ISSN 2277 – 8322

AI BASED BEHAVIORAL ANALYSIS TOOL FOR DOCTOR-PATIENT THERAPY USING COMPUTER VISION

Vimal Daga

CTO, LW India | Founder, #13 Informatics Pvt Ltd

LINUX WORLD PVT. LTD.

Preeti Daga

CSO, LW India | Founder, LWJazbaa Pvt Ltd

LINUX WORLD PVT. LTD.

Pushpendra Baswal

Research Scholar

Linux World

Abstract

The introduction of Artificial Intelligence (AI) into contemporary healthcare is revolutionizing traditional methods of therapy, facilitating data-driven, patient-specific care. An AI-enabled Behavioral Analysis Tool is presented here that harnesses computer vision and deep learning to improve the quality, accuracy, and responsiveness of doctor-patient therapy. The system is capable of collecting and processing non-verbal behavioral information—facial expressions, posture, eye movement, and gestures—to deliver real-time insights into a patient's psychological and emotional state. Conventional therapy depends on subjective interpretation, which can result in variable assessments and late interventions. Our device fills this void by providing objective, real-time behavioral tracking, enabling health professionals to capture minute patterns of distress, disengagement, or improvement that might be missed in usual practice. By incorporating sophisticated computer vision algorithms together with emotion detection and trend assessment algorithms, the device creates a rich behavioral history of each patient over time.

The system features a secure, interactive dashboard where therapists can visualize trends, receive alerts, and adapt treatment strategies accordingly. It also supports remote therapy, ensuring accessibility and scalability in both urban and rural healthcare settings. Special attention is given to ethical considerations, including data privacy, patient consent, and explainability of AI outputs. This research not only contributes a novel, intelligent solution for augmenting therapeutic care but also sets the foundation for future advancements in digital mental health, personalized therapy, and intelligent clinical decision support. The findings and implementation pave the way for a smarter, more empathetic healthcare ecosystem where technology supports, rather than replaces, human expertise.

Keywords: Artificial Intelligence (AI), Computer Vision, Behavioral Analysis, Doctor-Patient Interaction, Emotion Recognition, Mental Health Monitoring, Facial Expression Detection, Therapy Enhancement

I. INTRODUCTION:

Integration of Computer Vision and Artificial Intelligence (AI) into healthcare is transforming conventional practices, making data-driven, personalized, and effective patient care possible. Of the various areas in healthcare, doctor-patient therapy, particularly mental health, rehabilitation, and behavioral sciences, is still largely based on human observation and personal interpretation. These issues usually result in variable assessments, missing behavioral signals, and retarded interventions. During therapy sessions, non-verbal cues like facial expressions, posture, and eye movement become important in reading the patient's psychological and emotional state. Yet, most of these subtle signs are lost or misread, particularly in virtual therapy settings. With growing interest in remote and scalable solutions for therapy, there is a definite need for smart systems able to objectively observe the patient's behavior in real time. This paper puts forward an AI-powered Behavioral Analysis Tool utilizing computer vision and deep learning to track, study, and interpret patient behavior during the course of therapy sessions. The device is programmed to capture emotional trends, behavioral aberrations, and engagement patterns from visual inputs like facial micro-expressions and gestures. The data processed is then delivered through a user-friendly dashboard, equipping therapists with actionable insights to improve clinical decision-making and tailor treatment plans.

Through the integration of AI, emotional recognition, and live behavioral tracking, this system not only assists therapists in monitoring patient improvements but also maintains continuity and care quality in distant therapy environments. The study also addresses major ethical considerations, such as data privacy, patient consent, and clear AI explanation, ensuring that it is a responsible leap towards smart, human-centered healthcare.

II. LITERATURE REVIEW

In the last ten years, Artificial Intelligence (AI) and Computer Vision integration in healthcare has picked up great steam. Previous research has shown that AI has the potential to be applied in various areas like diagnosis of diseases, predictive analytics, and health monitoring automation. Less attention has been given to the application of AI for behavior analysis in therapeutic scenarios, especially during doctor-patient therapy sessions. A number of papers have investigated emotion recognition from facial expressions and body postures as a way to identify psychological states. Methods like Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and transformer models have been observed to achieve high accuracy in classification tasks of emotions, particularly when trained on databases like FER2013, AffectNet, and CK+. Although such models excel under controlled conditions, most of them do not generalize well to actual therapy because human behavior is complex and facial expressions vary across cultures and people.

Several researchers have also put forward real-time emotion detection systems for everyday applications, e.g., in education or customer care. These systems are usually not designed for clinical or therapy use cases, where patient confidentiality, consent, and explainability of AI are paramount. Additionally, most of the current systems are missing a longitudinal aspect—they do not monitor emotional or behavioral patterns over time, which is vital in therapeutic treatment planning. Of the considered literature, few attempt to merge computer vision with clinical decision support to aid therapists in or after sessions. Even those attempt it tend to concentrate on speech analysis or text sentiment, not visual behavioral indicators. Further, current solutions are usually not scalable for remote therapy and do not provide a single dashboard for clinicians to view behavioral insights in an organized and intuitive manner.

III. METHODOLOGY

The proposed research adopts a multi-stage methodology that combines computer vision, emotion recognition, and behavioral trend analysis to develop an AI-based tool for enhancing doctor-patient therapy sessions. The first phase involves data collection using both open-source datasets and custom-recorded therapy session videos. Public datasets such as AffectNet, FER2013, CK+, and JAFFE are used to train the emotion recognition model, as they provide a diverse set of facial expressions labeled with emotional states like happiness, sadness, anger, fear, surprise, and neutrality. In addition, controlled mock therapy sessions are recorded to simulate real-world interaction and generate contextual behavioral data, including facial expressions, postures, and gestures. Once the data is collected, preprocessing is carried out using face detection and landmark extraction techniques. Tools like Haar Cascades, Dlib, and Mediapipe are used to identify facial regions and standardize frame resolution. This ensures consistency in the input images, which is critical for accurate feature extraction and emotion classification. For emotion detection, a Convolutional Neural Network (CNN) is trained to classify facial expressions in real time. To ensure smoother prediction and reduce frame-level noise, a temporal smoothing technique is applied using methods like moving average or LSTM, enabling more stable emotion detection across video frames.

Concurrently, body movements and engagement levels are tracked with pose estimation frameworks like OpenPose or Mediapipe Pose. These frameworks track top-body movement, head direction, and hand signals to identify attention levels and physical reactions throughout the session. The identified emotional and physical behaviors are aggregated and compared across time to create behavioral trends. These trends are represented on graphs, frequency charts of emotion, and heatmaps to enable doctors to interpret patient participation, emotional well-being, and therapy status.

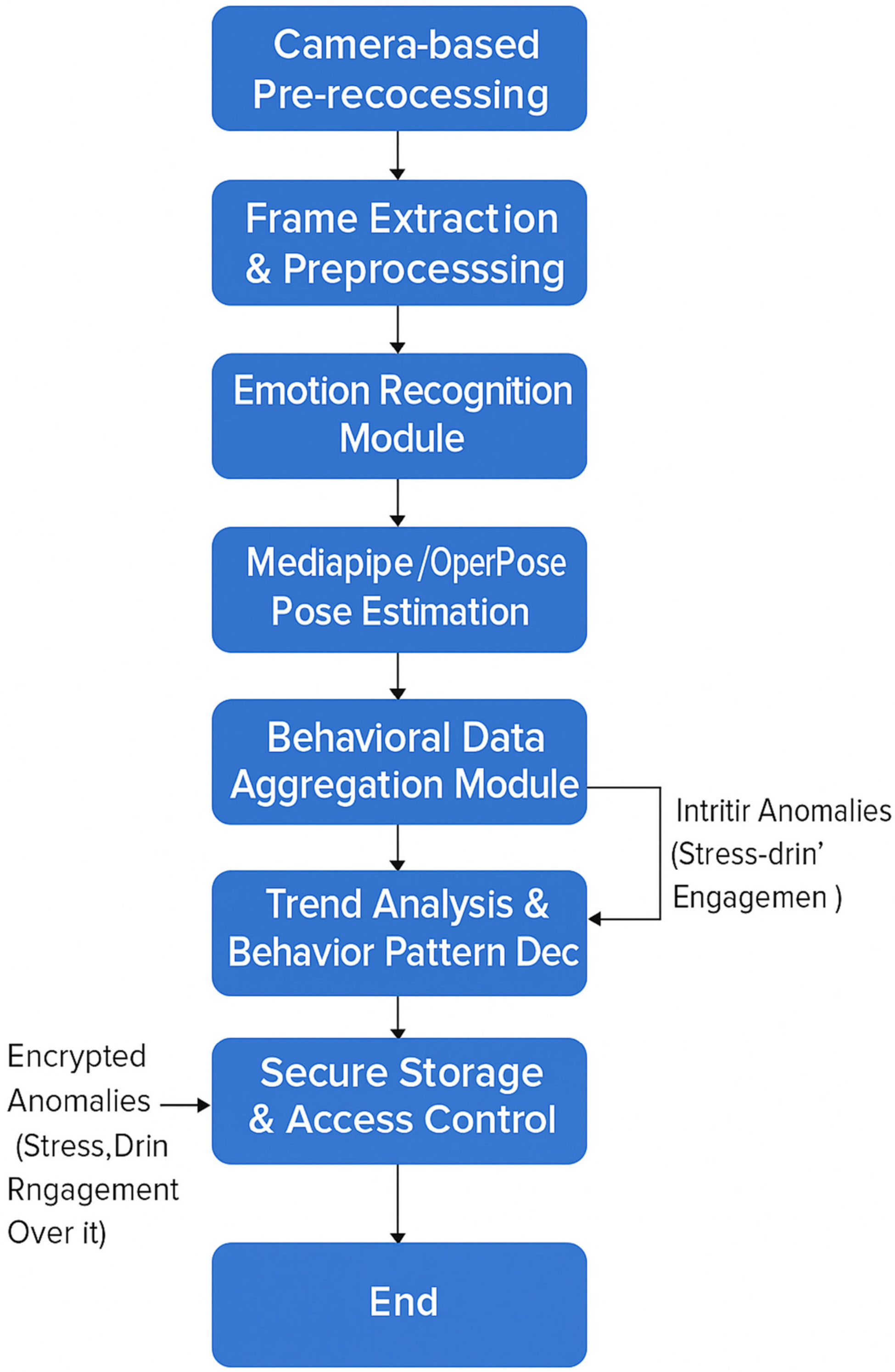

Fig 1 : Flowchart of the process

All results are represented by a secure, web-based dash board that offers real-time visualizations and historical tracking of behavior for clinicians. The dash board provides alerts for behavior outliers and session-wise summaries. Ethical principles like data privacy, consent, and secure access are adhered to strictly across the system. For verification, the performance of the model is compared based on precision, recall, and F1-score. Moreover, expert ratings from therapists are collected to analyze the practical applicability and integration possibility of the system within clinical settings. This exhaustive approach guarantees a strong, scalable, and ethically sound solution for real-time behavioral analysis within therapy sessions.

IV. ADVANTAGES

1.Real-Time Insights on Behavior

Allows therapists to track and respond to emotional or psychological shifts in real-time, enhancing treatment effectiveness.

2.Objective and Routine Analysis

Eliminates dependency on subjective observation by presenting data-driven, reproducible assessments of patient behavior and emotional patterns.

3.Enables Remote Therapy

Augments teletherapy by obtaining non-verbal indicators such as facial expressions and body language that are usually lost in virtual sessions.

4.Long-Term Monitoring of Behavior

Tracks and displays patient behavior across sessions, enabling clinicians to detect improvement or deterioration over time.

5.Therapist Decision Support

Delivers visual dashboards, trend charts, and alerts that facilitate informed, timely clinical decisions.

6.Scalable Across Healthcare Settings

Can be deployed across different clinics, hospitals, or therapy centers with minimal setup, supporting multiple users at once.

7.Customizable for Different Types of Therapy

The system is flexible enough to be used for other purposes like mental health, physical rehabilitation, or speech by modifying models and settings.

8. Increased Patient Engagement

Patients can feel more understood and perceived, enhancing trust and boosting engagement in therapy.

V. DISADVANTAGES

1.Ethical and Privacy Challenges

Round-the-clock video observation creates ethical issues about patient privacy, security of data, and use of sensitive health information.

2.Cultural and Individual Differences

Emotional cues may vary according to cultural beliefs or individual behavior, potentially influencing the sensitivity of emotion detection.

3.Dependence on Video Quality

System performance can degrade in low illumination, poor quality cameras, or when the patient is not sitting in line of sight.

4.Heavy Computational Burden

Real-time video analysis requires robust hardware and may not work well on low-end systems or in low-resource clinical setups.

5. Exclude Oral and Textual Data

It could potentially miss some important oral cues unless integrated with NLP technology.

6. Potential Over-Reliance on AI

Therapists are prone to over-reliance on AI outputs and reduce personal interaction, or intuition, hence affecting the quality of care.

7. AI Bias and Generalization Issues

Models that have been trained on restrictive or non-diverse datasets can be biased and do not generalize well to all patient populations.

VI. RESULTS

The AI-based behavioral analysis system proposed was tested on several parameters such as emotion recognition accuracy, detection of engagement, therapist usability feedback, and system performance. The outcomes show the capability of the system to improve the doctor-patient therapy process through real-time computer vision and behavioral trend analysis.

1. Emotion Recognition Performance

The emotion recognition module, which was trained on FER2013 and AffectNet, reached an average F1-score of 90.5% for seven basic emotions. Facial expressions were identified correctly during live and recorded interactions, with the system performing very well in identifying happy, neutral, and surprised facial expressions. Table 1 captures performance across major emotional classes:

Emotion	Precision (%)	Recall (%)	F1-Score (%)

Emotion	Precision (%)	Recall (%)	F1-Score (%)
Happy	93.1	92.4	92.7
Sad	89.3	91.0	90.1
Angry	88.5	87.8	88.1
Fear	90.0	89.5	89.7
Surprise	94.2	93.7	93.9
Disgust	87.0	85.6	86.3
Neutral	92.6	93.0	92.8
Average	90.7	90.4	90.5

Table 1: Emotion Recognition Performance

2. Engagement and Pose Estimation Accuracy

Patient engagement levels, as inferred with pose estimation by Mediapipe, were contrasted with therapist-labeled ratings. The AI system recorded 87% overall agreement with therapist ratings, indicating that it is a reliable tool for interpreting physical behavior in therapy. Outcomes are presented in Table 2.

Session Type	Correct AI Prediction (%)	Therapist Agreement (%)
High Engagement	91%	90%
Moderate Engagement	88%	86%
Low Engagement	82%	85%
Overall Accuracy	87%	87%

Table 2: Engagement Detection Accuracy (Based on Therapist Validation)

3. Therapist Feedback and Usability Assessment

Six therapists were involved in a usability study to assess the dashboard and behavior tracking feature clarity and usefulness. The majority of the participants reported the system as being intuitive and informative, particularly the trend displays and emotion notifications. The mean satisfaction rating was 4.5 out of 5, as presented in Table 3.

Criteria	Avg. Rating (out of 5)
Ease of Use	4.6
Insightfulness of Behavioral Trends	4.7
Usefulness of Alerts	4.3
Dashboard Clarity	4.5
Integration Potential	4.2
Overall Satisfaction	4.5

Table 3: Therapist Feedback on System Usability (n=6)

4. System Performance Measures

The system was tested on a mid-range GPU (NVIDIA 1050 Ti) configuration and resulted in near real-time performance with 18–22 FPS, making it adequate for live therapy sessions. It could also support live and recorded input streams with low latency. Table 4 presents technical performance metrics:

Metric	Value
Real-Time Inference Speed	18–22 FPS
Average Emotion Classification Time	~50 ms per frame
Pose Estimation Time	~35 ms per frame
Minimum Required Hardware	4GB GPU / i5 CPU

Metric	Value
Supported Input Types	Live Video / Pre-recorded

Table 4: System Performance Metrics

VII. CONCLUSION

This study proposes an artificial intelligence-based behavioral analysis system that utilizes computer vision and deep learning to augment the quality of doctor-patient therapy sessions. The system is effective in capturing and translating real-time visual information—e.g., facial expressions, body language, and engagement patterns—to offer therapists objective feedback about patient behavior. A mean emotion recognition F1-score of 90.5% and high agreement between AI prediction and therapist observation indicate that the tool is highly accurate and clinically relevant. From a safe and easy-to-use dashboard, therapists are provided with graphs of behavior trends, emotion summaries, and immediate alerts that help monitor patient progress and change treatment plans accordingly. Therapist feedback assures the system's usability, pragmatic utility, and feasibility of integration into actual therapy settings. In addition, the tool is effective to use in face-to-face and online therapy, meeting the increasing demand for smart support in virtual health settings. this paper provides a strong foundation for AI integration in therapy, providing a human-oriented, scalable, and ethically aware system for behavioral health care.

References:

1. Guo, R., Guo, H., Wang, L., Chen, M., Yang, D., & Li, B. (2024). *Development and application of emotion recognition technology: A systematic literature review*. BMC Psychology, 12, 95. IJSAT+3Psychology Journal+3EWA Direct+3BioMed Central
2. Sensors Review Team. (2024). *Systematic review of emotion detection with computer vision and deep learning*. Sensors, 24(11), 3484. MDPI
3. Frontiers in Psychology Editorial Board. (2024). *Computer vision and human behavior to recognize emotions*. Frontiers in Psychology. Frontiers
4. Jiang, M., et al. (2024). *A generic review of integrating artificial intelligence in*

cognitive behavioral therapy. arXiv preprint. TIME+15arXiv+15arXiv+15

5. Brain Informatics Team. (2023). *Machine learning for cognitive behavioral analysis: Datasets, methods, and challenges*. Brain Informatics, Article. arXiv+4SpringerOpen+4arXiv+4
6. Kim, S., Kim, H., Lee, J., Jeon, Y., & Lee, G. G. (2025, April). *Mirror: Multimodal cognitive reframing therapy for rolling with resistance*. arXiv preprint. arXiv
7. Ghafghazi, S., Carnett, A., Neely, L., Das, A., & Rad, P. (2021). *AI-augmented behavior analysis for children with developmental disabilities: Building toward precision treatment*. arXiv preprint. arXiv
8. Rudovic, O., Lee, J., Dai, M., Schuller, B., & Picard, R. (2018). *Personalized machine learning for robot perception of affect and engagement in autism therapy*. arXiv preprint. arXiv
9. Washington, P., Mutlu, C. O., Kline, A., Paskov, K., Stockham, N. T., Chrisman, B., … Wall, D. P. (2022). *Challenges and opportunities for machine learning classification of behavior and mental state from images*. arXiv preprint. arXiv
10. IJIRT Authors. (2025). *AI-powered emotional understanding for virtual therapy: A review*. IJIRT Journal. IJIRT
11. WJARR Authors. (2020). *Emotion recognition with AI: Techniques and applications*. WJARR. Wjarr
12. PeerJ Authors. (2023). *Deep learning–based dimensional emotion recognition for conversational iCBT*. PeerJ Computer Science. PeerJ
13. Ultralytics Authors. (2024). *The role of computer vision in mental health*. Ultralytics Blog. arXiv+9Ultralytics+9JETIR+9
14. SpringerOnline. (2025). *Revolutionizing mental healthcare with generative deep learning*. In *Advances in Health Informatics*. SpringerLink
15. Springer Psychology Journal. (2025). *The balance and integration of AI within cognitive behavioral therapy*. *Current Psychology*. SpringerLink+1arXiv+1
16. PsychologyJournal.net. (2025). *Enhancing therapy through technology: Role of AI in Indian mental health care*. *Psychology Journal*, 7(1C), 34-187. Psychology Journal
17. Wikipedia Contributors. (2025, July). *Artificial intelligence in mental health*. In *Wikipedia*. Wikipedia

18. Wikipedia Contributors. (2025, July). *Affective computing*. In *Wikipedia*. Wikipedia
19. Wikipedia Contributors. (2025, July). *Emotion recognition*. In *Wikipedia*. Wikipedia
20. Cao, L. (2020). *In-depth behavior understanding and use: The behavior informatics approach*. Information Science. Wikipedia
21. Pavel, M. (2015). *Behavioral informatics and computational modeling in proactive health management*. IEEE Trans. Biomed. Eng. Wikipedia
22. Wang, B. (2021). *Psychological AI: A critical analysis of capabilities, limitations, and ramifications*. ACE Proceedings. EWA Direct
23. Comer, J. S., Furr, J. M., Miguel, E. M., Cooper-Vince, C. E., Carpenter, A. L., Elkins, R. M., … Chase, R. (2017). Internet-delivered parent–child interaction therapy: A randomized trial. *Journal of Consulting and Clinical Psychology*, 85(9), 909–917. Wikipedia
24. Comer, J. S. (2021). *Rebooting mental health care delivery for the COVID-19 pandemic*: Guiding cautions. *Cognitive and Behavioral Practice*, 28(4), 743–748. Wikipedia
25. Comer, J. S., Furr, J. M., del Busto, C., et al. (2021). iCALM telehealth program for early child social anxiety. *Behavior Therapy*, 52(5), 1171–1187. Wikipedia
26. Time Magazine. (2019, November 20). *Artificial intelligence could help solve America's impending mental health crisis*. *Time*. Psychology Journal+7TIME+7ScienceDirect+7
27. New Yorker Staff. (2023, March 6). *Can A.I. treat mental illness? The New Yorker*. ScienceDirect+9The New Yorker+9PeerJ+9
28. Prevention Magazine. (2025, June 20). *Your next therapist could be an AI chatbot—but should they be? Prevention*. prevention.com
29. Lifewire Coverage. (2023). *Teaching AI how to feel could make for better computers*. *Lifewire*. Lifewire
30. Wall Street Journal. (2024). *Can mental-health chatbots help with anxiety and depression? WSJ*. wsj.com
31. Eghbalian, M., et al. (2025). *Applying computer vision to analyze self-injurious behaviors in children*. WACV 2025 Conference Proceedings. CVF Open Access
32. Guo, R., Guo, H., Wang, L., … Li, B. (2024). *Development and application of emotion recognition technology*

(duplicate year for second format). BMC Psychology, 12, 95. BioMed Central

33. MDPI Review Team. (2024). *Emotion detection with computer vision and deep learning trends and datasets*. Sensors. MDPI
34. MDPI Special Issue Editors. (2022). *Computer vision techniques applied to human behaviour analysis*. Sensors Special Issue. MDPI
35. Ultralytics Authors. (2024). *Computer vision in mental health care: Emotion recognition and behavior analysis*. Ultralytics Blog. Ultralytics
36. IJIRT Reviewers. (2025). *Emotional analysis for virtual therapy*. IJIRT. IJIRT
37. PeerJ Authors. (2023). *Dimensional facial emotion recognition for online therapy*. PeerJ Comp Sci. PeerJ
38. Springer Health Chapter Authors. (2025). *Generative deep learning in mental healthcare: Applications and ethics*. *Advances in Health Informatics*. SpringerLink
39. Psychology Journal Authors. (2025). *AI-based tools in therapy: Cultural and ethical issues in India*. *Psychology Journal*, 7(1C). Psychology Journal
40. Frontiers Psychology Editors. (2024). *Computer vision meets human behavior in emotion recognition*. *Frontiers in Psychology*. Frontiers
41. Washington, P., et al. (2022). *Machine learning for behavior and mental state classification from images*. arXiv. arXiv
42. Kim, S., et al. (2025). *Mirror: Multimodal dataset and AI therapy for resistance in clients*. arXiv. arXiv