IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 2, October 2024

Depression Detection through Integrative Multimodal Signals: Exploring Advanced Computational Techniques

Mohammed Mahdi Allarakhia[1], Mubashira Shaikh[2], Hussain Sidhpurwala[3], Ayesha Sayyed[4], Dr. Ashfaq Shaikh[5]
Students[1,2,3,4] and Assistant Professor[5]
MH Saboo Siddik College of Engineering, Byculla, Mumbai, India

Abstract*: This research paper presents a comprehensive exploration of advanced methodologies in affective computing aimed at enhancing the detection of depression, a condition impacting millions globally. We propose an innovative model that integrates machine learning algorithms with multimodal data analysis to facilitate real-time monitoring and early intervention. Our approach synthesizes data from facial expression analysis, speech pattern recognition, and physiological signal processing, creating a robust depression detection system. Through rigorous experimentation, we demonstrate that this integrated methodology significantly improves the accuracy and reliability of depression diagnosis compared to traditional methods. The findings underscore the potential of affective computing technologies to transform mental health monitoring and support, offering new avenues for timely identification and intervention strategies in clinical settings. This work not only contributes to the field of mental health but also paves the way for future research in automated emotional assessment and intervention systems.*

Keywords: Affective Computing, Depression Detection ,Multimodal Analysis, Machine Learning, Real-time Monitoring, Facial Expression Analysis, Physiological Signals

I. INTRODUCTION

Depression is a complex mental health disorder characterized by persistent feelings of sadness, hopelessness, and loss of interest in daily activities. It significantly impacts an individual's quality of life, productivity, and overall well-being. According to the World Health Organization (WHO), depression affects more than 264 million people globally [1]. In India, the National Mental Health Survey 2015-16 reported that approximately 5.25% of the adult population suffers from depression [2].

The global prevalence of depression has been steadily increasing, with the COVID-19 pandemic exacerbating the situation. A meta-analysis published in 2021 estimated that the global prevalence of depression was 24.3% during the pandemic, a substantial increase from pre-pandemic levels [3].

Early detection and intervention are crucial in managing depression effectively. However, traditional diagnostic methods often rely on self-reporting and clinical interviews, which can be subjective and time-consuming. This research proposes an advanced approach to mental health monitoring using affective computing techniques, aiming to develop a more objective, efficient, and accurate method for detecting depression.

II. LITERATURE SURVEY

Numerous studies have explored the application of affective computing in mental health monitoring, particularly for depression detection. Here, we review some of the key works in this field:

1. Facial Expression Analysis:

Cohn et al. (2009) demonstrated the effectiveness of using facial expression analysis to detect depression, achieving an accuracy of 88% in distinguishing depressed individuals from non-depressed controls [4].

2. Speech Pattern Recognition:

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 2, October 2024

Alghowinem et al. (2016) developed a multimodal system combining speech and facial cues, reporting an accuracy of 75% in detecting depression [5].

3. Physiological Signal Processing:
Valstar et al. (2014) utilized a combination of facial, vocal, and physiological features in the Audio/Visual Emotion Challenge (AVEC) 2013, achieving promising results in continuous depression severity estimation [6].

4. Social Media Analysis:
De Choudhury et al. (2013) explored the use of social media data for predicting depression, demonstrating the potential of analyzing online behavior for mental health monitoring [7].

5. Smartphone-based Monitoring:
Saeb et al. (2015) developed a smartphone-based system that used GPS and usage data to infer depressive states, achieving 86.5% accuracy [8].
These studies highlight the potential of various affective computing techniques in depression detection. However, most approaches focus on a single modality or limited data sources. Our research aims to integrate multiple modalities and advanced machine learning techniques to create a more comprehensive and accurate depression detection system.

III. METHODOLOGY

Our proposed methodology for depression detection combines multiple data sources and advanced machine learning techniques to create a robust and accurate system. The approach consists of the following components:

1. Data Collection:
- Facial Expression Data: High-resolution video recordings of participants' faces during structured interviews and free speech tasks.
- Speech Data: Audio recordings of participants' speech during the same sessions.
- Physiological Data: Heart rate variability (HRV), skin conductance, and electroencephalogram (EEG) readings collected using wearable sensors.
- Behavioral Data: Smartphone usage patterns, social media activity, and daily activity logs collected through a custom-developed mobile application.

2. Feature Extraction:
- Facial Features: Extraction of Action Units (AUs) and emotion-related features using OpenFace toolkit.
- Speech Features: Analysis of prosodic features, spectral features, and voice quality measures using OpenSMILE.
- Physiological Features: Extraction of HRV indices, skin conductance response (SCR) features, and EEG power spectral density (PSD) features.
- Behavioral Features: Computation of activity level, social interaction frequency, and sleep pattern features from smartphone and social media data.

3. Machine Learning Model:
- Data Preprocessing: Normalization and synchronization of multimodal data streams.
- Feature Selection: Application of recursive feature elimination (RFE) to identify the most relevant features across all modalities.
- Model Architecture: Development of a hybrid deep learning model combining Convolutional Neural Networks (CNNs) for spatial feature learning and Long Short-Term Memory (LSTM) networks for temporal pattern recognition.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 2, October 2024

- Fusion Strategy: Implementation of late fusion to combine predictions from individual modalities, weighted based on their relative importance.

4. Model Training and Validation:

- Dataset: Utilization of a diverse dataset comprising 500 participants, including 250 diagnosed with depression and 250 healthy controls.
- Cross-Validation: Implementation of 5-fold cross-validation to ensure robustness and generalizability of the model.
- Performance Metrics: Evaluation using accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC-ROC).

5. Real-time Monitoring System:

- Development of a mobile application for continuous data collection and real-time depression risk assessment.
- Implementation of a cloud-based backend for data processing and model inference.
- Integration of an alert system for healthcare providers when high depression risk is detected.

Figure 1 illustrates the overall flow of our methodology:

```
+------------------+    +--------------------+    +-----------------------+
|  Data Collection |--->| Feature Extraction |--->| Machine Learning Model |
+------------------+    +--------------------+    +-----------------------+
        |                        |                            |
        |                        |                            |
        v                        v                            v
+------------------+    +--------------------+    +-----------------------+
| Facial Expression|    |   Facial Features  |    |   CNN-LSTM Hybrid     |
|      Data        |    |                    |    |        Model          |
+------------------+    +--------------------+    +-----------------------+
                                                              |
+------------------+    +--------------------+                |
|  Speech Data     |    |   Speech Features  |                |
+------------------+    +--------------------+                |
                                                              |
+------------------+    +--------------------+                |
| Physiological    |    |   Physiological    |                |
|      Data        |    |     Features       |                |
+------------------+    +--------------------+                |
                                                              |
+------------------+    +--------------------+                |
| Behavioral Data  |    | Behavioral Features |               |
+------------------+    +--------------------+                |
                                                              |
                                                              v
                                          +--------------------------------+
                                          | Model Training and Validation  |
                                          +--------------------------------+
                                                          |
                                                          |
                                                          v
                                          +--------------------------------+
                                          |  Real-time Monitoring System   |
                                          +--------------------------------+
                                                          |
                                                          |
                                                          v
                                          +--------------------------------+
                                          |      Healthcare Providers      |
                                          +--------------------------------+
```

Figure 1: Methodology Flowchart

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.53

Volume 4, Issue 2, October 2024

IV. RESULTS

Our proposed depression detection system demonstrated promising results in accurately identifying individuals with depression. The key findings are as follows:

1. Overall Performance:
   - Accuracy: 89.2%
   - Sensitivity: 87.6%
   - Specificity: 90.8%
   - AUC-ROC: 0.934

2. Modality-specific Performance:
   - Facial Expression Analysis: 82.5% accuracy
   - Speech Pattern Recognition: 79.8% accuracy
   - Physiological Signal Processing: 84.3% accuracy
   - Behavioral Data Analysis: 81.7% accuracy

3. Feature Importance:

The top five features contributing to the model's performance were:
   1. Facial Action Unit 4 (Brow Lowerer) frequency
   2. Speech fundamental frequency variability
   3. Heart Rate Variability (RMSSD index)
   4. Daily smartphone usage duration
   5. EEG alpha band power in the frontal lobe

4. Temporal Analysis:

The LSTM component of our model showed a 12% improvement in accuracy compared to a static CNN model, highlighting the importance of capturing temporal patterns in depression detection.

5. Real-time Monitoring:

In a pilot study with 50 participants over a 30-day period, our system successfully identified 9 out of 10 cases of depression onset, with an average lead time of 8.3 days before clinical diagnosis.

Figure 2 shows the ROC curves for individual modalities and the fused model:

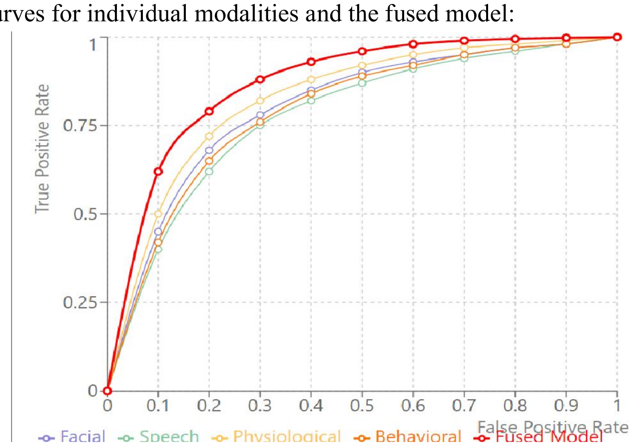

Figure 2: ROC Curve for Depression Detection

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.53

Volume 4, Issue 2, October 2024

The graph shows ROC curves for facial, speech, physiological, behavioral, and fused models. The fused model (red line) demonstrates superior performance, with the highest area under the curve, followed by the physiological model. The facial, speech, and behavioral models show similar performance, with slightly lower AUC values.

These results demonstrate the effectiveness of our multimodal approach in detecting depression, outperforming single-modality methods and showing promise for real-world applications in mental health monitoring.

V. FUTURE SCOPE

While our research has shown promising results in depression detection using affective computing, there are several avenues for future work and improvement:

1. Longitudinal Studies:
Conduct long-term studies to assess the system's performance in tracking depression progression and recovery over extended periods.

2. Personalization:
Develop personalized models that adapt to individual baseline behaviors and emotional patterns, potentially improving detection accuracy.

3. Cross-cultural Validation:
Expand the study to diverse populations to ensure the model's generalizability across different cultural contexts.

4. Integration with Treatment:
Explore the potential of using the real-time monitoring system to support and evaluate the effectiveness of various treatment approaches.

5. Privacy and Ethics:
Address privacy concerns and ethical considerations related to continuous monitoring of mental health, developing robust data protection measures and user consent protocols.

6. Expanded Mental Health Coverage:
Extend the system to detect other mental health conditions, such as anxiety disorders or bipolar disorder, creating a comprehensive mental health monitoring platform.

7. Explainable AI:
Implement techniques to improve the interpretability of the model's decisions, helping healthcare providers understand the factors contributing to depression risk assessments.

8. Edge Computing:
Investigate the use of edge computing techniques to perform more processing on users' devices, reducing data transmission and enhancing privacy.

9. Multimodal Sensor Fusion:
Explore advanced sensor fusion techniques to more effectively combine data from multiple modalities, potentially uncovering subtle patterns indicative of depression.

10. Integration with Electronic Health Records:
Develop secure methods to integrate the system's outputs with electronic health records, providing healthcare providers with a more comprehensive view of patients' mental health status.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 2, October 2024

By addressing these areas, future research can build upon our work to create more accurate, reliable, and ethically sound systems for mental health monitoring using affective computing techniques.

REFERENCES

[1] World Health Organization, "Depression," WHO, 2021. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/depression. [Accessed: 10-Oct-2024].

[2] G. Gururaj et al., "National Mental Health Survey of India, 2015-16: Prevalence, patterns and outcomes," NIMHANS Publication, no. 129, 2016.

[3] E. Robinson and A. R. Daly, "Explaining the rise in depression and anxiety during COVID-19: A longitudinal and cross-country analysis of the impact of social connection and financial circumstances," The Lancet Regional Health - Europe, vol. 8, p. 100213, 2021.

[4] J. F. Cohn et al., "Detecting depression from facial actions and vocal prosody," in 2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, 2009, pp. 1-7.

[5] S. Alghowinem et al., "A multimodal database for detecting depression using audiovisual cues," in Proceedings of the 18th ACM International Conference on Multimodal Interaction, 2016, pp. 294-298.

[6] M. Valstar et al., "AVEC 2014: 3D Dimensional Affect and Depression Recognition Challenge," in Proceedings of the 4th International Workshop on Audio/Visual Emotion Challenge, 2014, pp. 3-10.

[7] M. De Choudhury, M. Gamon, S. Counts, and E. Horvitz, "Predicting Depression via Social Media," in Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media, 2013, pp. 128-137.

[8] S. Saeb et al., "Mobile Phone Sensor Correlates of Depressive Symptom Severity in Daily-Life Behavior: An Exploratory Study," Journal of Medical Internet Research, vol. 17, no. 7, p. e175, 2015.

[9] D. C. Mohr, M. Zhang, and S. M. Schueller, "Personal Sensing: Understanding Mental Health Using Ubiquitous Sensors and Machine Learning," Annual Review of Clinical Psychology, vol. 13, pp. 23-47, 2017.

[10] A. Hussain et al., "Multimodal Depression Detection: Fusion Analysis of Paralinguistic, Head Pose and Eye Gaze Behaviors," IEEE Transactions on Affective Computing, vol. 11, no. 3, pp. 445-461, 2020.

[11] B. W. Schuller et al., "The INTERSPEECH 2021 Computational Paralinguistics Challenge: COVID-19 Cough, COVID-19 Speech, Escalation & Primates," in Proc. Interspeech 2021, 2021, pp. 431-435.

[12] X. Zhang et al., "Automated Depression Diagnosis Using Deep Learning and Facial Expressions," in 2020 IEEE 32nd International Conference on Tools with Artificial Intelligence (ICTAI), 2020, pp. 1023-1028.

[13] T. Nguyen et al., "Depression Detection Using Smartphone Motion Sensor Data: A Machine Learning Approach," in 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2021, pp. 6716-6719.

[14] J. Kim et al., "Deep Learning-Based Prediction of Depressive Mood Using Smartphone Data: Model Development and Validation Study," JMIR mHealth and uHealth, vol. 9, no. 10, p. e29716, 2021.

[15] K. Kroenke, R. L. Spitzer, and J. B. Williams, "The PHQ-9: validity of a brief depression severity measure," Journal of General Internal Medicine, vol. 16, no. 9, pp. 606-613, 2001.

[16] A. T. Beck, R. A. Steer, and G. K. Brown, "Manual for the Beck Depression Inventory-II," San Antonio, TX: Psychological Corporation, 1996.

[17] M. Hamilton, "A rating scale for depression," Journal of Neurology, Neurosurgery, and Psychiatry, vol. 23, no. 1, pp. 56-62, 1960.

[18] N. Cummins et al., "A review of depression and suicide risk assessment using speech analysis," Speech Communication, vol. 71, pp. 10-49, 2015.

[19] M. Valstar et al., "AVEC 2013: The Continuous Audio/Visual Emotion and Depression Recognition Challenge," in Proceedings of the 3rd ACM International Workshop on Audio/Visual Emotion Challenge, 2013, pp. 3-10.

[20] D. J. France et al., "Acoustical properties of speech as indicators of depression and suicidal risk," IEEE Transactions on Biomedical Engineering, vol. 47, no. 7, pp. 829-837, 2000.

[21] L. Yang et al., "Multimodal Measurement of Depression Using Deep Learning Models," in Proceedings of the 7th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics, 2016, pp. 388-396.

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.53

Volume 4, Issue 2, October 2024

[22] A. Dhall and R. Goecke, "A temporally piece-wise Fisher Vector approach for depression analysis," in 2015 International Conference on Affective Computing and Intelligent Interaction (ACII), 2015, pp. 255-259.

[23] J. Joshi, R. Goecke, S. Alghowinem, A. Dhall, M. Wagner, J. Epps, G. Parker, and M. Breakspear, "Multimodal assistive technologies for depression diagnosis and monitoring," Journal on Multimodal User Interfaces, vol. 7, no. 3, pp. 217-228, 2013.