International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(23s) 2026 pp. 169-178
https://doi.org/10.70917/ijcisim-2026-5562
Article

An Integrated AI Framework for Depression Detection Using Audio, Visual and Textual Data

***Mukundam Uduta[1], Dr. K. Srujan Raju[2]**

[1]Research Scholar, Bharatiya Engineering Science & Technology Innovation University (BESTIU) Andhra Pradesh, India
[2]Dean, Research & Development, CMR Technical Campus, Hyderabad, Telangana, India
*Corresponding Author: Mukundam Uduta

Abstract: Depression, as a mental health condition, often manifests through diverse verbal, non-verbal, and behavioral signals. Early and accurate detection of these indicators is critical to ensuring timely and effective intervention. This work introduces a multimodal Integrated AI framework for depression detection that uses: audio, visual, and textual data to improve classification performance. The proposed framework integrates features through ensemble stacking techniques by combination of depressive indicators such as facial expressions, speech patterns, and linguistic text content to form a holistic representation of depression detection. Conducted experiments demonstrates that the by integration of these modalities significantly enhanced classification performance. In all the conducted experiments, multimodal feature vectors outperformed unimodal and bimodal approaches. This best performance was achieved by this novel framework and reported an accuracy of 88.97%, precision of 91.42%, recall of 89.31%, and F1-score of 92.75%. By utilizing Distress Analysis Interview Corpus (DAIC), an open source dataset and proposed feature extraction methods, the system effectively fuses audio, visual, and textual cues, establishing a more reliable approach for detecting depression.
Keywords: Ensemble Learning, Stacking Method, Machine Learning, Multimodal Depression, Unimodal, Bimodal

1. Introduction

Depression, a widespread mental health condition affecting millions, has worsened due to COVID-19 [1]. The pandemic introduced uncertainties and isolation, adding challenges for those with depression and potentially triggering symptoms in new individuals. Recognizing depression's risk factors, from genetics and brain chemistry to life events, is essential for timely interventions. COVID-related stressors—social distancing, economic strain, health worries—complicate depression's assessment and treatment [2]. Integrating Machine Learning (ML) with WHO data enables improved risk identification, moving beyond conventional assessments. This study aims to advance adaptive intervention frameworks, supporting early recovery and resilience [3].

Recent efforts emphasize using technology and data to improve mental health care. COVID-19 has intensified challenges around depression, making it crucial to explore new predictive approaches examines depression, pandemic stressors, and machine learning (ML), referencing WHO statistics to include relevant risk factors in models for better risk identification. The goal is to support personalized interventions that address mental health complexities heightened by the pandemic. Using complex algorithms and large datasets, this work aims to enhance resilience by offering timely support. The proposed multimodal depression detection workflow involves dataset processing, feature extraction, normalization, ensemble learning, classification, and evaluation. Three possible research questions arising from these activities are as follows:

***Research Question* 1:** How does the use of multimodal features, including facial, acoustic, and textual data, affect the performance of systems detecting depression?

This investigates the impact of integrating multimodal data on system accuracy, comparing the improvement gained by combining facial expressions, speech, and text data with the accuracy of single-modal systems.

***Research Question* 2**: How robust are depression detection models when employing multimodal feature normalization and pre-processing techniques?

This evaluates the effectiveness of various pre-processing and normalization techniques on the quality of feature extraction. It examines how these processes enhance the overall performance of the classification system in accurately identifying individuals with depression.

***Research Question* 3:** How does ensemble learning through stacking improve the classification of depressed versus non-depressed populations compared to individual classification models?

This explores the benefits of ensemble stacking techniques in improving classification performance. It focuses on how combining models enhances the accuracy and reliability of distinguishing between depressed and non-depressed individuals compared to using single models.

2. Literature Review

Facial markers are widely used in diagnosing depression for several key reasons [1][4]. Firstly, individuals with depression tend to display typical facial expressions, such as reduced smiling, frequent lip pressing, and increased corrugator muscle activity, which are linked to frowning. Additionally, they often exhibit sad, negative, or neutral expressions and display irregular eye movements, like rapid or delayed blinking. Advances in technology have made it easier and more efficient to capture facial images through web cameras. Numerous tools now allow for detailed analysis and extraction of visual features from facial images, with researchers using software such as the Computer Expression Recognition Toolbox (CERT), OpenFace, and iMotions to study facial expressions and movements in depth[2][3][4].

Wang et al. [5] conducted a study on facial expression differences between depressed and non-depressed individuals in controlled settings. Using a person-specific active appearance model, 68 facial landmarks were identified, and statistical features were calculated from distances between key facial points (e.g., near the eyes, eyebrows, and mouth corners). These features were then input into an SVM classifier, achieving 78% test accuracy.

Sharifa Alghowinem et al. [7] studied eyelid behavior and blinking patterns in depressed individuals, finding that they have a smaller eyelid distance when open and longer blink durations compared to non-depressed individuals.

In a follow-up study, Alghowinem et al. [8] explored variations in head pose and movement between depressed and non-depressed individuals. They found that, compared to non-depressed participants, those with depression demonstrated longer gaze durations, particularly in rightward and downward directions, slower head movements, and fewer shifts in head orientation.

Acoustic speech features, which include both linguistic (what is said) and paralinguistic (how it is said) elements, are crucial in diagnosing depression. These features, influenced by mental state, are commonly used by clinicians. Key prosodic features like loudness, pitch, formants, energy, shimmer, and jitter help differentiate between depressed and non-depressed individuals. Speech data collection is straightforward, and tools like openSMILE, PRAAT, and COVAREP assist in extracting these acoustic features for detailed analysis.

Nicholas Cummins et al. [9][10] investigated acoustic features differentiating normal and depressed speakers, focusing on spectral centroid frequencies and amplitudes, normalized via Mel-Frequency Cepstral Coefficients (MFCCs). Their research showed that multidimensional feature sets, combining multiple acoustic characteristics, outperformed single-dimensional ones.

They used Gaussian mixture models to classify depressed and non-depressed speakers. In a follow-up study, they explored how depression reduces acoustic variability in phonetic events, using metrics like Average Weighted

Variance (AWV), Acoustic Movement (AM), and Acoustic Volume (AV). They found that depressed individuals tend to have a smaller vowel space, an important depression marker.

Recent studies have explored multimodal approaches, combining facial expressions, speech, and linguistic features for more accurate depression diagnosis [11]. This strategy enhances accuracy by leveraging each modality's strengths and complementarity.

Williamson et al. [12] combined facial movement and acoustic verbal cues to detect psychomotor retardation, a depression symptom. Using PCA for data reduction and Gaussian mixture model classification, they found multimodal integration outperformed individual modality analysis. Similarly, Alghowinem et al. [13] showed that combining head pose, eye gaze, and verbal cues resulted in higher classification accuracy than single-modality analysis, demonstrating the effectiveness of multimodal approaches in depression diagnosis.

3. Methodology

The essential steps of the multimodal depression detection method, as illustrated in Fig. 1, encompass several critical processes that must be clearly understood, as they significantly enhance the system's performance regarding accuracy and robustness. These steps include:

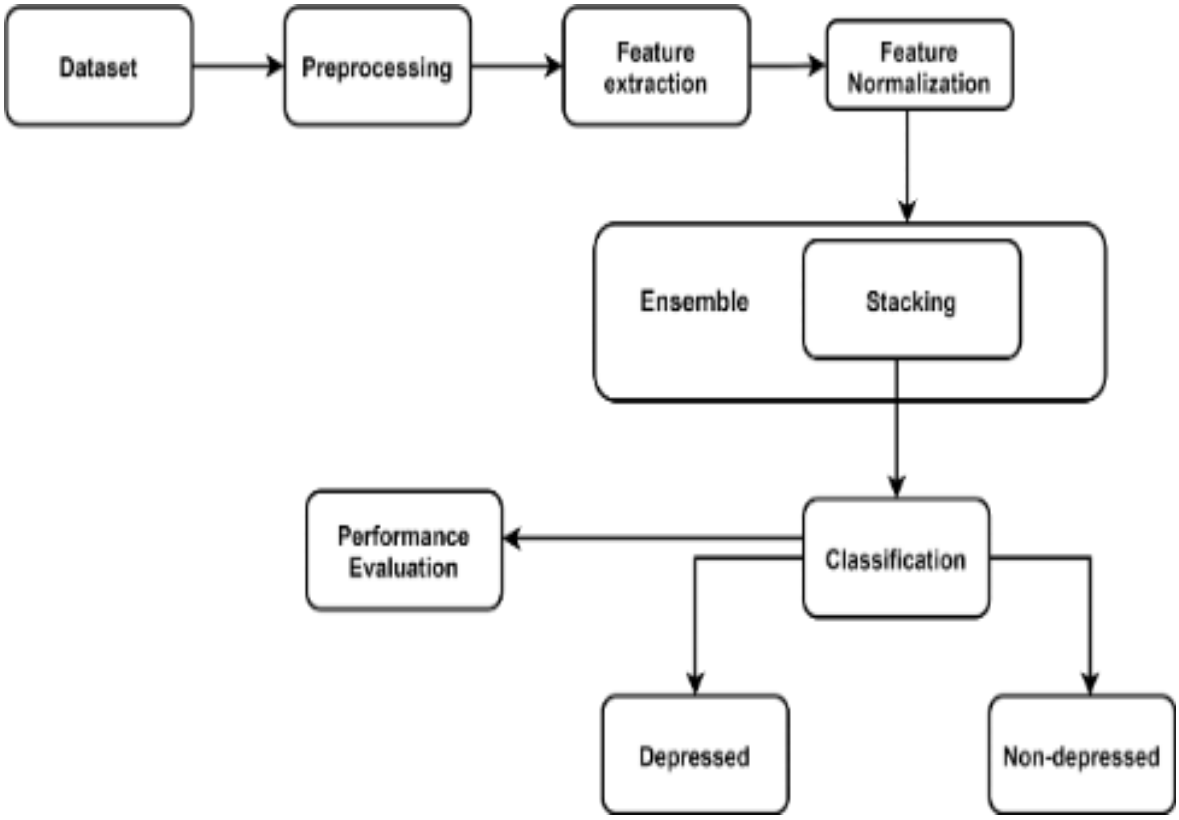

Figure. 1 Overview of the proposed work

Step 1 - Data Collection: Gathering multimodal data, including audio, visual, and textual information from subjects, ensuring a comprehensive dataset that captures various aspects of depression.

Step 2 - Dataset Processing: Pre-processing the collected data to remove noise and irrelevant information, standardizing formats, and preparing the data for feature extraction.

Step 3 - Feature Extraction: Identifying and extracting relevant features from each modality, such as facial landmarks, acoustic parameters, and linguistic characteristics, which are essential for effective classification.

Step 4 - Feature Extraction: Normalizing the extracted features to ensure consistency across modalities, enhancing the comparability of different feature sets.

Step 5 - Ensemble Learning: Implementing ensemble techniques, such as stacking, to combine the strengths of individual classifiers trained on different modalities, thereby improving overall classification performance.

Step 6 - Classification: Using ML algorithms, such as SVMs or Gaussian mixture models, to classify the processed and integrated features into depressed and non-depressed categories.

Step 7 - Performance Evaluation: Assessing the system's performance using metrics such as accuracy, precision, recall, and F1-score to ensure the model's effectiveness in real-world applications.

3.1 Multimodal Dataset Collection

The process starts with collecting an appropriate dataset, ideally audio-visual, but it may also include text-based data (such as transcripts or features) and behavioural indicators. The dataset should be labelled to indicate whether subjects are depressed, making it suitable for supervised learning. The data source includes a publicly available depression dataset, such as the Distress Analysis Interview Corpus Wizard of Oz (DAIC-WOZ). Splitting the dataset into training and testing segments, allocating 80% for training and 20% for testing.

In the multimodal context, this dataset should require Visual, audio and text data.

3.2 Pre-Processing

The goal at this stage is to prepare a clean, well-structured dataset for feature extraction. Pre-processing varies across modalities.

Visual data: Steps include face detection, facial landmarking, alignment, and resizing to standardize features like eye, mouth, and eyebrow movements for reliable feature extraction.

Audio data*:* Pre-processing includes noise reduction, Voice Activity Detection (VAD), and normalization to improve the extraction of pitch, energy, and other paralinguistic features.

Text data: Text transcription is pre-processed by removing stop words, normalizing punctuation, and tokenizing. Stemming or lemmatization reduces words to a base form for further linguistic feature extraction.

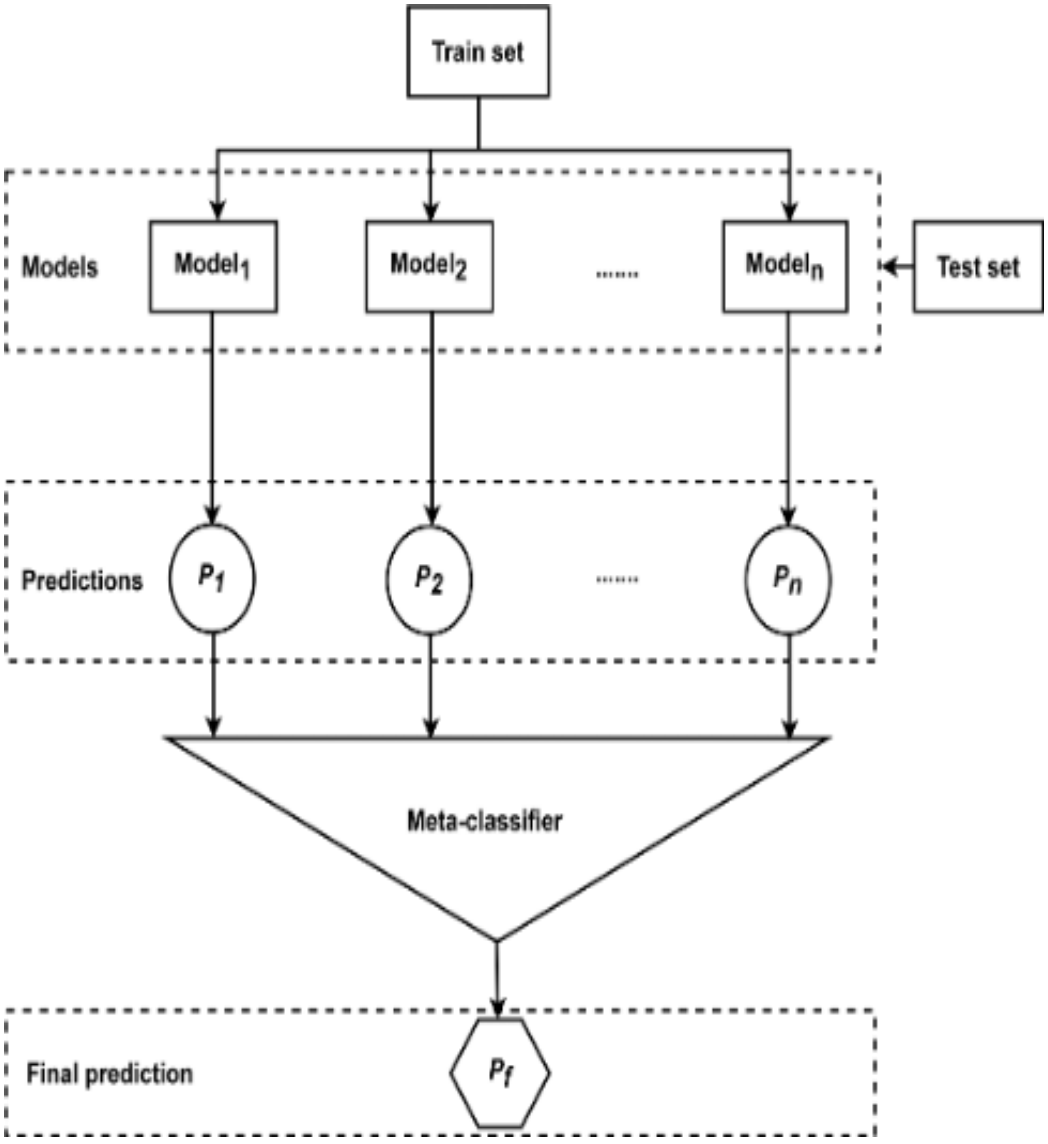

Figure. 2 Block diagram of the Stacking ensemble

Method

3.3 Feature Normalization

It is of utmost importance to normalize all the features before feeding them to the classification model since all features should be scaled equally to reduce possible biases introduced by individual variations. Techniques like z-score normalization or min-max scaling may be used to adjust the range of features such that no single modality dominates the function of classification. For multimodal data combinations, feature normalization is very important because the ranges and distributions of visual, acoustic, and textual features differ vastly.

3.4 Ensemble Learning

To maximize the benefits of integrating multiple modalities, ensemble learning techniques, like stacking [14-16], combine predictions from various classifiers to improve accuracy. This approach enhances model performance by leveraging diverse base models.

Stacking involves training multiple classifiers on the same dataset and using their predictions as input for a Meta-learner as shown in Fig. 2, which synthesizes the outputs for better decision-making. This method enhances performance by combining the strengths of different classifiers, making it effective for multimodal depression detection. Base models like SVM, Random Forests, Decision Trees, or Neural Networks are trained on various feature subsets (visual, acoustic, textual), and a Meta-model optimizes their combined predictions for more accurate classification. Stacking is particularly useful in multimodal contexts, where each modality provides complementary information, improving overall performance.

3.5 Classification

After feature normalization and stacking through ensemble learning, classification follows, aiming to categorize subjects as depressed or not. Common classifiers used in machine learning [17-19] include:

Support Vector Machines: Effective in high-dimensional spaces, widely used for depression detection in both images and acoustic signals.

Random Forests or Decision Trees: Capture non-linear data patterns well and are highly effective in ensemble settings.

Deep Learning Models (e.g., CNNs, RNNs)*:* Highly effective for multimodal large datasets, as they automatically learn and optimize feature representations from raw data [20-21].

With the comprehensive multimodal feature set, the classifier makes its final prediction. Stacking-based ensemble learning strengthens this process, improving classification accuracy, especially for borderline or subtle signs of depression.

3.6 Performance Evaluation

The final step in the proposed method is performance evaluation, using metrics to assess the depression detection model's accuracy and reliability:

Accuracy: Percentage of correctly classified instances.

Precision, Recall and F1-score: These metrics assess false positives (non-depressed classified as depressed) and false negatives (missed detection of depressed individuals), important in healthcare applications like depression detection.

Area under the Curve (AUC): Measures the classifier’s ability to distinguish between depressed and non-depressed classes.

Cross-validation: Helps prevent overfitting by training the model on some data folds and testing on others, improving its reliability on unseen data [22-23].

This evaluation ensures the model is optimized for real-world mental health diagnosis applications.

3.7 Advantages of the Proposed Method

The proposed system improves depression detection accuracy and adaptability by integrating multiple data modalities and applying ensemble machine learning techniques. This combination enhances performance and scalability, making it suitable for real-world applications. Key benefits include:

Multimodal Integration: This system can trace a wider range of depression indicators with the integration of visual, acoustic, and textual features, hence making it more accurate.

Ensemble Learning with Stacking: With stacking, the system combines predictions from multiple classifiers to improve overall performance and reduce the likelihood of classification errors.

Scalability and Flexibility: The methodology is directly applicable to different datasets and easily allows the inclusion of additional modalities (such as physiological signals) to further enhance robustness [24].

In summary, the proposed method employs a systematic, multimodal approach to depression detection, combining advanced ML techniques with domain-specific knowledge on facial, acoustic, and textual cues. This is enhanced by ensemble learning, particularly stacking, which effectively manages the complexities of multimodal data while ensuring accurate and reliable predictions [25].

4. Implementation and Results

This section reviews Table 1, summarizing the performance metrics of depression detection using unimodal, bimodal, and multimodal feature sets. The advantages of combining audio, visual, and textual cues are evident when these modalities are stacked in an ensemble approach [26]. This analysis aligns with the research questions in multimodal depression detection, highlighting how multimodal methods outperform unimodal and bimodal techniques. Discussion and Justification are presented below:

4.1 Research Question 1: Impact of Multimodal Features on Performance

The results in Table 1 show that multimodal systems significantly improve detection accuracy over unimodal systems. This supports the hypothesis that integrating diverse data sources captures a broader range of depressive symptoms, enhancing diagnostic precision.

The 88.97% accuracy demonstrates the model's effectiveness in correctly classifying individuals with and without depression. The precision score of 91.42% reflects a high level of accuracy in positive predictions, minimizing false positives, which is crucial in depression diagnosis to avoid inappropriate interventions.

A recall rate of 89.31% shows that the model effectively identifies most depression cases, minimizing missed diagnoses. This is crucial for a sensitive mental health detection system that ensures individuals in need of support are not overlooked. The F1-score of 92.75% indicates a strong balance between precision and recall, confirming the multimodal approach's reliability in identifying true positives and avoiding false negatives.

The multimodal system outperforms individual modalities, with audio achieving 60.39% accuracy, and text 54.35%. While visual data performs better than audio and text alone, it still lags behind the combined modalities. This indicates that each modality captures unique depression-related aspects, and when integrated, they offer a more comprehensive understanding, detecting subtle signals that may be missed individually.

4.2 Research Question 2: Robustness of Normalization Techniques

The comparative performance metrics indicate that multimodal feature normalization leads to improved feature quality, which directly affects classification outcomes. The analysis shows that systems employing advanced normalization techniques on multimodal data consistently outperform those relying on unimodal or bimodal data, reinforcing the importance of robust pre-processing in enhancing model reliability.

While Table 1 presents performance metrics, it implicitly underscores the significance of feature normalization and pre-processing in enhancing the model's effectiveness. The observed improvement in metrics as features are combined indicates that effective pre-processing has facilitated the integration of multiple modalities.

When dealing with diverse data types—such as raw audio, facial videos, and textual information—normalization is crucial to ensure that features can be analysed together without introducing biases related to scale or distribution. Here are some key considerations:

Visual Features: Facial action units, like eye, eyebrow, and lip movements, can vary in intensity. Normalizing these features ensures consistent interpretation of facial expressions across subjects.

Acoustic Features: Parameters like pitch, loudness, jitter, and prosodic patterns differ from visual or textual features. Normalization aligns these acoustic metrics with other modalities.

Textual Features: Text features, such as negative sentiment frequency and syntax complexity, should align with acoustic and visual modalities to represent the subject's mental state comprehensively.

Normalization techniques, like standardization or z-score normalization, ensure all features contribute equally to the model's classification. The stacking method in the ensemble model leverages these normalized features, enhancing the discriminative power of each modality. This results in improved performance when combining all three modalities, leading to the highest performance metrics. Normalizing features effectively enables the model to achieve more accurate and robust depression classification by utilizing the unique information from each modality.

4.3 Research Question 3: Effectiveness of Ensemble Learning

The metrics show that ensemble methods, especially stacking, perform better in classifying depressed versus non-depressed populations [27-28]. By combining classifiers trained on different modalities, the ensemble approach enhances classification accuracy, demonstrating the value of integrating multiple modalities for more precise diagnostic results.

The effectiveness of ensemble learning, particularly through stacking, is evident in the multimodal model's superior performance compared to unimodal and bimodal approaches. The core principle of stacking involves integrating separate classifiers that may excel in capturing different features or modalities, resulting in a Meta-model that is both accurate and robust. This strategy is particularly well-suited for multimodal depression detection for several reasons:

i. Limitations of Unimodal Classifiers: Unimodal classifiers focus on a single feature—audio, video, or text—capturing only specific aspects of depressive behavior. For example, a visual classifier may detect facial expressions but miss key prosodic features in audio data, while an acoustic model might analyze speech patterns but overlook visual cues or textual context.

ii. Advantages of Bimodal Classifiers: Bimodal classifiers, which combine two sources of information (e.g., audio and visual, visual and textual, or audio and textual), typically demonstrate improved performance over unimodal systems. By leveraging the complementary nature of the two modalities, these classifiers can enhance overall predictive accuracy. For example, a combination of visual and text features (V+T) achieves an accuracy of 80.77%, significantly outperforming any single modality.

iii. Strength of Multimodal Classifiers: Multimodal classifiers that integrate all three modalities (A+V+T) achieve the highest performance metrics. This demonstrates that stacking these modalities allows the ensemble method to effectively capture complementary information from each feature set. Since depression manifests differently across individuals, combining auditory, visual, and textual cues enables the model to detect a wider range of depressive symptoms.

In summary, the application of stacking within the multimodal framework enhances the detection of depression by capitalizing on the unique strengths of each modality, resulting in more accurate and comprehensive assessments. This approach not only addresses the limitations inherent in unimodal and bimodal classifiers but also aligns with the complex nature of depression, ultimately leading to better outcomes in mental health diagnostics.

Table 1 supports the research questions, highlighting the significant advantages of multimodal depression detection systems and emphasizing the value of integrated approaches in mental health assessments.

The formula for accuracy is as shown in (1).

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \qquad (1)$$

Where:

- TP = True Positives (correctly predicted positive instances)

- TN = True Negatives (correctly predicted negative instances)
- FP = False Positives (incorrectly predicted positive instances)
- FN = False Negatives (incorrectly predicted negative instances)

Precision: Measures the proportion of recommended articles that are relevant and is calculated as shown in (2).

$$Precison(P) = \frac{TP}{TP+FP} \quad (2)$$

Recall: Indicates the percentage of relevant articles that occurred effectively recommended and is calculated as shown in Eq. (3)

$$Recall(R) = \frac{Total\ Num\_Rel\_Articles\ in\ Dataset}{Num\ of\ Relevant\ Articles\ Recommended} \quad (3)$$

F1-Score: Harmonic mean of precision and recall and is calculated as shown in Eq. (4)

$$F1\ Score\ = 2 * \frac{P * R}{P + R} \quad (4)$$

5. Conclusion

This study presents a multimodal framework for detecting depression by integrating audio, visual, and textual features using the ensemble stacking method. This approach significantly enhances the accuracy and reliability of depression diagnosis. The investigation addresses three research questions, each thoroughly explored and justified based on the experimental results obtained. The findings effectively answer the research questions posed and validate the importance of multimodal feature integration and ensemble learning in improving depression detection systems.

Table 1. Results obtained using individual and multimodal feature sets by proposed ensemble stacking method

Input Features	**Performance metrics**			
	Accuracy	Precision	Recall	F1-score
Audio (A)	60.39	63.68	82.73	71.57
Visual (V)	72.17	74.13	83.22	78.47
Textual (T)	54.35	58.69	78.85	65.11
A+V	79.15	81.06	85.17	87.21
V+T	80.77	79.46	83.71	80.01
A+T	76.37	77.78	84.86	80.98
A+V+T	88.97	91.42	89.31	92.75

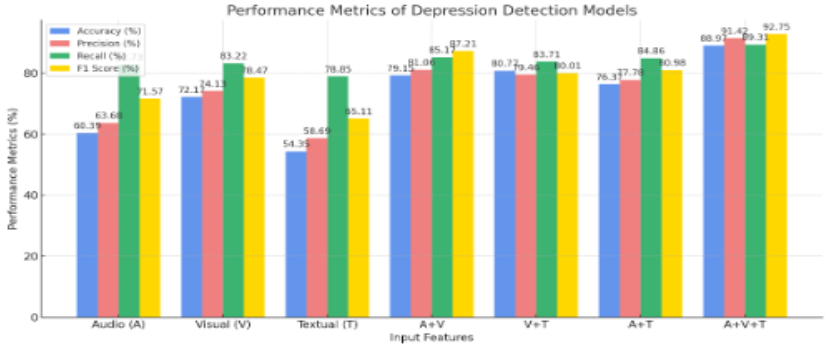

Figure. 3 Performance metrics of unimodal, bimodal, and multimodal approaches

REFERENCES

1. https://www.who.int/news-room/fact-sheets/detail/depression.
2. M. Suresh, A. M. Reddy, "Multimodal Depression Detection Using Audio, Visual and Textual Cues: A Survey.", *NeuroQuantology*, vol. 20, no. 4, pp. 325-336, 2022, doi:10.14704/nq.2022.20.4.nq22127.
3. WHO. (2021). Mental health preparedness and response for the COVID-19 pandemic: *Report by the director-general*. https://apps.who.int/iris/handle/10665/359717.
4. M. Suresh, A. M. Reddy, "Enhancing Depression Prediction Accuracy Using Filter and Wrapper-Based Visual Feature Extraction," *J. Adv. Inf. Technol*., vol. 14, no. 6, 2023, doi: 10.12720/jait.14.6.1425-1435.
5. Q. Wang, H. Yang, and Y. Yu, "Facial expression video analysis for depression detection in Chinese patients," *J. Vis. Commun. Image Represent.*, vol. 57, pp. 228–233, 2018, doi: 10.1016/j.jvcir.2018.11.003.
6. T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance and models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 681–685, 2001.
7. Alghowinem S, Goecke R, Wagner M, Parker G, (2013) "Eye movement analysis for depression detection", 4220–4224.
8. Alghowinem S, Goecke R, et al., (2013) "Head pose and movement analysis as an indicator of depression", 283–288.
9. Cummins, Nicholas \& Epps, Julien (2011). An Investigation of Depressed Speech Detection: Features and Normalization. In: *Proc. Interspeech.* 2997-3000. 10.21437/Interspeech.2011-750.
10. Suresh M, A. M. Reddy, "Correlation-Based Attention Weights Mechanism in Multimodal Depression Detection: A Two-Stage Approach," *Int. J. Intell. Eng. Syst.,* vol. 17, no. 5, pp. 350-365, 2024, doi: 10.22266/ijies2024.1031.28.
11. Williamson JR, et al., "Vocal and facial biomarkers of depression based on motor incoordination and timing", pp. 65–72, 2014. Doi:10.1145/2661806.266180.
12. S. Alghowinem et al., "Multimodal depression detection: Fusion analysis of paralinguistic, head pose and eye gaze behaviors," *IEEE Transactions on Affective Computing*, vol. 9, no. 4, pp. 478–490, 2018.
13. Suresh M., A. M. Reddy, "A Stacking-based Ensemble Framework for Depression Detection using Audio Signals," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 7, pp. 603-612, 2023, doi: 10.14569/IJACSA.2023.0140767.
14. Hema, M.S., et al., "Prediction analysis for Parkinson disease using multiple feature selection classification methods", *Multimedia Tools and Applications*, vol. 82, pp. 42995–43012, 2023, doi:10.1007/s11042-023-15280-6.
15. Mamatha, K., et al., "Fostering Employability - Synergy between Industry and Academia", *Journal of Engineering Education Transformations*, vol. 33(0), pp. 319-322. Doi:10.16920/jeet/2020/v33i0/150178.
16. D. L. Padmaja, et.al., "A Comparative Study on Natural Disasters," In: *Proc. of 2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC), Salem, India,* pp. 1704-1709, doi: 10.1109/ICAAIC53929.2022.9793039.
17. Subhrajyoti Deb et al., "Colour Image Encryption Using an Improved Version of Stream Cipher and Chaos," *International Journal of Ad Hoc and Ubiquitous* Computing, Vol. 41, No. 2, 2022, pp. 118–133, doi:10.1504/ijahuc.2022.125428.
18. Kumar C, Reddy, K. S. (2019). Effective Data Analytics on Opinion Mining. *International Journal of Innovative Technology and Exploring Engineering*, 8, 2073-2080. doi: 10.35940/ijitee.J9332.0881019.
19. A. Mallikarjuna Reddy, et al.,"Face recognition based on stable uniform patterns" *International Journal of Engineering Technology*, Vol.7, No.(2), pp. 626-634, 2018, doi: 10.14419/ijet.v7i2.9922
20. Sudeepthi Govathoti, et al., "Data Augmentation Techniques on Chilly Plants to Classify Healthy and Bacterial Blight Disease Leaves" *International Journal of Advanced Computer Science and Applications,* 13(6), 2022. doi: 10.14569/IJACSA.2022.0130618
21. Swarajya Lakshmi V. Papineni, et al., "Big Data Analytics Applying the Fusion Approach of Multicriteria Decision Making with Deep Learning Algorithms," *Int. J. Eng. Trends Technol*., vol. 69, no. 1, pp. 24-28, doi: 10.14445/22315381/IJETT-V69I1P204.
22. A Mallikarjuna Reddy, et al., "Age Classification Using Motif and Statistical Features Derived On Gradient Facial Images", *Recent Advances in Computer Science and Communications* (2020) 13: 965. doi:10.2174/2213275912666190417151247.
23. A.Mallikarjuna, B. Karuna Sree, "Security towards Flooding Attacks in Inter Domain Routing Object using Ad hoc Network" *International Journal of Engineering and Advanced Technology*, Volume-8 Issue-3, February 2019.
24. Mallikarjuna Reddy, A., et al., (2019), "Generating cancelable fingerprint template using triangular structures", *Journal of Computational and Theoretical Nanoscience*, Volume 16, Numbers 5-6, pp. 1951-1955(5), doi: 10.1166/jctn.2019.7830.
25. V. Venkata Krishna, L. Sumalatha, "Efficient Face Recognition by Compact Symmetric Elliptical Texture Matrix (CSETM)", *Jour.of Adv Research in Dynamical Control Systems*, Vol. 10, 4, 2018.
26. K. Sudheer Reddy and C. Srinagesh, "Fostering Problem Solving through Innovative Knowledge Events," In: *Proc. of 8th ICCSE, Colombo, Sri Lanka*, 2013, pp. 1233-1238, doi: 10.1109/ICCSE.2013.6554108.
27. Prasanth Rao, Adiraju, et al., "Automated Soil Residue Levels Detecting Device With IoT Interface," in *Advances in IoT and Its Applications*, 2021, doi: 10.4018/978-1-7998-2566-1.ch007.
28. C. N. S. Kumar et al., "Similarity Matching of Pairs of Text Using CACT Algorithm," *Int. J. Eng. Adv. Technol*., vol. 8, no. 6, pp. 2296-2298, 2019, doi: 10.35940/ijeat.F8685.088619.