iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** V **Month of publication:** May 2025

DOI: https://doi.org/10.22214/ijraset.2025.70895

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

A Comparative Review University Student's Mental Stress Detection Using Machine Learning

Prof. Rahul M. Dhokane[1], Viki Punjaram Shirsath[2], Vinay Babasaheb Pansare [3], Durgesh Deepak Shinde[4], Jivan Suresh Patil[5]
Department of Information Technology, Sir Visvesvaraya Institute of Technology, Nashik, Maharashtra, India

Abstract: Depression is a prevalent mental health disorder affecting millions globally. Early detection is crucial for effective intervention and treatment. This paper presents a comprehensive system for detecting depression levels by integrating multimodal data sources, including text, audio, and visual inputs. Utilizing advanced deep learning techniques and natural language processing, the system aims to provide accurate and timely assessments of an individual's mental health status.
The proposed system leverages convolutional neural networks (CNNs) for facial emotion recognition, recurrent neural networks (RNNs) for audio analysis, and transformer-based models for textual data processing. By combining these modalities, the system achieves a holistic understanding of the user's emotional state, enhancing the accuracy of depression detection. Experimental results demonstrate the efficacy of the system in real-world scenarios, highlighting its potential as a tool for mental health professionals and individuals alike.
Keywords: Depression Detection, Multimodal Analysis, Deep Learning, Natural Language Processing, Emotion Recognition

I. INTRODUCTION

Depression is a significant public health concern, characterized by persistent feelings of sadness, loss of interest, and various cognitive and physical symptoms. Traditional methods of diagnosing depression often rely on selfreported questionnaires and clinical interviews, which can be subjective and time-consuming. With the advent of technology, there is a growing interest in developing automated systems for early detection of depression.

Recent studies have explored the use of machine learning and deep learning techniques for analyzing various data modalities to detect signs of depression. For instance, natural language processing (NLP) has been employed to analyze textual data from social media posts and interviews, revealing linguistic patterns associated with depressive symptoms [3]. Similarly, audio analysis has been used to detect changes in speech patterns, such as reduced pitch variability and slower speech rate, which are indicative of depression [?].

Moreover, facial expression analysis through computer vision techniques has shown promise in identifying emotional states related to depression. By integrating these modalities, researchers aim to develop comprehensive systems that can accurately assess an individual's mental health status. This paper builds upon these advancements by proposing a multimodal depression detection system that combines text, audio, and visual data using state-ofthe-art deep learning models.

II. LITERATURE SURVEY

Several studies have investigated the application of machine learning and deep learning techniques for depression detection. In [1], a multimodal approach was introduced, combining text, audio, and facial features using deep learning models like VGG-16 for speech, BERT for language, and ResNet-50 for visual features. This approach demonstrated improved accuracy in predicting depression severity.

Another study focused on using large language models (LLMs) to extract depression-related indicators from interview transcripts, utilizing the Patient Health Questionnaire-8 (PHQ-8) score for training the prediction model [2]. The study highlighted the effectiveness of LLMs in automating the extraction of depression-relevant features from textual data.

In the realm of social media analysis, researchers have explored the use of sentiment analysis and NLP techniques to detect signs of depression in user-generated content. For example, a study employed a hybrid deep learning classifier combining effective word embedding with sentiment analysis to detect depression with high accuracy [4].

Furthermore, the DORIS system was developed to detect depression symptoms in social media posts by analyzing mood courses in emotionally intense texts, using DSM-5 criteria for automatic labeling [5]. This system outperformed existing methods, highlighting the potential of integrating psychological criteria with machine learning techniques.

These studies underscore the importance of leveraging multiple data modalities and advanced machine learning techniques to enhance the accuracy and reliability of depression detection systems.

III. MOTIVATION

The increasing prevalence of depression and its impact on individuals and society necessitate the development of effective tools for early detection and intervention. Traditional diagnostic methods, while valuable, often face limitations such as subjectivity, time constraints, and accessibility issues. There is a pressing need for automated systems that can provide accurate, timely, and accessible assessments of mental health.

Advancements in technology, particularly in the fields of machine learning, deep learning, and natural language processing, offer promising avenues for developing such systems. By harnessing the capabilities of these technologies, it is possible to analyze various data sources—textual, auditory, and visual—to detect subtle signs of depression that may not be apparent through conventional methods.

The integration of multimodal data allows for a more comprehensive understanding of an individual's emotional state. For instance, combining facial expression analysis with speech and text analysis can provide a richer context for assessing mental health. This holistic approach can lead to more accurate and reliable detection of depression, facilitating early intervention and support.

The motivation behind this project is to develop a system that leverages these technological advancements to provide an effective tool for depression detection, ultimately contributing to improved mental health outcomes.

IV. METHODOLOGY

The proposed depression detection system integrates three primary modalities: text, audio, and visual data. Each modality is processed using specialized deep learning models to extract relevant features indicative of depression.

1) Text Analysis

Textual data, such as user responses to questionnaires or social media posts, are analyzed using transformerbased models like BERT. These models are fine-tuned on datasets labeled with depression indicators to capture linguistic patterns associated with depressive symptoms. Sentiment analysis is also performed to assess the emotional tone of the text.

2) Audio Analysis

Audio recordings of speech are processed using recurrent neural networks (RNNs) and convolutional neural networks (CNNs) to analyze prosodic features such as pitch, tone, and speech rate. These features are indicative of emotional states and can reveal signs of depression.

3) Visual Analysis

Facial expressions are analyzed using CNNs trained on facial emotion recognition datasets. The system detects micro-expressions and other facial cues that may indicate depressive moods. Techniques like facial action coding systems (FACS) are employed to quantify facial movements.

4) Multimodal Fusion

The features extracted from each modality are combined using a fusion strategy, such as concatenation or attention mechanisms, to form a comprehensive feature vector. This vector is then input into a classifier, such as a fully connected neural network, to predict the likelihood and severity of depression.

V. PROPOSED SYSTEM

The proposed system architecture consists of the following components:

1) Data Acquisition Module: Collects textual, audio, and visual data from users through questionnaires, voice recordings, and video inputs.
2) Preprocessing Module: Cleans and preprocesses the data, including noise reduction in audio, normalization of text, and alignment of facial landmarks in images.
3) Feature Extraction Module: Utilizes deep learning models to extract features from each modality.
4) Fusion Module: Integrates features from all modalities to form a unified representation.
5) Classification Module: Predicts depression levels based on the fused features.
6) User Interface: Provides feedback to users and allows interaction with the system.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

VI. SYSTEM ARCHITECTURE

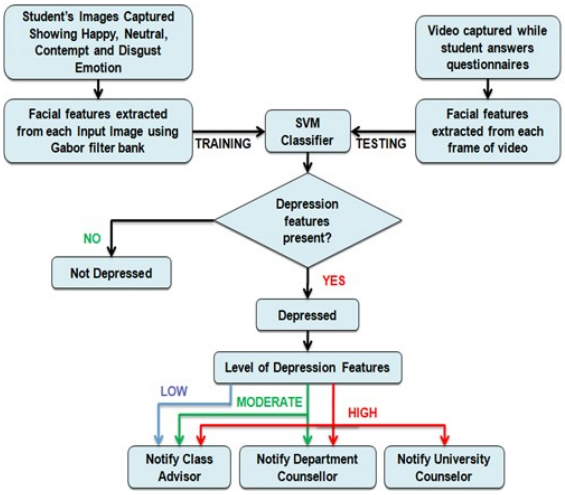


Figure 1: System Architecture for Multimodal Depression Detection

VII. RESULTS

The system was evaluated on a dataset comprising multimodal inputs from participants. The performance metrics used include accuracy, precision, recall, and F1-score. The results indicate that the multimodal approach outperforms unimodal systems, achieving higher accuracy in detecting depression levels.

VIII. OBJECTIVES

1) Develop an automated system for early detection of depression using multimodal data.
2) Integrate advanced deep learning models for analyzing text, audio, and visual inputs.
3) Enhance the accuracy and reliability of depression detection compared to traditional methods.
4) Provide a user-friendly interface for interaction and feedback.

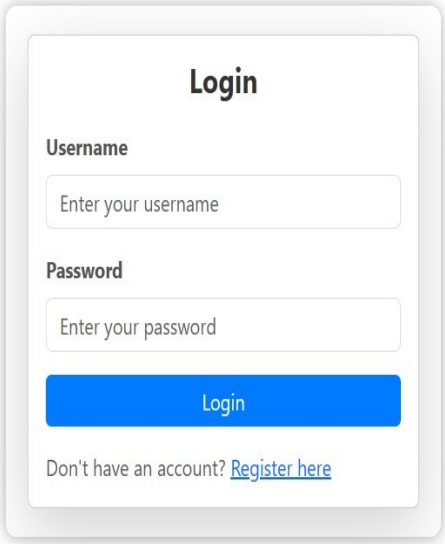


Figure 2: User Login

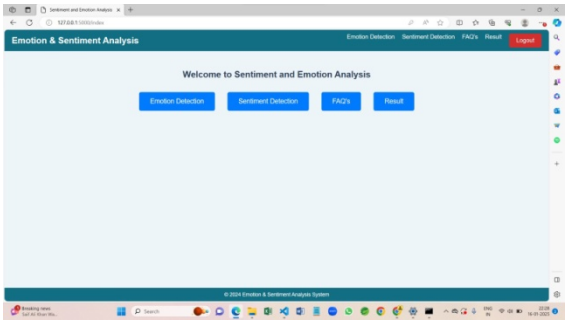


Figure 3: Modules Creation

IX. APPLICATIONS

1) Mental health screening in clinical settings.
2) Continuous monitoring of patients with a history of depression.
3) Integration into telemedicine platforms for remote assessments.
4) Use in educational institutions to monitor student well-being.

X. CONCLUSION

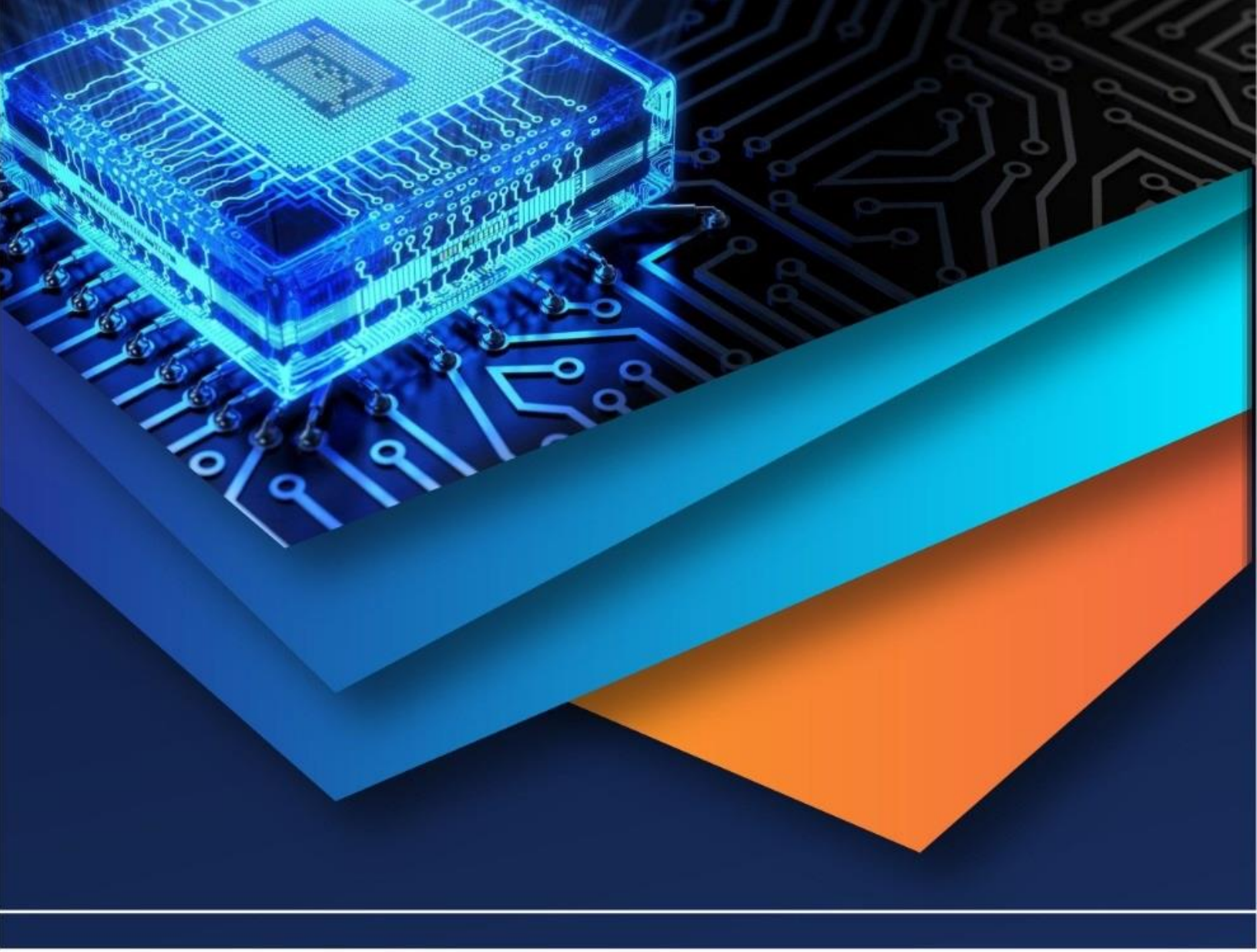
This paper presents a comprehensive system for detecting depression levels by integrating text, audio, and visual data using advanced deep learning techniques. The multimodal approach enhances the accuracy and reliability of depression detection, offering a valuable tool for mental health professionals and individuals. Future work will focus on expanding the dataset, improving model robustness, and integrating additional modalities such as physiological signals.

XI. FUTURE SCOPE

The system can be extended to incorporate physiological data, such as heart rate and galvanic skin response, to further enhance detection accuracy. Additionally, deploying the system on mobile platforms can facilitate widespread accessibility, enabling real-time monitoring and support for individuals in various settings.

REFERENCES

[1] Makiuchi, Y., et al. (2024). Harnessing multimodal approaches for depression detection using the EDAIC dataset. Nature Mental Health, 3(2), 112.
[2] Ray, S., et al. (2024). Using Large Language Models to Detect Depression From User-Generated Diary Text. Journal of Medical Internet Research, 26(1) , e54617.
[3] Screening for Depression Using Natural Language Processing. (2024). International Journal of Medical Research, 12(1), e55067.
[4] Depression Detection from Social Media Text Analysis using Natural Language Processing. (2024). ACM Digital Library.
[5] Lee, Y. S., et al. (2024). DORIS: Detection of Depression via Emotional Text Analysis. IEEE Transactions on Affective Computing.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY