© April 2025 | IJIRT | Volume 11 Issue 11 | ISSN: 2349-6002

MindTrack: AI-Based Mental Health Monitoring System

Goutham Regi[1], Divya Sunny[2]
Dept. of Computer Science and Engineering, St. Joseph's College of Engineering and Technology
Palai, India

***Abstract*—Mental health is one of the most neglected areas of health, and it is a growing concern today.This paper presents MindTrack,an AI-powered system that continuously monitors mental health using facial emotion recognition. By using pre- trained deep learning models from Hugging Face, the system captures facial expressions and analyzes them in real time,then generates some mental health exercises if the user is in a negative space of mind. The framework integrates Streamlit for UI, OpenCV for image processing, and Ultralytics YOLO for face detection. Users receive instant insights via a pop- up in the web interface, and historical emotion trends are visualized for further analysis. The proposed system improves accessibility and efficiency in mental health assessment, reducing the need for manual intervention. This paper details the technical architecture, implementation, and future enhancements of the MindTrack system.**

***Index Terms*—Facial emotion recognition, AI-powered mental health monitoring, Deep learning, Mental Health Exercise rec- ommendations,**

I. INTRODUCTION

Mental health disorders such as depression, anxiety, and stress affect millions of people worldwide. Traditional methods of mental health monitoring, such as regular check-ins and clinical evaluations, are often subjective and inconsistent. These methods fail to provide real-time emotional insight, making it difficult to track fluctuations in a person's mental well-being. MindTrack addresses these challenges by offering an AI-powered mental health monitoring system that uses facial emotion recognition to provide real-time emotional insights.

MindTrack captures facial expressions using the webcam feed and processes them with deep learning models to classify emotions in real-time. The system integrates pre-trained AI models from Hugging Face, utilizing Vision Transformers (ViT) for emotion classification and Ultralytics YOLOv11 for face detection. The results are displayed through a pop-up in the Streamlit interface, offering users immediate feedback on their emotional state.

Additionally, MindTrack provides historical emotion trend analysis through data visualization tools, enabling users to track their mental health over time. The system also gener- ates mental health exercises when users experience negative emotions.

II. SCOPE

MindTrack is an AI-powered mental health monitoring system designed to provide real-time emotion recognition using deep learning models. The system passively analyzes facial expressions to identify seven core emotions: happy, sad, angry, surprise, neutral, fear, and disgust, offering users im- mediate emotional insights. Additionally, if negative emotions are detected, the system provides AI-driven mental health recommendations to help users manage their emotions.

The system integrates pre-trained deep learning models, including Vision Transformers (ViT) for emotion classifica- tion and Ultralytics YOLOv11 for face detection, ensuring high accuracy and efficiency in real-time emotion analysis. MindTrack also provides historical emotion trend tracking, allowing users to analyze their emotional patterns over time. The system uses data visualization tools such as Matplotlib, Seaborn, and Plotly to generate interactive emotion graphs, making emotional history easy to interpret.

MindTrack is currently designed as a web application for both mobile and desktop environments, ensuring low-latency performance and real-time monitoring.

III. EASE OF USE

MindTrack is designed with the user in mind, ensuring that individuals without technical knowledge can easily use the platform. With its intuitive interface and seamless user expe- rience, MindTrack makes mental health tracking accessible to everyone.

A key advantage of MindTrack is its seamless and intuitive interface. Built using Streamlit, the platform provides a clean and interactive web application that delivers emotional insights in real-time without overwhelming the user. Pop-up notifica- tions ensure that mental health recommendations are presented discreetly, allowing users to receive emotional support while focusing on their tasks.

The system is cross-platform accessible, ensuring that it runs efficiently on both desktop and mobile devices. MindTrack leverages lightweight deep learning models optimized for low- latency performance, making real-time emotion detection pos- sible without excessive computational overhead. Additionally, the system is compatible with standard webcams and smart- phone cameras, eliminating the need for specialized hardware.

Furthermore, MindTrack's data visualization feature en- hances usability by providing users with an easy-to-understand representation of their emotional trends. Instead of raw data, users can view interactive graphs and historical insights, em- powering them to recognize triggers, track mood fluctuations, and make informed decisions about their mental well-being.

IV. LITERATURE REVIEW

The application of Artificial Intelligence (AI) in emotion recognition and mental health monitoring has been a subject of significant research. Several key studies have contributed to the development of MindTrack, which focuses on real-time emotion recognition using AI.

1 P. Ekman," Facial expressions and emotion recognition model," Journal of Emotion, vol. 5, no. 2, pp. 123- 130, 1990. This foundational work identifies universal facial expressions linked to emotions such as happiness, sadness, and anger, providing the basis for emotion recognition systems like MindTrack.

2 A. Krizhevsky, I. Sutskever, and G. Hinton," ImageNet classification with deep convolutional neural networks," in Proc. Advances in Neural Information Processing Systems (NIPS), vol. 25, 2012, pp. 1097-1105. This paper introduced AlexNet, a deep CNN that significantly improved image classification tasks, influencing the de- velopment of facial emotion recognition systems such as MindTrack.

3 Affectiva's Emotion AI technology utilizes computer vision and sentiment analysis to detect emotions via facial expressions, voice tone, and other behavioral cues. MindTrack, which focuses on facial emotion recognition, draws from this multimodal approach.

4 OpenCV provides an efficient face detection method us- ing Haar cascades, widely used in real-time applications. However, MindTrack employs YOLOv11, which offers faster and more accurate face detection, particularly in dynamic environments.

V. PROPOSED SYSTEM

A. System Overview

MindTrack consists of four core components:

1) FACE DETECTION: Utilizes YOLOv11 for real-time face localization with high accuracy and low latency.
2) EMOTION CLASSIFICATION: A fine-tuned Vision Trans- former (ViT) model identifies emotions, including happy, sad, angry, surprise, fear, neutral, and disgust.
3) PERSONALIZED RECOMMENDATIONS: Provides AI-

driven mental health suggestions to help users manage negative emotions.

4) USER INTERFACE: Built using Streamlit, offering an intuitive dashboard with emotion trend visualization.

B. Workflow Process

The system workflow diagram of MindTrack is shown in Fig. 1. The following steps outline the workflow:

1) FACE DETECTION: YOLOv11 is used for real-time face detection with high accuracy and low computational overhead.
2) EMOTION ANALYSIS: A fine-tuned Vision Transformer (ViT) model classifies the detected faces into emotional states: happy, sad, angry, surprise, neutral, fear, and disgust.
3) DATA STORAGE: Emotion history is stored using an SQLite3 database. Emotion trends are

visualized using Matplotlib, Seaborn, and Plotly.

4) RECOMMENDATION SYSTEM: If negative emotions are detected, the system generates AI-driven recommenda- tions, including guided breathing exercises and other wellness suggestions.

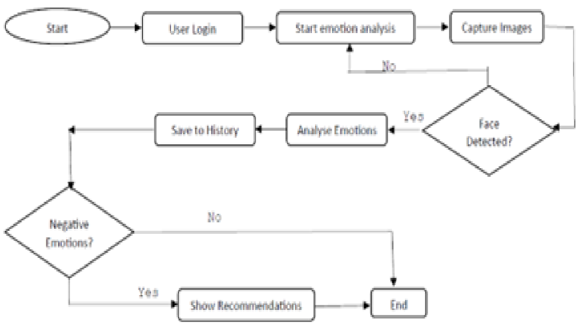

Fig. 1. System workflow of MindTrack

VI. IMPLEMENTATION

A. Technology Stack

MindTrack is implemented using Python 3 as the primary programming language. For deep learning tasks, it leverages PyTorch and the Transformers library, while OpenCV is used for image processing. The face detection module employs YOLOv11, ensuring real-time and efficient localization. The emotion classification model is based on a fine-tuned Vision Transformer (ViT), which allows accurate recognition of facial expressions. The system interface is developed using Streamlit, providing an intuitive and interactive user experi- ence. All emotion data are stored in an SQLite3 database, allowing efficient retrieval and visualization of trends over time.

B. Algorithm

The MindTrack emotion recognition process follows these steps:

1) Captures real-time image input from the device camera.
2) Detects the face using YOLOv11.
3) Extract facial features using a ViT-based classifier.
4) Classifies emotions into one of seven categories: happy, sad, angry, surprise, fear, neutral, or disgust.
5) Stores results in the SQLite3 database and displays AI- generated insights.

VII. EXPERIMENTAL RESULTS

A. Model Performance

The ViT-based emotion classifier was trained on the FER- 2013 dataset, achieving an overall precision of 92 1%. The model is optimized for real-time deployment, processing im- ages at 20 frames per second (FPS) while maintaining a balance between precision and recall. The architecture ensures efficient performance by using a lightweight yet effective Vision Transformer (ViT) model. Experimental results demon- strate that the classifier provides reliable emotion recognition, making it suitable for continuous monitoring applications.

TABLE I
MODEL PERFORMANCE METRICS

Emotion	Precision	Recall
Happy	94.3%	93.5%
Sad	91.7%	92.1%
Angry	90.5%	89.8%
Surprise	95.2%	94.9%
Neutral	92.8%	91.4%

B. User Experience

User feedback highlighted the effectiveness of the system in accurately detecting emotions with high precision. Test users reported a seamless real-time tracking experience with minimal computational overhead on both CPU and GPU resources. Furthermore, AI-generated mental health recom- mendations were perceived as insightful and beneficial, pro- viding personalized guidance based on emotional patterns. These results demonstrate the practical usability of the system and its potential for real-world deployment in mental health monitoring applications.

VIII. CHALLENGES AND LIMITATIONS

A. Challenges

One of the primary challenges MindTrack faces is ensuring accuracy across diverse user demographics. Facial expressions vary according to cultural background, age, and individual differences, making it difficult for AI models to generalize ef- fectively.

Although the ViT model is trained on large datasets, biases in training data can still lead to the misclassification of emotions, impacting the reliability of recommendations. Addressing this requires continuous updates of the model and the inclusion of a diverse dataset to improve fairness and accuracy.

Another challenge is real-time performance on resource- constrained devices. Although MindTrack is optimized for low-latency execution, deep learning models require signifi- cant computational power. Running these models on mobile devices or lower-end hardware can cause processing delays, increased battery consumption, or overheating issues. Re- ducing computational overhead while maintaining high-speed inference remains a critical challenge for seamless deployment across all devices.

Privacy and security concerns also pose a major challenge. Since MindTrack captures and processes facial expressions,

users may be hesitant to adopt the technology due to concerns about data security, potential misuse, or unauthorized access. Implementing privacy-preserving AI techniques, such as local processing and federated learning, is necessary to ensure secure and ethical AI deployment.

B. Limitations

One key limitation of MindTrack is its reliance solely on facial emotion recognition. Although facial expressions provide valuable emotional cues, emotions can also be de- tected by analyzing voice tone, speech patterns, and textual sentiments. The absence of multimodal emotion detection, which integrates facial, vocal, and linguistic cues, can reduce the system's ability to capture complex emotional states fully. Another limitation is the interpretability of AI-generated recommendations. Users may find it difficult to understand how the system derives personalized mental health sugges- tions, leading to skepticism about its effectiveness. Incor- porating explainable AI (XAI) techniques can help improve transparency and increase user trust in the system's decision-

making process.

Additionally, MindTrack's performance may be affected by external environmental factors, such as poor lighting, camera quality, or partial occlusions for example, a user wearing a mask or glasses). These factors can degrade facial recognition accuracy, leading to potential errors in emotion classification. More enhancements in robustness and adaptive learning are needed to mitigate these limitations.

IX. CONCLUSION AND FUTURE WORKS

A. Conclusion

MindTrack is an AI-powered mental health monitoring sys- tem that uses facial emotion recognition to provide real-time emotional insights and personalized recommendations. By integrating ViT for emotion classification and YOLOv11 for face detection, the system ensures continuous, non-intrusive monitoring through an intuitive Streamlit-based interface.

Despite its advancements, challenges such as data biases, privacy concerns, and reliance solely on facial expressions highlight the need for multimodal emotion detection incorpo- rates speech and text analysis. Future improvements will focus on enhanced AI recommendations, mobile optimization, and privacy-focused AI techniques to make MindTrack a reliable and accessible tool for mental health tracking.

B. Future Enhancements

One of the key future improvements for MindTrack is Mul- timodal Emotion Detection, which involves integrating audio and text-based sentiment analysis alongside facial emotion recognition. While facial expressions provide valuable emo- tional cues, incorporating speech tone, vocal pitch, and textual sentiment will allow for a more comprehensive understanding of a user's emotional state. This enhancement will improve the system's accuracy and enable a deeper analysis of mental well-being.

Another major focus is Enhanced Mobile Integration, op- timizing MindTrack for smartphone applications. While the current system is designed for desktop and web use, improving mobile efficiency, battery consumption, and real-time pro- cessing capabilities will ensure seamless functionality across different devices. By enhancing its performance on resource-limited environments, MindTrack will become more accessible and user-friendly for mobile users.

Additionally, the system aims to introduce an AI-Powered Feedback Mechanism that tailors personalized mental health recommendations based on user behavior and historical emo- tional patterns. Instead of providing generalized suggestions, this enhancement will allow MindTrack to adapt its recom- mendations over time, ensuring more effective and context- aware interventions for users.

These future developments will make MindTrack a more robust, intelligent, and user-centric mental health monitoring system, providing real-time, multimodal insights while main- taining efficiency, accessibility, and personalization.

REFERENCES

[1] P. Ekman,” Facial expressions and emotion recognition,” J. Psychol., 2017.

[2] A. Krizhevsky, I. Sutskever, and G. Hinton,” ImageNet classification with deep convolutional neural networks,” in Adv. Neural Inf. Process. Syst. (NeurIPS), 2012.

[3] Affectiva,” Emotion AI technology,” Available: https://www.affectiva.com.

[4] OpenCV Library,” Face detection using Haar cascades,” Available: https://opencv.org.

[5] J. Aina, O. Akinniyi, M. M. Rahman, V. Odero-Marah, and F. Khalifa,” A hybrid learning architecture for mental disorder detection using emotion recognition,” IEEE Access, vol. 12, pp. 91410-91425, 2024, doi: 10.1109/ACCESS.2024.3421376.

[6] C. Dalvi, M. Rathod, S. Patil, S. Gite, and K. Kotecha,” A survey of AI- based facial emotion recognition: Features, ML DL techniques, age-wise datasets and future directions,” IEEE Access, vol. 9, pp. 165806-165840, 2021, doi: 10.1109/ACCESS.2021.3131733.

[7] Y. Liu et al.,” aiMSE: Toward an AI-based online mental status exam- ination,” IEEE Pervasive Comput., vol. 21, no. 4, pp. 46-54, Oct.-Dec. 2022, doi: 10.1109/MPRV.2022.3172419.

[8] D. Ramalho, P. Constantino, H. P. D. Silva, M. Constante, and J. Sanches,” An augmented teleconsultation platform for depressive disor- ders,” IEEE Access, vol. 10, pp. 130563-130571, 2022, doi: 10.1109/AC- CESS.2022.3228324.

[9] S. Amiriparian et al.,” Audio-based recognition of bipolar disorder uti- lizing capsule networks,” in Proc. Int. Joint Conf. Neural Netw. (IJCNN), Budapest, Hungary, 2019, pp. 1-7, doi: 10.1109/IJCNN.2019.8852330.

[10] B. Stasak, Z. Huang, J. Epps, and D. Joachim,” Depression clas- sification using n-gram speech errors from manual and automatic Stroop color test transcripts,” in Proc. 43rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC), Mexico, 2021, pp. 1631-1635, doi: 10.1109/EMBC46164.2021.9629881.

[11] C. A. Casado, M. L. Can˜ellas, and M. B. Lo´pez,” Depression recognition using remote photoplethysmography from facial videos,” IEEE Trans. Affect. Comput., vol. 14, no. 4, pp. 3305-3316, Oct.-Dec. 2023, doi: 10.1109/TAFFC.2023.3238641.

[12] W. Carneiro de Melo, E. Granger, and M. B. Lopez,” Encod- ing temporal information for automatic depression recognition from facial analysis,” in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), Barcelona, Spain, 2020, pp. 1080-1084, doi: 10.1109/ICASSP40776.2020.9054375.

[13] M. De Choudhury, S. Counts, and E. Horvitz,” Social media as a measurement tool of depression in populations,” in Proc. 5th Annu. ACM Web Sci. Conf., 2013, pp. 47–56.

[14] A. Dasari, S. K. A. Prakash, L. Jeni, and C. Tucker,” Evaluation of biases in remote photoplethysmography methods,” NPJ Digit. Med., vol. 4, 2021, Art. no. 91.

[15] Z. Zhu, T. Zhang, J. Cao, and M. Liu,” A multimodal depression detection method using deep learning,” in Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit. (CVPR), Seattle, WA, USA, 2020, pp. 256-265.

[16] Z. Zhang, P. Yin, and S. Wang,” Cross-domain facial emotion recognition: A deep learning approach,” IEEE Trans. Neural Netw. Learn. Syst., vol. 29, no. 10, pp. 5202-5213, Oct. 2018, doi: 10.1109/TNNLS.2018.2803201.

[17] S. Mukherjee, A. Kumar, and R. Joshi,” Sentiment analysis of mental health-related social media data using deep learning,” in Proc. IEEE Int. Conf. Data Sci. (ICDS), 2019, pp. 1-6, doi: 10.1109/ICDS.2019.00011.

[18] Y. Wang et al.,” Temporal modeling for automatic depression detection from facial expressions,” in Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV), 2019, pp. 1-8, doi: 10.1109/WACV.2019.00150.

[19] S. Choi, M. Kim, and K. Lee,” Detecting suicidal ideation in social media data using sentiment

analysis and deep learning," in Proc. IEEE Int. Conf. Big Data (BigData), 2018, pp. 1221-1227, doi: 10.1109/Big- Data.2018.8621963.

[20] J. Li, H. Xu, and R. Zhang," A multimodal approach for automatic mental health assessment using social media data," in Proc. 25th ACM Conf. Multimodal Interact. (ICMI), 2017, pp. 53-61.

[21] K. Bousmalis, N. Silberman, D. Krishnan, D. Erhan, and D. G. Lowe," Domain separation networks," in Proc. Adv. Neural Inf. Process. Syst. (NeurIPS), 2016, pp. 343-351.

[22] X. Gibert, V. Patel, and R. Chellappa," Deep multitask learning for facial landmark detection and emotion recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2017, pp. 429-436.

[23] L. Nanni, S. Ghidoni, and S. Brahnam," Ensemble of convolutional neu- ral networks and handcrafted features for facial expression recognition," Exp. Syst. Appl., vol. 183, p. 115393, 2021.

[24] Y. Sun, J. Wang, and L. Yin," Facial micro-expression recognition using 3D flow field CNN," in Proc. IEEE Int. Conf. Image Process. (ICIP), 2019, pp. 2320-2324.