enabled significant advances in depression detection through the analysis of multiple data types [10, 11], including:

- Voice analysis technologies that identify depression markers in speech
- Systems that automatically detect depression indicators in facial expressions
- Methods for processing biological and physiological signals
- Techniques that analyze social media content for signs of depression

This systematic review synthesizes current research findings and develops a structured framework for effectively implementing these intelligent detection technologies in real-world clinical settings.

2 Multimodal Feature Analysis of Intelligent Depression Detection

The clinical manifestations of depression are complex and diverse, including changes in multiple dimensions such as emotions, cognition, behavior, and physiology. Based on artificial intelligence, depression detection technology can capture the characteristic manifestations of depression from different perspectives by analyzing multimodal data such as patients' voice, facial expressions, physiological signals, and social media behaviors. Studies have shown that the comprehensive analysis of multimodal features can provide more comprehensive and reliable detection results [16].

2.1 Voice Feature Analysis

Voice feature analysis is one of the earliest and most mature research directions in intelligent depression detection. Cummins et al. [19] systematically summarized the application of voice analysis in depression assessment and found that the voice of depressed patients has obvious changes in prosodic features, including a decrease in fundamental frequency (F0), a weakening of voice energy, and a slowdown in speech rate. These features can be objectively quantified through digital signal processing technology. Low et al. [10] demonstrated how to use natural language processing technology to analyze voice features and verified that these features have a significant correlation with the severity of depressive symptoms. In addition to basic acoustic features, more complex features are also included.

2.2 Facial Expression Analysis

Facial expressions are important external manifestations of emotional states. Depressed patients often exhibit reduced facial expressions and dull emotional expressions. The facial expression analysis system developed by Ringeval et al. [12] can track the movement of key facial feature points in real-time. By analyzing the spatio-temporal change patterns of these feature points, the facial expression features of depressed patients can be effectively identified. The system proposed by Ringeval et al. [12] can not only track the movement of facial key points but also analyze micro-expression changes and emotional transition processes. This study found that depressed patients have specific features in the dynamic change patterns of facial expressions, such as delayed expression transitions and the duration of emotional expressions. Facial expression analysis can also be combined with temporal information for dynamic modeling.

2.3 Physiological Signal Analysis

Physiological signal analysis provides an important basis for the objective assessment of depression. Cai et al. [14] demonstrated a depression detection method based on electroencephalogram (EEG) signals and verified the value of physiological signal analysis in depression diagnosis. Jacobson et al. [15] explored the application of digital biomarkers in emotional disorder assessment and demonstrated the advantages of wearable devices in continuously monitoring physiological indicators. These studies provide new technical means for the dynamic assessment of depressive states.

2.4 Social Media Behavior Analysis

Social media behavior analysis provides a new research perspective for depression detection. De Choudhury et al. [11] developed a depression prediction model based on social media by analyzing users' posting content, interaction behavior, and activity patterns. Tadesse et al. [17] specifically studied depression-related posts on the Reddit platform and achieved high detection accuracy through natural language processing technology, verifying the feasibility of social media data analysis in depression screening. Guntuku et al. [16] systematically summarized the application of social media data in mental health assessment and demonstrated the potential of this method in early depression screening. Additionally, De Choudhury et al. [29] explored the use of social media as a measurement tool for assessing depression prevalence in populations, demonstrating its potential for large-scale mental health monitoring through behavioral pattern analysis.

2.5 Multimodal Feature Fusion

Multimodal feature fusion aims to comprehensively utilize feature information from different sources to improve the comprehensiveness and accuracy of detection. From an algorithmic perspective, multimodal feature fusion mainly includes two strategies: feature-level fusion and decision-level fusion. Yang et al. [18] proposed a decision-level fusion method based on decision trees. This method first processes and predicts audio, visual, and language features separately, and then integrates the PHQ - 8 score prediction results of each modality and participant features (such as PTSD diagnosis, sleep status, etc.) through a gender - specific decision tree model [18].

2.6 Datasets for Depression Research

The development of AI-based depression detection systems relies heavily on high-quality, diverse datasets. The DAIC-WOZ dataset [40] contains clinical interviews with 189 participants, providing audio, video, and text data along with PHQ-8 depression scores, serving as a gold standard for multimodal depression research. The AVEC challenge datasets [12] offer annually updated resources with synchronized audio-visual recordings specifically designed for affective computing research.

As summarized in Table 1, more recently, the D-VLOG dataset [31] has emerged as a valuable resource, containing 961 video blogs (approximately 160 hours) from 816 subjects in naturalistic settings, offering a more ecologically valid alternative to laboratory-collected data. The Pittsburgh Depression Dataset [32] includes 49 subjects with 130 audio-visual samples recorded during clinical interviews, with expert annotations of depression severity.

For physiological signal analysis, several datasets are available: MODMA [41] provides 53-channel EEG recordings from clinically diagnosed patients and matched controls; DEAP [33] contains EEG and peripheral physiological signals from 32 participants during emotion elicitation; and SEED [34] offers EEG data specifically collected for emotion and depression analysis.

Text-based depression research is supported by datasets such as the eRisk CLEF collection [35], which contains early risk detection data from social media platforms; the Reddit Depression Corpus [42], which includes user posts from depression-related subreddits; and various Twitter depression datasets [36] that enable large-scale analysis of linguistic patterns.

Emotion-focused datasets that support depression research include EmotionLines [37] (containing 29,245 labeled utterances from 2,000 dialogues), IEMOCAP [38] (Interactive Emotional Dyadic Motion Capture Database), and MELD [39] (Multimodal EmotionLines Dataset), which provide emotional annotations valuable for understanding affective patterns in depression.

These diverse datasets collectively address different aspects of depression manifestation while presenting unique challenges in data quality, annotation consistency, and ethical considerations. Researchers must carefully select appropriate datasets based on their specific research questions and methodological approaches.

3 Algorithm Research on Intelligent Depression Detection

The evolution of intelligent depression detection algorithms has paralleled the broader advancements in machine learning technology, progressing from conventional machine learning approaches to sophisticated deep learning architectures. Each algorithmic paradigm offers distinct advantages in feature extraction, pattern recognition, and predictive classification, collectively providing a robust technical foundation for depression detection systems [20].

3.1 Traditional Machine Learning Methods

Conventional machine learning algorithms established the methodological groundwork for early depression detection research. Yang et al. [1] demonstrated the efficacy of support vector machines (SVM) in depression diagnosis using electroencephalography

Table 1. Commonly used depression prediction datasets.

Dataset Name	Data Type	Key Features
DAIC-WOZ	Audio/Video/Text	189 participants in clinical interviews with PHQ-8 depression scores
AVEC Challenge	Multimodal	Annual competition datasets with synchronized audio-visual recordings for affective computing research
D-VLOG	Video/Audio	961 vlogs (160+ hours) from 816 subjects in naturalistic settings
Pittsburgh Depression	Audio/Video	49 subjects with 130 samples from clinical interviews
MODMA	EEG signals	53-channel EEG recordings from clinically diagnosed patients and matched controls
DEAP	EEG/Physiological	EEG and peripheral physiological signals from 32 participants during emotion elicitation
SEED	EEG	EEG data collected specifically for emotion and depression analysis
eRisk CLEF	Text	Early risk detection data from social media platforms
Reddit Depression Corpus	Text	User posts from depression-related subreddits with temporal information
Depression Twitter	Social media text	Large-scale collection of tweets with depression indicators
EmotionLines	Text	29,245 labeled utterances from 2,000 dialogues with emotion annotations
IEMOCAP	Audio/Video/Motion	Interactive Emotional Dyadic Motion Capture Database with depression-relevant emotional states
MELD	Multimodal	Multimodal EmotionLines Dataset with rich contextual information

(EEG) data. Their approach employed a radial basis kernel function to effectively map high-dimensional, non-linear EEG features into a more discriminative feature space. Through systematic grid search optimization of the penalty factor C and kernel parameter γ, they achieved an optimal balance between model complexity and generalization capability.

In another significant application, Sau and Bhakta [28] implemented artificial neural networks (ANN) to predict depression among elderly residents in resource-limited settings in Kolkata, India. Their innovative approach integrated standardized clinical assessment tools with multimodal feature analysis. To overcome challenges associated with feature heterogeneity, they developed a hierarchical encoding framework that normalized diverse input data types and applied one-hot encoding where appropriate. Their model architecture incorporated dropout regularization to mitigate overfitting risks, resulting in a computationally efficient prediction system particularly suited for community-based mental health screening programs. Furthermore, a meta-analysis by Lee et al. [27] systematically reviewed the application of machine learning algorithms in predicting therapeutic outcomes for depression, highlighting their potential in enhancing clinical decision-making by identifying effective treatment strategies.

3.2 Deep Learning Methods

Deep learning has transformed depression detection through its powerful feature extraction and representation learning capabilities. Low et al. [10] advanced this field by applying BERT (Bidirectional Encoder Representations from Transformers) to analyze linguistic patterns in social media content. Their approach fine-tuned pre-trained language models on depression-specific texts, enabling the identification of subtle linguistic markers of depression that traditional methods might miss. The contextual understanding provided by these models significantly improved detection accuracy over conventional techniques. Guntuku et al. [13] further developed specialized frameworks for mental health monitoring on social platforms. Their approach combined convolutional neural networks (CNNs) with recurrent neural networks (RNNs) to capture both semantic content and temporal patterns in user-generated text. This architecture effectively detected changes in linguistic patterns that often precede clinical depression symptoms.

Multimodal deep learning has shown particular promise by integrating diverse data sources—audio, visual, and physiological signals—to create comprehensive representations of depressive symptoms. Yang et al. [18] demonstrated that neural networks could effectively combine features from speech, facial expressions, and interview content, outperforming single-modality approaches. Attention mechanisms represent another significant advancement, enhancing both model performance and interpretability. Knyazev et al. [21] provided insights into how these mechanisms improve generalization in neural networks, with important implications for depression detection systems. These approaches not only improve accuracy but also help clinicians understand which features influence the model's decisions—critical for clinical adoption. The ongoing development of explainable AI techniques continues to improve transparency, building trust and facilitating integration of these tools into clinical practice.

3.3 Algorithm Optimization and Improvement

To address clinical implementation challenges, researchers have pursued targeted algorithm optimizations that balance performance with practical constraints. Roy et al. [22] conducted a comprehensive review of deep learning applications in EEG signal analysis, identifying key optimization strategies for neural network architectures processing neurophysiological data. Their analysis highlighted the importance of temporal convolutional networks and recurrent architectures in capturing the dynamic patterns characteristic of depression-related EEG signals, while also emphasizing computational efficiency considerations essential for real-time monitoring applications.

Kingphai and Moshfeghi [23] extended these optimization approaches to mental workload assessment using EEG signals, demonstrating novel regularization techniques and parameter optimization methods that significantly improved model robustness across diverse patient populations. Their work introduced adaptive learning rate schedules and domain-specific data augmentation strategies that effectively addressed the data scarcity challenges common in clinical depression research. These optimization techniques have broader implications for depression detection systems, particularly in resource-constrained environments where computational efficiency and model reliability are paramount concerns.

4 Existing Problems and Challenges

Despite significant advancements in AI-based depression detection, several critical challenges remain across multiple domains including data management, technological development, clinical integration, and ethical considerations [20, 24, 25]. Addressing these challenges is essential for translating research innovations into practical clinical applications that can meaningfully impact mental healthcare delivery.

4.1 Data Quality and Standardization Issues

High-quality, standardized data represents the cornerstone of effective intelligent detection systems. Larson et al. [24] identified several fundamental challenges in medical data acquisition and utilization that directly impact model performance and clinical validity. These include inconsistent data collection protocols, variable annotation standards, and limited representation of diverse patient populations.

To overcome these limitations, researchers advocate for establishing unified acquisition standards and rigorous quality control processes. This approach necessitates the formation of specialized annotation teams working with comprehensive guidelines and protocols. Implementing multi-rater annotation procedures with formal consistency assessments significantly enhances the reliability of training data, which directly translates to improved model performance in clinical applications.

4.2 Technical Reliability and Robustness

The reliability and robustness of depression detection technologies remain significant concerns for real-world deployment. In their comprehensive review, Shatte et al. [20] highlighted three primary technical challenges: limited model generalization across diverse populations and clinical contexts, heightened sensitivity to data noise and environmental interference, and insufficient stability in multimodal feature fusion frameworks. These limitations can lead to inconsistent performance in naturalistic settings, potentially undermining clinical trust and adoption.

4.3 Clinical Practice and Application Issues

Bridging the gap between technological innovation and clinical implementation presents multifaceted challenges. Kellogg and Sadeh-Sharvit [30] emphasized that successful clinical integration depends on addressing model interpretability, seamless workflow integration, and real-time performance capabilities. Clinicians require systems that not only provide accurate assessments but also offer transparent reasoning that aligns with established clinical frameworks. Well-designed AI-assisted diagnosis systems have demonstrated potential to enhance diagnostic efficiency and accuracy while providing reliable decision support, but their implementation requires careful attention to clinical workflows and practitioner needs.

4.4 Ethical Privacy and Security Issues

The deployment of intelligent detection systems raises significant ethical considerations regarding patient privacy and data security. Larson et al. [24] proposed a comprehensive framework for protecting patient privacy in medical data utilization, emphasizing that technological innovation must proceed in parallel with robust ethical safeguards. This framework highlights the importance of informed consent, data minimization principles, and transparent algorithmic processes. Establishing sound ethical codes and privacy protection mechanisms not only safeguards patient rights but also significantly influences social acceptance and long-term technological adoption.

4.5 Cost-Benefit and Implementation Considerations

The widespread adoption of intelligent detection technologies necessitates careful evaluation of economic factors and implementation requirements. Huckvale et al. [26] advocated for comprehensive assessment frameworks that consider system development and maintenance costs, infrastructure requirements, training needs, and measurable clinical outcomes. Successful implementation strategies must balance technological capabilities with resource constraints, particularly in diverse healthcare settings. This requires establishing unified data standards, developing robust and interpretable algorithms, conducting rigorous clinical validation studies, implementing appropriate ethical safeguards, and creating evidence-based evaluation systems to demonstrate clinical and economic value [20, 24, 25].

5 Future Development Trends

Based on the analysis of the current research status of intelligent depression detection, this section will explore the future development trends of this field from three dimensions: technological innovation, multidisciplinary integration, and clinical practice.

5.1 Technological Innovation Directions

The continuous innovation of artificial intelligence technology will bring new development opportunities for intelligent depression detection. In terms of data sharing and privacy protection, Graham et al. [25] emphasized the need to pay special attention to privacy protection, data security, and algorithmic bias when applying AI in the mental health field. Larson et al. [24] proposed an ethical framework for medical data sharing, emphasizing that data sharing should be part of clinical research rather than a commercial transaction.

5.2 Multidisciplinary Cross - Integration

The development of intelligent depression detection requires the deep integration of multidisciplinary knowledge. The guidance of psychological theories is crucial for improving the scientific nature of detection models. Shatte et al. [20] pointed out in a review study that combining psychological theories with machine learning methods can significantly improve the clinical practicality and interpretability of models. Cummins et al. [19] presented a depression detection method based on voice analysis in a systematic review and proposed a complete technical roadmap from voice features to clinical application. Combining this method with traditional clinical assessments can provide more objective diagnostic evidence.

5.3 Clinical Practice Promotion

Promoting the clinical application of intelligent detection technology is a key direction for future development. Graham et al. [25] showed that the application of artificial intelligence in the mental health field requires a multi - dimensional evaluation system. Technical verification should be carried out in a single - center to evaluate the basic performance and usability of the system. Through multi - center collaborative research, the applicability and consistency of the technology in different scenarios should be verified, and a standardized clinical application process and quality control system should be established.

6 Conclusion and Outlook

This paper systematically reviewed the research status, key technologies, and future trends in the field of intelligent depression detection. Current research still faces the following main challenges: difficulty in obtaining high - quality data, the need to improve the stability and generalization ability of algorithm models, and practical problems such as system integration and workflow adaptation in clinical practice. The development of intelligent depression detection will pay more attention to technological innovation, clinical practice, and ethical norms.

Achieving these goals requires the continuous efforts and in - depth collaboration of all parties in industry, academia, research, and medicine. Through the organic combination of technological innovation and clinical practice, intelligent detection technology will surely play a greater role in the early screening, diagnosis, assessment, and prognosis monitoring of depression, and make positive contributions to improving the diagnosis and treatment level of depression and improving patient prognosis.

Data Availability Statement

Not applicable.

Funding

This work was supported by the Key Project of Natural Science Foundation of Xinjiang Hetian College under Grant 2025ZR002.

Conflicts of Interest

The authors declare no conflicts of interest.

Ethical Approval and Consent to Participate

Not applicable.

References

[1] Yang, C. Y., & Chen, Y. Z. (2024). Support vector machine classification of patients with depression based on resting-state electroencephalography. *Asian Biomedicine: Research, Reviews and News, 18*(5), 212-223. [CrossRef]

[2] Su, C., Xu, Z., Pathak, J., & Wang, F. (2020). Deep learning in mental health outcome research: a scoping review. *Translational psychiatry, 10*(1), 116. [CrossRef]

[3] World Health Organization. (2022). *World mental health report: Transforming mental health for all*. WHO Press.

[4] Santomauro, D. F., Herrera, A. M. M., Shadid, J., Zheng, P., Ashbaugh, C., Pigott, D. M., ... & Ferrari, A. J. (2021). Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. *The Lancet, 398*(10312), 1700-1712. [CrossRef]

[5] Wang, J., Zhao, Z., Yang, J., Wang, L., Zhang, M., & Zhou, M. (2024). The Association Between Depression and All-Cause, Cause-Specific Mortality in the Chinese Population—China, 2010–2022. *China CDC Weekly, 6*(40), 1022. [CrossRef]

[6] The Lancet Psychiatry. (2024). Global burden of disease 2021: Mental health messages. *The Lancet Psychiatry, 11*(8), 573. [CrossRef]

[7] Chinese Medical Association, Chinese Medical Association Publishing House, & Chinese Medical Association General Practice Branch. (2021). Guidelines for the diagnosis and treatment of depression in primary care (2021). *Chinese Journal of General Practitioners, 20*(12), 1101–1116. [CrossRef]

[8] Torous, J., Myrick, K. J., Rauseo-Ricupero, N., & Firth, J. (2020). Digital mental health and COVID-19: using technology today to accelerate the curve on access and quality tomorrow. *JMIR mental health, 7*(3), e18848. [CrossRef]

[9] Sadeghi, M., Richer, R., Egger, B., Schindler-Gmelch, L., Rupp, L. H., Rahimi, F., ... & Eskofier, B. M. (2024). Harnessing multimodal approaches for depression detection using large language models and facial expressions. *npj Mental Health Research, 3*(1), 66. [CrossRef]

[10] Low, D. M., Rumker, L., Talkar, T., Torous, J., Cecchi, G., & Ghosh, S. S. (2020). Natural language processing reveals vulnerable mental health support groups and heightened health anxiety on reddit during covid-19: Observational study. *Journal of medical Internet research, 22*(10), e22635. [CrossRef]

[11] De Choudhury, M., Gamon, M., Counts, S., & Horvitz, E. (2013). Predicting depression via social media. In *Proceedings of the international AAAI conference on web and social media* (Vol. 7, No. 1, pp. 128-137). [CrossRef]

[12] Ringeval, F., Schuller, B., Valstar, M., Cummins, N., Cowie, R., Tavabi, L., ... & Pantic, M. (2019, October). AVEC 2019 workshop and challenge: state-of-mind, detecting depression with AI, and cross-cultural affect recognition. In *Proceedings of the 9th International on Audio/visual Emotion Challenge and Workshop* (pp. 3-12). [CrossRef]

[13] Guntuku, S. C., Sherman, G., Stokes, D. C., Agarwal, A.

K., Seltzer, E., Merchant, R. M., & Ungar, L. H. (2020). Tracking mental health and symptom mentions on Twitter during COVID-19. *Journal of general internal medicine*, *35*(9), 2798-2800. [CrossRef]

[14] Cai, H., Han, J., Chen, Y., Sha, X., Wang, Z., Hu, B., ... & Gutknecht, J. (2018). A pervasive approach to EEG-based depression detection. *complexity*, *2018*(1), 5238028. [CrossRef]

[15] Jacobson, N. C., Weingarden, H., & Wilhelm, S. (2019). Digital biomarkers of mood disorders and symptom change. *NPJ digital medicine*, *2*(1), 3. [CrossRef]

[16] Guntuku, S. C., Yaden, D. B., Kern, M. L., Ungar, L. H., & Eichstaedt, J. C. (2017). Detecting depression and mental illness on social media: an integrative review. *Current Opinion in Behavioral Sciences*, *18*, 43-49. [CrossRef]

[17] Tadesse, M. M., Lin, H., Xu, B., & Yang, L. (2019). Detection of depression-related posts in reddit social media forum. *Ieee Access*, *7*, 44883-44893. [CrossRef]

[18] Yang, L., Jiang, D., He, L., Pei, E., Oveneke, M. C., & Sahli, H. (2016, October). Decision tree based depression classification from audio video and language information. In *Proceedings of the 6th international workshop on audio/visual emotion challenge* (pp. 89-96). [CrossRef]

[19] Cummins, N., Scherer, S., Krajewski, J., Schnieder, S., Epps, J., & Quatieri, T. F. (2015). A review of depression and suicide risk assessment using speech analysis. *Speech communication*, *71*, 10-49. [CrossRef]

[20] Shatte, A. B., Hutchinson, D. M., & Teague, S. J. (2019). Machine learning in mental health: a scoping review of methods and applications. *Psychological medicine*, *49*(9), 1426-1448. [CrossRef]

[21] Knyazev, B., Taylor, G. W., & Amer, M. (2019). Understanding attention and generalization in graph neural networks. *Advances in neural information processing systems*, *32*.

[22] Roy, Y., Banville, H., Albuquerque, I., Gramfort, A., Falk, T. H., & Faubert, J. (2019). Deep learning-based electroencephalography analysis: a systematic review. *Journal of neural engineering*, *16*(5), 051001. [CrossRef]

[23] Kingphai, K., & Moshfeghi, Y. (2024). Mental workload assessment using deep learning models from eeg signals: A systematic review. *IEEE Transactions on Cognitive and Developmental Systems*. [CrossRef]

[24] Larson, D. B., Magnus, D. C., Lungren, M. P., Shah, N. H., & Langlotz, C. P. (2020). Ethics of using and sharing clinical imaging data for artificial intelligence: a proposed framework. *Radiology*, *295*(3), 675-682. [CrossRef]

[25] Graham, S., Depp, C., Lee, E. E., Nebeker, C., Tu, X., Kim, H. C., & Jeste, D. V. (2019). Artificial intelligence for mental health and mental illnesses: an overview. *Current psychiatry reports*, *21*(11), 116. [CrossRef]

[26] Huckvale, K., Venkatesh, S., & Christensen, H. (2019). Toward clinical digital phenotyping: a timely opportunity to consider purpose, quality, and safety. *NPJ digital medicine*, *2*(1), 88. [CrossRef]

[27] Lee, Y., Ragguett, R. M., Mansur, R. B., Boutilier, J. J., Rosenblat, J. D., Trevizol, A., ... & McIntyre, R. S. (2018). Applications of machine learning algorithms to predict therapeutic outcomes in depression: a meta-analysis and systematic review. *Journal of affective disorders*, *241*, 519-532. [CrossRef]

[28] Sau, A., & Bhakta, I. (2017). Artificial neural network (ANN) model to predict depression among geriatric population at a slum in Kolkata, India. *Journal of clinical and diagnostic research: JCDR*, *11*(5), VC01. [CrossRef]

[29] De Choudhury, M., Counts, S., & Horvitz, E. (2013, May). Social media as a measurement tool of depression in populations. In *Proceedings of the 5th annual ACM web science conference* (pp. 47-56). [CrossRef]

[30] Kellogg, K. C., & Sadeh-Sharvit, S. (2022). Pragmatic AI-augmentation in mental healthcare: key technologies, potential benefits, and real-world challenges and solutions for frontline clinicians. *Frontiers in psychiatry*, *13*, 990370. [CrossRef]

[31] Yoon, J., Kang, C., Kim, S., & Han, J. (2022, June). D-vlog: Multimodal vlog dataset for depression detection. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 36, No. 11, pp. 12226-12234). [CrossRef]

[32] Girard, J. M., Cohn, J. F., Mahoor, M. H., Mavadati, S. M., Hammal, Z., & Rosenwald, D. P. (2014). Nonverbal social withdrawal in depression: Evidence from manual and automatic analyses. *Image and vision computing*, *32*(10), 641-647. [CrossRef]

[33] Koelstra, S., Muhl, C., Soleymani, M., Lee, J. S., Yazdani, A., Ebrahimi, T., ... & Patras, I. (2011). Deap: A database for emotion analysis; using physiological signals. *IEEE transactions on affective computing*, *3*(1), 18-31. [CrossRef]

[34] Zheng, W. L., & Lu, B. L. (2015). Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks. *IEEE Transactions on autonomous mental development*, *7*(3), 162-175. [CrossRef]

[35] Losada, D. E., & Crestani, F. (2016, August). A test collection for research on depression and language use. In *International conference of the cross-language evaluation forum for European languages* (pp. 28-39). Cham: Springer International Publishing. [CrossRef]

[36] Coppersmith, G., Dredze, M., Harman, C., & Hollingshead, K. (2015). From ADHD to SAD: Analyzing the language of mental health on Twitter through self-reported diagnoses. In *Proceedings of the 2nd workshop on computational linguistics and clinical psychology: from linguistic signal to clinical reality* (pp. 1-10).

[37] Chen, S. Y., Hsu, C. C., Kuo, C. C., & Ku, L. W. (2018). Emotionlines: An emotion corpus of multi-party conversations. *arXiv preprint arXiv:1802.08379*.

[38] Busso, C., Bulut, M., Lee, C. C., Kazemzadeh, A., Mower, E., Kim, S., ... & Narayanan, S. S. (2008). IEMOCAP: Interactive emotional dyadic motion capture database. *Language resources and evaluation, 42*(4), 335-359. [CrossRef]

[39] Poria, S., Hazarika, D., Majumder, N., Naik, G., Cambria, E., & Mihalcea, R. (2018). Meld: A multimodal multi-party dataset for emotion recognition in conversations. *arXiv preprint arXiv:1810.02508*.

[40] Gratch, J., Artstein, R., Lucas, G. M., Stratou, G., Scherer, S., Nazarian, A., ... & Morency, L. P. (2014, May). The distress analysis interview corpus of human and computer interviews. *In LREC* (Vol. 14, pp. 3123-3128).

[41] Cai, H., Yuan, Z., Gao, Y., Sun, S., Li, N., Tian, F., ... & Hu, B. (2022). A multi-modal open dataset for mental-disorder analysis. *Scientific Data, 9*(1), 178. [CrossRef]

[42] Pirina, I., & Çöltekin, Ç. (2018, October). Identifying depression on reddit: The effect of training data. In *Proceedings of the 2018 EMNLP workshop SMM4H: the 3rd social media mining for health applications workshop & shared task* (pp. 9-12). [CrossRef]

Liangguo Wang received the PhD degree in Computer Science and Technology from Beijing Institute of Technology, Beijing 100081, China, in 2018. His research interests include multimodal analysis and artificial intelligence applications.(Email: wlgbit@163.com)

Jiaqian Wu received the B.S. degree in Clinical Medicine from Anhui Medical University, Hefei 230032, China, in 2016. (Email:1310908555@qq.com)