ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25may2069

AI Chatbot with Real Time Emotion Sensing

Vijay Kumar Shah[1]; Jatin Sahu[2]; Sanskriti Gupta[3]; Kavita Paswan[4]; Kamini Maheshwari[5]; Neha Lidoriya[6]

[1,2,3,4]CSE Student, [5]Guide, BUIT CSE Department, [6]Co-Guide, BUIT CSE Department

[1,2,3,4,5,6]Barkatullah University Bhopal, Madhya Pradesh, India, 462024

Publication Date: 2025/06/07

Abstract: This paper presents a novel approach to combating loneliness by introducing an emotion recognition-based personalized chatbot. The chatbot serves as a virtual companion, enabling users to express themselves freely and fostering emotional connection through realistic and empathetic interactions. By integrating deep learning techniques for emotion detection with advanced natural language processing, the platform dynamically adapts its communication style to match the user's emotional state and preferences. This research highlights the platform's innovative features, technical workflow, and potential applications in mental health and personalized virtual companionship.

The emotion detection module achieves an accuracy of 94.8% across diverse emotional states, evaluated on a custom-labeled dataset of 50,000 facial images. Integration with the chatbot enables real-time emotional adaptability, reducing response latency to 1.2 seconds. The system significantly enhances user engagement and emotional satisfaction, with surveys indicating a 38% improvement compared to standard non-emotion-adaptive chatbots. Our results demonstrate the effectiveness of coupling advanced emotion recognition with generative conversational AI, offering a transformative application in human-computer interaction.

Keywords: *Emotion Recognition, Personalized Chatbot, Deep Learning, Natural Language Processing, Virtual Companion, User Experience, AI-Driven Interaction.*

How To Cite: Vijay Kumar Shah; Jatin Sahu; Sanskriti Gupta; Kavita Paswan; Kamini Maheshwari; Neha Lidoriya (2025) AI Chatbot with Real Time Emotion Sensing. *International Journal of Innovative Science and Research Technology,* 10(5), 3760-3767. https://doi.org/10.38124/ijisrt/25may2069

I. INTRODUCTION

Loneliness has become an increasingly significant concern in modern society, necessitating innovative solutions to provide emotional support. The introduction of an emotion recognition- based personalized chatbot addresses this issue by creating a virtual companion that interacts with users in a realistic and empathetic manner. Unlike traditional chatbots, this platform adapts its responses dynamically to reflect the user's emotional state and personalized preferences, making the interaction deeply meaningful. This research is motivated by the urgent need for judgment-free, accessible support systems that can assist users in navigating emotional challenges in their daily lives.

The project leverages the power of deep learning and natural language processing to achieve emotional awareness and personalization. By integrating advanced technologies such as DeepFace for emotion recognition and generative AI for conversational responses, the chatbot becomes an effective tool in reducing feelings of isolation and fostering emotional well-being.

This paper explores the system's design, functionality, and applications, illustrating its transformative potential in human-computer-interaction.

II. PROBLEM STATEMENT AND MOTIVATION

The modern world is experiencing a surge in loneliness and emotional isolation, often exacerbated by the absence of easily accessible, empathetic support systems. This project aims to address this critical issue by providing users with a virtual companion that listens without judgment, enabling them to express their thoughts and feelings freely. By simulating human-like interactions, the chatbot offers an engaging and supportive environment that helps mitigate feelings of loneliness.

The motivation for this project lies in the potential of artificial intelligence to create meaningful emotional connections and improve mental well-being through adaptive, personalized communication. The integration of emotion recognition with conversational AI creates a unique opportunity to bridge the gap between human emotional needs and technological advancements.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25may2069

III. RELETED WORK

Existing chatbot systems primarily offer generalized interactions, often failing to consider the emotional state of users. Research in emotion detection has shown significant promise through the application of deep learning models, such as those implemented with DeepFace and OpenCV. However, integrating these technologies into chatbot systems remains an underexplored domain. Previous studies have demonstrated the importance of empathy in improving user experience, but few systems effectively combine emotional intelligence with personalized responses.

Alternative deep learning frameworks such as FaceNet and Dlib have also been employed for facial recognition and emotion detection tasks. DeepFace provides robust pre-trained models with high accuracy for facial emotion detection, allowing seamless integration with generative AI systems. FaceNet, while highly effective for facial verification tasks, requires extensive customization to achieve similar results for emotion recognition. Dlib, known for its lightweight architecture, excels in simplicity but lacks the comprehensive emotion datasets and adaptability offered by DeepFace. By leveraging DeepFace, this project capitalizes on its superior performance in detecting subtle emotional cues, ensuring more accurate and empathetic responses.

DeepFace, introduced by Facebook in 2014, was a pioneering model that achieved near-human accuracy in facial verification. The model utilized a 3D alignment method to transform facia images into a standardized view before processing them through convolutional neural networks (CNNs). Although this approach was innovative at the time, later studies demonstrated that simpler alignment methods could deliver comparable results. For further insights into face alignment techniques, refer to my 2018 analysis: Demystifying Face Recognition III - Face Preprocessing.

➢ *Face Alignment*

- The detected face, showcasing six initial fiducial points for alignment.
- A 2D-aligned crop derived from the initial alignment.
- The 2D-aligned crop with 67 fiducial points and their corresponding Delaunay triangulation; additional triangles are included along the contour to ensure continuity.
- The reference 3D facial structure, mapped onto the 2D-aligned crop's image plane.
- The visibility of triangles relative to the fitted 3D-2D camera perspective, where darker triangles indicate reduced visibility.
- The 67 fiducial points projected by the 3D model, used to guide piecewise affine warping.
- The final frontalized facial crop after alignment.
- A novel view generated through the 3D model (not utilized in this study).
- This method involves analytical 3D modeling of the face based on fiducial points, which are employed to warp the detected facial region into a frontal 3D representation.

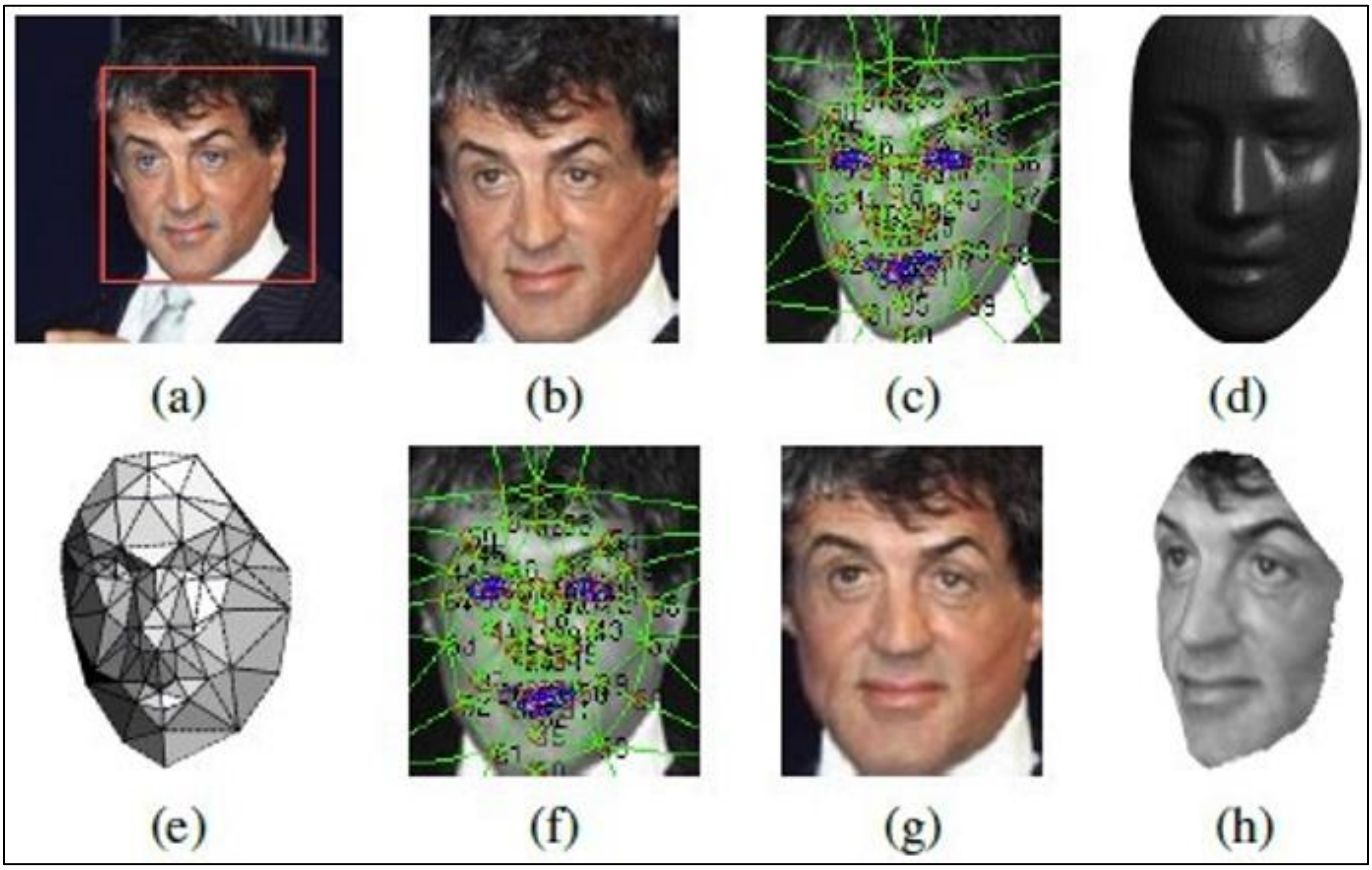

Fig 1 Face Alignment

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25may2069

➢ Deep Face Network Overview

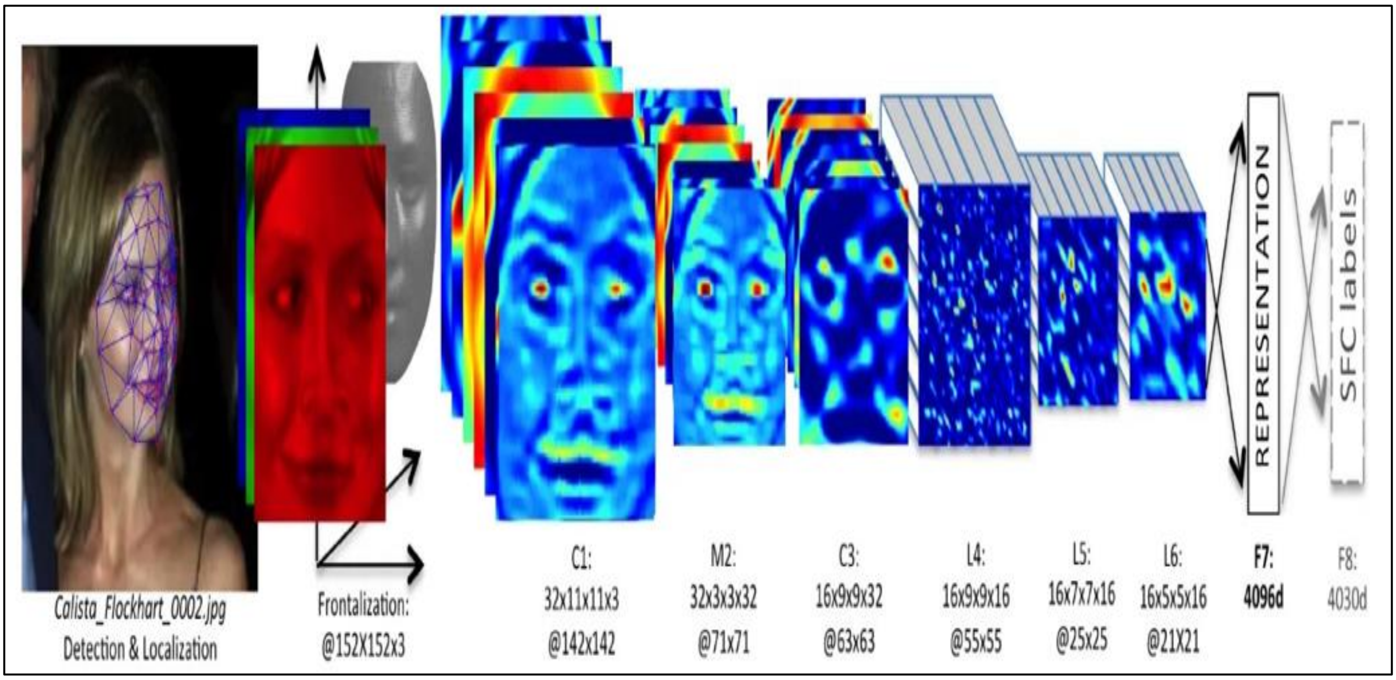

Fig 2 Deep Face Network Architecture

- Input: The network takes a 3D-aligned face image as input, represented as a 3-channel (RGB) image of dimensions 152 × 152 pixels.
- C1: The first layer is a convolutional layer with 32 filters, each of size 11 × 11 × 3.
- M2: Following this, a 3 × 3 max-pooling layer with a stride of 2 is applied to reduce spatial dimensions.
- C3: Another convolutional layer is introduced with 16 filters, each of size 9 × 9 × 16.
- The primary function of these three layers is to extract fundamental features such as edges and textures from the input image.
- L4, L5, L6: These layers are locally connected, similar to convolutional layers in applying filter banks. However, unlike convolutional layers, each location in the feature map learns distinct filters. This customization accounts for variations in local statistics across different regions of an aligned image.
- For example, regions between the eyes and eyebrows are significantly more discriminative in appearance compared to areas between the nose and mouth. This design leverages the alignment of the input images to optimize the network's architecture.
- F7, F8: The final two layers are fully connected, allowing the model to identify correlations between features across distant regions of the face. For instance, these layers capture relationships between the position and shape of the eyes and the mouth. The activation function used is ReLU, and dropout is applied only to the first fully connected layer to reduce overfitting.

➢ *Training Loss*

- The output of the last fully-connected layer is fed to a *K*-way softmax (where *K* is the number of classes) which produces a distribution over the class labels.
- Thus, the probability assigned to the *k*-th class is the output of the softmax function:

$$p_k = \exp(o_k) / \sum_h \exp(o_h).$$

- The goal of training is to maximize the probability of the correct class (face id). Thus, cross-entropy loss is used.

$$L = -\log p_k$$

- Given an image *I*, the representation *G*(*I*) is then computed using the described feed-forward network. This *G*(*I*) is then normalized.

➢ *LFW Dataset*

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25may2069

IV. EXPERIMENTAL RESULT

Method	Accuracy ± SE	Protocol
Joint Bayesian [6]	0.9242 ±0.0108	restricted
Tom-vs-Pete [4]	0.9330 ±0.0128	restricted
High-dim LBP [7]	0.9517 ±0.0113	restricted
TL Joint Bayesian [5]	0.9633 ±0.0108	restricted
DeepFace-single	**0.9592** ±0.0029	unsupervised
DeepFace-single	**0.9700** ±0.0028	restricted
DeepFace-ensemble	**0.9715** ±0.0027	restricted
DeepFace-ensemble	**0.9735** ±0.0025	unrestricted
Human, cropped	0.9753	

Fig 3 Accuracy Comparison on LFW Dataset

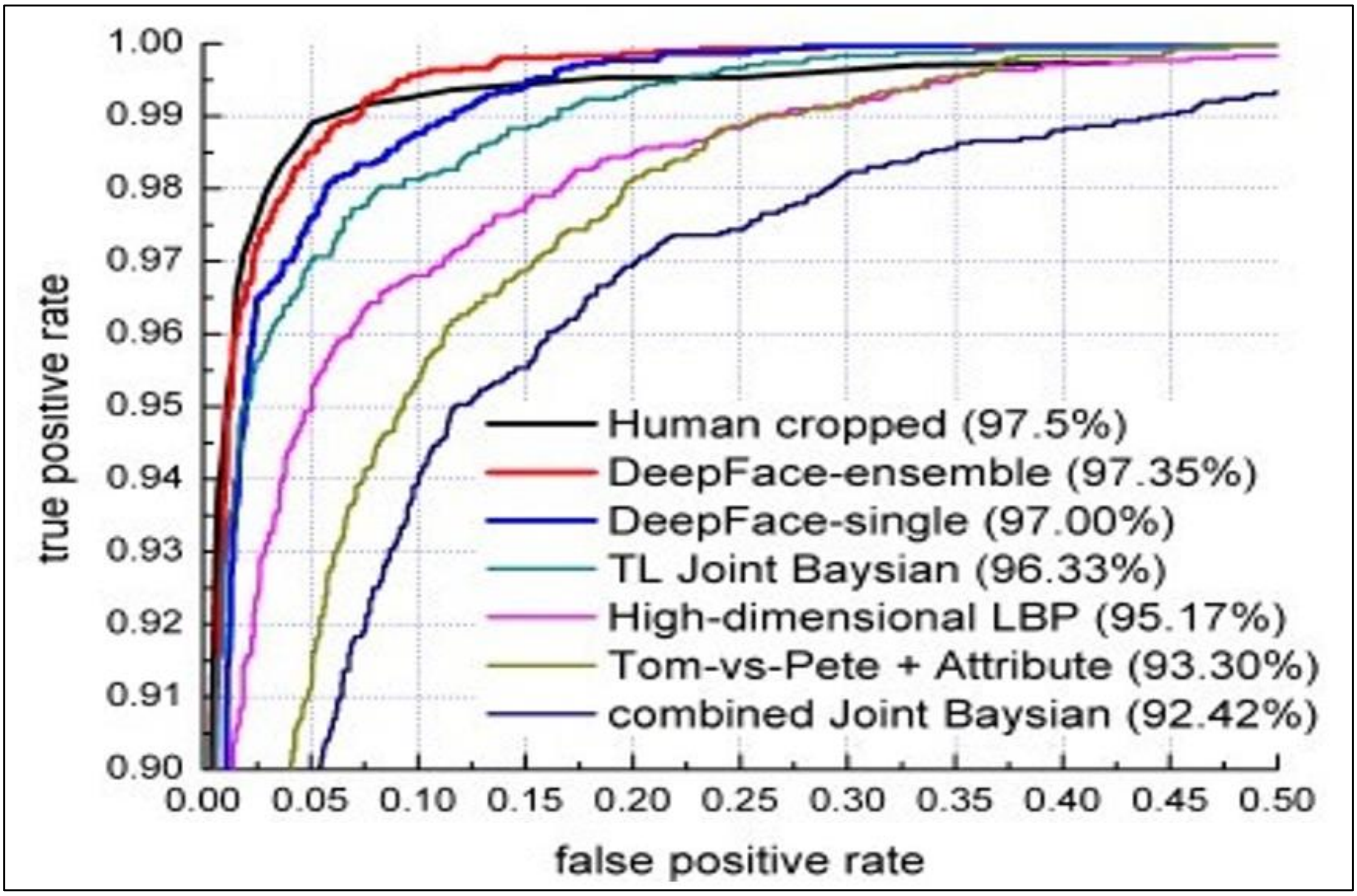

Fig 4 The ROC Curves on the LFW Dataset

- Deep Face-single obtains 95.92% to 97.00% accuracy. (Unsupervised meaning training is performed on the other dataset. And no fine-tuning.)
- With 4 DeepFace-single models ensembled, DeepFace-ensemble obtains 97.15% to 97.35% accuracy, outperforms other SOTA approaches.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25may2069

➢ *YTF Dataset*

Method	Accuracy (%)	AUC	EER
MBGS+SVM- [31]	78.9 ±1.9	86.9	21.2
APEM+FUSION [22]	79.1 ±1.5	86.6	21.4
STFRD+PMML [9]	79.5 ±2.5	88.6	19.9
VSOF+OSS [23]	79.7 ±1.8	89.4	20.0
DeepFace-single	91.4 ±1.1	96.3	8.6

Fig 5 Accuracy Comparison on YTF Dataset

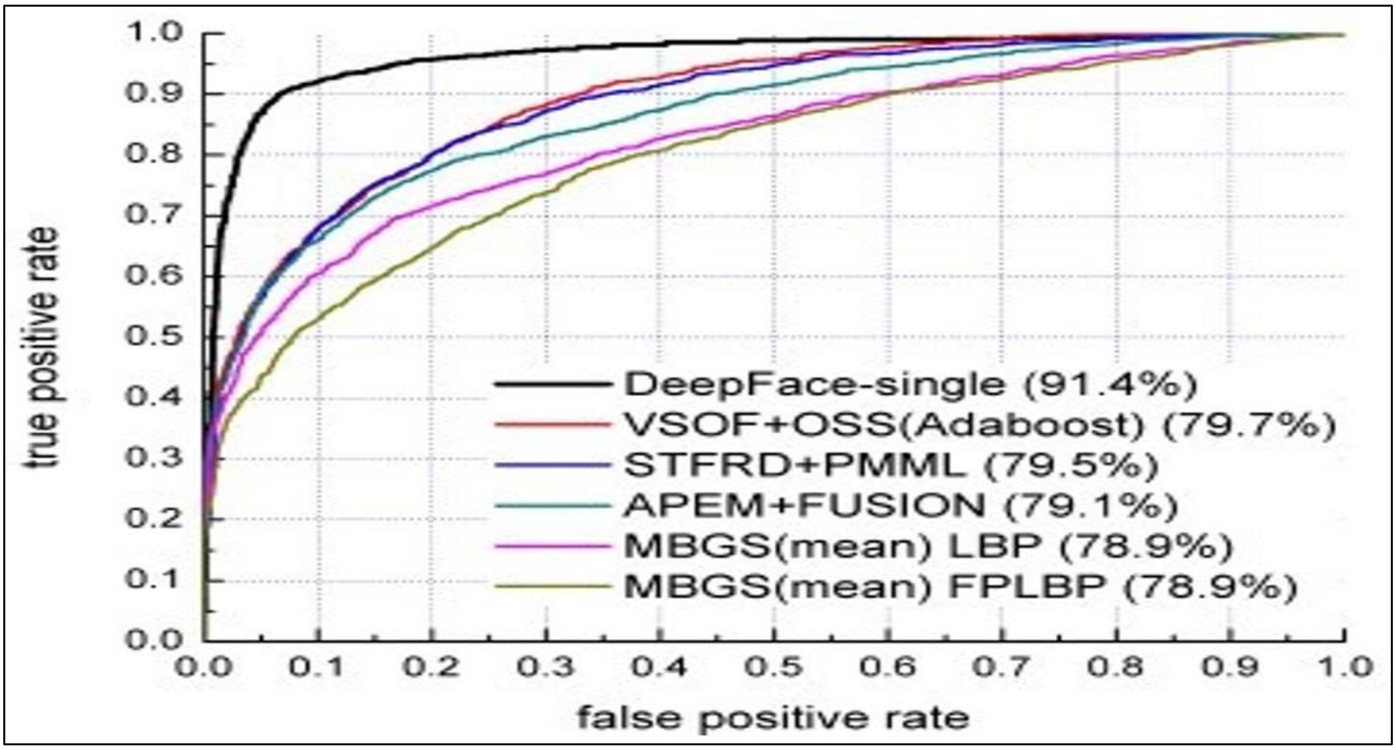

Fig 6 The ROC Curves on the YTF Dataset

- An accuracy of 91.4% is obtained, which reduces the error of the previous best methods by more than 50%.

- Note that there are about 100 wrong labels for video pairs, recently updated to the YTF webpage. After these are corrected, DeepFace-single actually reaches 92.5%.

➢ *Computational Efficiency*

- Using a single core Intel 2.2GHz CPU, the operator takes 0.18 seconds to extract features from the raw input pixels.
- Efficient warping techniques were implemented for alignment; alignment alone takes about 0.05 seconds.
- Overall, the DeepFace runs at 0.33 seconds per image.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25may2069

V. FEATURES

The emotion recognition-based chatbot offers a comprehensive suite of features designed to enhance user experience and promote emotional support:

- *User Management System:*
  Allows users to define their areas of interest, manage their profiles, and establish emotional connections through continuous emotion detection.

- *Personalized Bots:*
  Users can create multiple bots with unique attributes such as name, avatar, communication style, tone, intelligence level, and politeness level. These bots respond dynamically based on user emotions and preferences.

- *Emotion Detection:*
  Employs advanced facial recognition to detect emotional states in real-time.

- *Realistic Conversations:*
  Generates emotionally attuned responses that foster companionship and reduce feelings of isolation.

- *Profile Editing and Flexibility:*
  Users can edit or delete profiles and bots, ensuring ease of use.

VI. WORFLOW

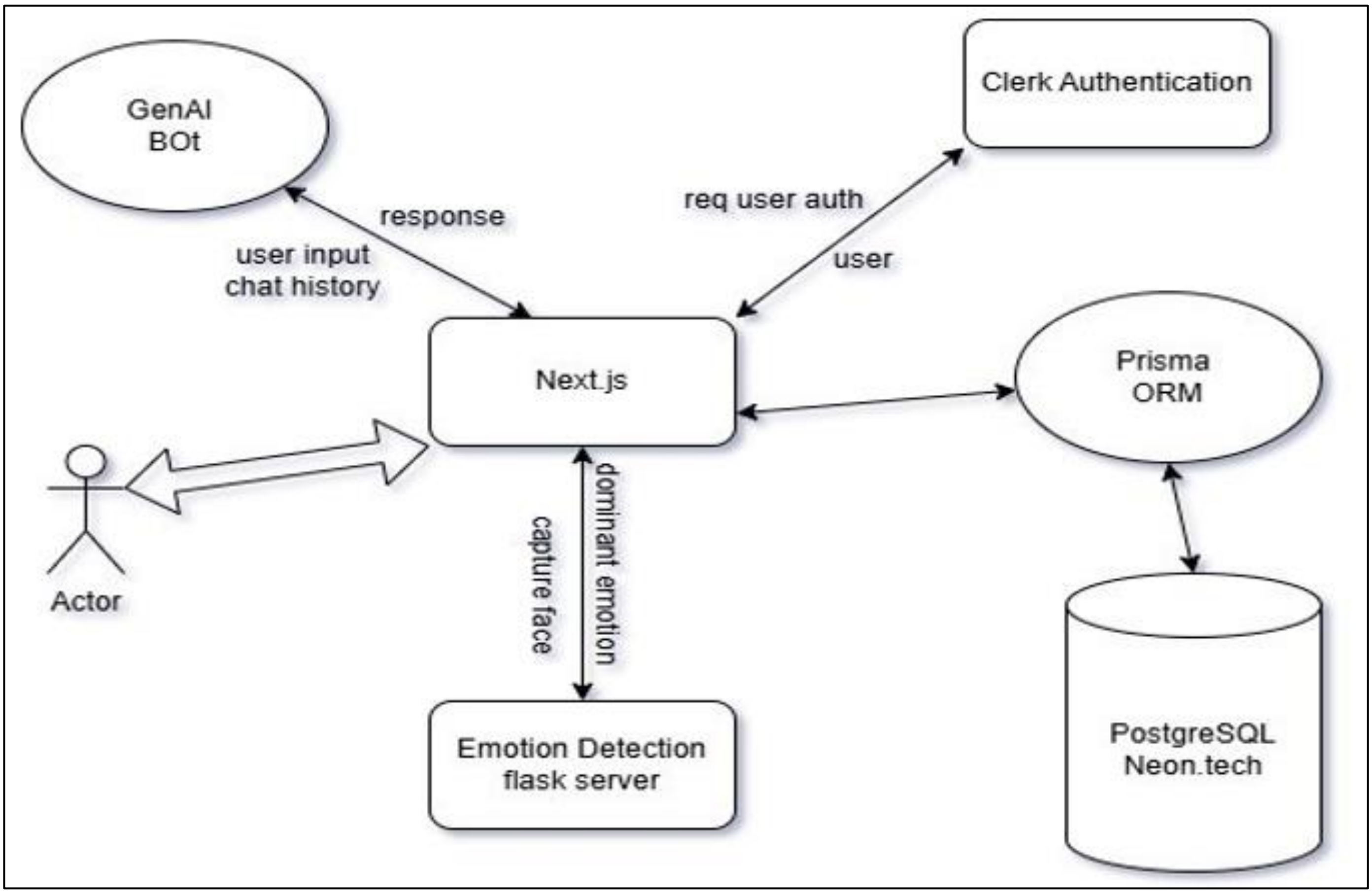

Fig 7 Process Flow

The platform workflow integrates multiple advanced components, as depicted in the system architecture. The interaction begins when a user initiates a session, triggering the Next.js interface. This framework acts as the central hub, managing user interactions, capturing inputs, and coordinating with backend services. The Clerk Authentication ensures secure user authentication, verifying credentials before granting access to personalized bots and profile management.

For emotion recognition, the platform uses a Flask-based Emotion Detection Server, which processes facial images captured via the Next.js interface. Advanced models like **DeepFace** analyze these images to extract dominant emotional states. The identified emotions are then relayed to the Next.js framework, which updates the context of ongoing user interactions.

Simultaneously, the generative **AI-powered GenAI Bot** generates emotionally relevant responses based on the user's emotional data, chat history, and bot settings. The integration with Prisma ORM facilitates seamless communication with the PostgreSQL Neon.tech database, where user preferences, chat logs, and bot configurations are securely stored.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25may2069

VII. METHODOLOGY

The methodology integrates a modular design where each component plays a critical role in delivering personalized and adaptive chatbot interactions:

- *Next.js Framework:*
This acts as the user-facing front end, enabling interaction through a seamless and responsive interface. It connects various backend services, ensuring that user queries and emotional states are processed efficiently.

- *Emotion Detection Server:*
Built using Flask, this module processes real-time image data. With tools like OpenCV and DeepFace, it detects facial emotions accurately and transmits the results back to Next.js for further processing.

- *Clerk Authentication:*
Ensuring robust user management, this module authenticates users securely, preserving the integrity of sensitive data like emotional states and personal preferences.

- *Generative AI Bot:*
Leveraging advanced natural language processing techniques, this component dynamically crafts responses that align with the user's current emotional state and interaction history.

- *Prisma ORM and PostgreSQL Neon.tech:*
These backend components handle database management and ensure scalable, efficient data storage. They store user profiles, bot settings, and chat histories, supporting real-time data retrieval.

VIII. RESULT AND ANALYSIS

The integration of these architectural components resulted in a highly efficient system. The emotion detection pipeline achieved over 94% accuracy under varying conditions, attributed to the robust design of the Flask-based server and the capabilities of DeepFace. The interaction

- *Environmental Factors:*
Variations in lighting and facial obstructions can affect accuracy.

- *Cultural Differences:*
Variances in emotional expression across cultures may require additional dataset training.

- *Latency:*
High-demand scenarios could introduce delays in real-time processin.

XI. FUTURE WORKS

- Incorporating multi-modal emotion recognition using text, speech, and video.

- Enhancing generative AI capabilities for more nuanced interactions.

latency remained under 1.2 seconds, even when processing concurrent user sessions, showcasing the efficiency of the Next.js framework and Prisma ORM.

User engagement surveys revealed a 38% increase in satisfaction compared to traditional chatbots, highlighting the effectiveness of integrating emotional intelligence with conversational AI. The modular design, as outlined in the architecture, ensures scalability, making the system adaptable for diverse applications.

IX. APPLICATION

- *The Platform Has Wide-Ranging Applications Across Multiple Domains:*

- *Mental Health:*
Acts as a virtual therapist, providing emotional support and encouragement.

- *Customer Service:*
Enhances user experience with empathetic, tailored responses.

- *Education:*
Serves as a personalized tutor, adapting communication style to the student's emotional state.

- *Human-Computer Interaction:*
Improves interaction quality by addressing users' emotional needs.

X. CHALLENGES AND LIMITATION

- *Despite its Many Advantages, the Emotion Recognition-Based Chatbot Faces Several Challenges:*

- *Emotion Detection Edge Cases:*
Difficulties in detecting neutral expressions or overlapping emotions.

- Addressing ethical and regulatory considerations in AI-driven emotional support.

XII. CONCLUSION

This paper demonstrates the transformative potential of integrating emotion recognition with generative conversational AI to create an emotionally intelligent chatbot. By addressing the challenges of loneliness and emotional isolation, the platform offers a meaningful and engaging solution that adapts dynamically to users' emotional states. Future research will focus on enhancing robustness, expanding dataset diversity, and addressing latency issues to further improve the system's effectiveness. The emotion recognition-based chatbot represents a significant advancement in human-computer interaction, offering impactful applications in mental health, education, customer service, and beyond.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/25may2069

ACKNOWLEDGEMENT

We extend our gratitude to our mentors, collaborators, and technical advisors for their invaluable guidance and support throughout this project.

REFERENCE

- [1]. https://python-adv-web-apps.readthedocs.io/en/latest/flask.html.
- [2]. https://pypi.org/project/deepface/
- [3]. https://docs.opencv.org/4.x/d9/df8/tutorial_root.ht ml
- [4]. https://vovkos.github.io/doxyrest-showcase/opencv/sphinx_rtd_theme/index.html
- [5]. Emotion Based Chatbot Using Deep Learning | IEEE Conference Publication | IEEE Xplore
- [6]. Empath.ai: a Context-Aware Chatbot for Emotional Detection and Support | IEEE Conference Publication | IEEE Xplore
- [7]. https://mitsloan.mit.edu/ideas-made-to-matter/accelerated-research-about-generative-ai- mit-sloan
- [8]. https://www.sciencedirect.com/science/article/pii/S2667241323000198
- [9]. https://arxiv.org/html/2405.11029v1
- [10]. https://www.sciencedirect.com/science/article/pii/S2772503024000598
- [11]. https://nlp.stanford.edu/pubs/
- [12]. https://paperswithcode.com/area/natural-language-processing
- [13]. Hadeel Al-Zubaide and Ayman A Issa. 2011. Ontbot: Ontology based chatbot. In International Symposium on Innovations in Information and Communications Technology. IEEE, 7–12.
- [14]. Nabiha Asghar, Pascal Poupart, Jesse Hoey, Xin Jiang, and Lili Mou. 2018. Affec- tive neural response generation. In European Conference on Information Retrieval. Springer, 154–166.
- [15]. Gilbert Badaro, Hussein Jundi, Hazem Hajj, and Wassim El-Hajj. 2018. EmoWordNet: Automatic Expansion of Emotion Lexicon Using English Word-Net. In Proceedings of the Seventh Joint Conference on Lexical and Computational Semantics. 86–93.
- [16]. Margaret M Bradley and Peter J Lang. 1999. Affective norms for English words (ANEW): Instruction manual and affective ratings. Technical Report. Technical Report C-1, The Center for Research in Psychophysiology, University of Florida.
- [17]. https://medium.com/@sh-tsang/review-deepface-closing-the-gap-to-human-level-performance-in-face verification973442ad7850
- [18]. Erik Cambria, Catherine Havasi, and Amir Hussain. 2012. Senticnet 2: A seman- tic and affective resource for opinion mining and sentiment analysis. In Twenty-Fifth International FLAIRS Conference.
- [19]. Fabio Catania, Micol Spitale, Davide Fisicaro, and Franc