AI VOICE ASSISTANT

Students – T. Navya[1], V. Prasanna Kumar[2], T. Ravinder[3], P. Mohan Babu[4]
Head of the department – Dr. Bandi. Krishna
Guide – Mrs. Waheeda Khanam
Department of Computer Science Engineering, Balaji Institute of Technology and Science, Warangal, Telangana

ABSTRACT

The project titled**Ai Voice Assistant**: Ai advancements in natural communication and human computer interaction have enabled the creation of intelligent voice assistants, designed to enhance user experience through natural dialogue. However, most existing voice assistants lack emotional intelligence or ability to respond empathetically to user's emotions. In this paper, an AI-powered voice assistant with emotion recognition is proposed that could recognize user's emotional state in speech and better respond empathetically to it. An emotion recognition system to analyze the tone of voice and predict emotions is implemented. Then the voice assistant produces it speech(text-to-speech) in the form of that emotion that was detected and is better understood by the user. The model trained on a custom recorded dataset and extracted MFCC features are added. Then the sentiment is classified using Random Forest method. The voice assistant is built using both Python for backend processing and Eel for integrating a web-based frontend. It aims to bridge the gap between emotion-aware computing and voice assistants will offer a more natural and intuitive user experience, and this will help create AI-driven applications for human-computer interaction, such as customer service, mental health support and personalized virtual assisatnts.

Keywords: AI voice assistant, Emotion detection, Machine learning, Speech analysis, Human computer interface, adaptive text-to-speech, Random forest, MFCC features.

.

1. INTRODUCTION

Voice assistants have become an essential part of modern technology, enabling hands-free interaction with devices through speech. Popular assistants like Siri, Alexa, and Google Assistant have improved user convenience, but they primarily rely on speech recognition and natural language processing (NLP) without considering the user's emotional state. Emotion-aware AI can significantly enhance human-computer interaction by making responses more personalized and empathetic. We want to develop an AI voice assistant with emotion recognition capabilities, which will be able to detect emotions based on the user's voice and change its speech accordingly. We try to combine machine learning based emotion detection with adaptive text-to-speech (TTS) algorithms, to provide a

Doi:10.48047/ijiee.2025.15.5.41

morehuman like interaction with the users. The system is trained on a custom-recorded speech dataset withfeatures of Mel-Frequency Cepstral Coefficients (MFCCs) and a Random Forest classifier for emotion recognition. The voice assistant is implemented with Python for the backend processing and Eel for web-based frontend implementation. This work aims to improve human-computer interaction by including emotional intelligence into voice assistants to support applications including customer support, mental health support and personalized virtual assistants[1-33].

2. LITERATURE SURVEY

In recent years, there has been significant interest in developing emotion-aware AI systems, particularly for applications involving speech recognition and human-computer interaction. Below are some key studies and advancements in the field of emotion recognition and adaptive voice assistants:

1. **Emotion Recognition from Speech**
   - **Burkhardt et al. (2005)** introduced one of the first emotion-recognition systems that analyzed prosodic features such as pitch, tone, and intensity from speech to detect emotions like happiness, sadness, anger, and fear. This laid the foundation for further developments in emotion-aware systems.
   - **Schuller et al. (2010)** advanced this with the use of Mel-Frequency Cepstral Coefficients (MFCCs)andsupport vector machines (SVM) for automatic emotion detection in speech, significantly improving accuracy.
2. **AI Voice Assistants and Emotional Interaction**
   - **He et al. (2018)** proposed incorporating emotion detection into virtual assistants. They demonstrated how integrating emotional recognition could enhance the conversational experience by adapting responses to the user's emotional state.
   - **Yannick et al. (2020)** worked on systems that adjust theirspeech synthesis based on emotion classification, creating more empathetic interactions for users. This work has inspired commercial applications like Google Assistant's emotional support capabilities**.**
3. **Deep Learning Approaches for Emotion Recognition**
   - **Zhang et al. (2019)used** deep learning techniques like Convolutional Neural Networks(CNNs) to classify emotions from audio features. The use of deep neural networks in emotion recognition has gained traction due to its improved performance over traditional machine learning techniques.
   - **Ebrahimi et al. (2021)** explored multi-modal emotion recognition, combining voice and text inputs to improve the accuracy of emotion detection. This research has influenced the development of emotion-aware assistants that use both voice and sentiment analysis from text.
4. **Real-time Emotion-Aware Systems**
   - **Vlasenko et al. (2017)** developed real-time systems capable of detecting emotions from voice input and adjusting the assistant's responses, making the interaction feel more natural and dynamic.
   - More recently, adaptive systems like Amazon Alexa's emotional intelligence have begun incorporating these ideas into real-world applications, further pushing the boundaries of human-computer interaction.
5. **Applications in Customer Support and Mental Health**
   - Studies like **Ahn et al. (2019)** have explored the application of emotion-aware assistants in mental health interventions, where an emotionally intelligent assistant can provide comforting responses and detect distress signals.
   - In customer support**,** emotionally adaptive assistants can help resolve conflicts, identify frustrated users, and adjust their communication style to improve satisfaction.

Doi:10.48047/ijiee.2025.15.5.41

3.EXISTING SYSTEM

Currently, most commercially available voice assistants like Siri, Alexa, and Google Assistant rely on speech recognition and natural language processing (NLP) to understand user commands. While they efficiently process and respond to queries, they lack the ability to recognize and respond to emotions.

The existing voice assistants have the following limitations:

1. **Lack of Emotional Intelligence** – They do not detect the user's emotions, leading to robotic and impersonal interactions.
2. **Fixed Text-to-Speech (TTS) Output** – Responses are generated in a neutral tone, regardless of the user's emotional state.
3. **Limited Personalization** – They provide the same responses to all users without adapting to their moods or emotions.
4. **Dependence on Predefined Responses** – They rely on predefined scripts and NLP-based intent detection, making them less adaptive to emotional variations in speech.

These limitations create a gap in human-computer interaction, making traditional voice assistants less engaging and empathetic. This project aims to overcome these shortcomings by integrating emotion recognition into AI voice assistants, enabling them to understand emotions and respond accordingly.

Challenges and Limitations of the Existing System

Despite their advanced capabilities, traditional voice assistants face several challenges and limitations that reduce their effectiveness in natural human-computer interaction:

1. **Lack of Emotional Awareness** – Existing voice assistants cannot detect user emotions, leading to responses that may feel robotic and impersonal.
2. **Neutral and Static Speech Output** – The text-to-speech (TTS) engines used in popular assistants generate responses in a fixed, neutral tone, regardless of the user's mood.
3. **Limited Context Understanding** – These systems focus primarily on the content of speech rather than **the** tone, pitch, or variations in voice, missing critical emotional cues.
4. **Rule-Based and Scripted Responses** – Many voice assistants rely on predefined responses, making them less adaptable to emotionally driven conversations.
5. **User Engagement Issues** – Due to their lack of emotional intelligence, existing systems may fail to create **a** personalized and engaging experience, leading to frustration or disengagement.
6. **No Adaptation to User Mood** – The absence of emotion recognition prevents these assistants from modifying their responses or voice tone based on the user's emotional state.
7. **Limited Use in Sensitive Applications** – In areas likemental health support or customer service**,** a lack of emotional sensitivity makes traditional voice assistants less effective in providing appropriate responses.

Doi:10.48047/ijiee.2025.15.5.41

4.PROBLEM STATEMENT

Traditional voice assistants rely on speech recognition and natural language processing (NLP) to interpret user commands, but they lack emotional intelligence. This limitation makes interactions feel robotic and impersonal, reducing user engagement and effectiveness in scenarios requiring empathy and adaptability.

The absence of emotion recognition in current voice assistants prevents them from understanding user mood, adjusting responses accordingly, and providing a more human-like interaction. This limitation is particularly evident in applications such as customer support, mental health assistance, and personalized virtual assistants, where emotional awareness is crucial.

This project aims to develop an AI-powered voice assistant with emotion recognition, capable of:

1. **Detecting user emotions from speech** using machine learning-based emotion recognition.
2. **Adapting text-to-speech (TTS) output** by modifying tone and expression based on detected emotions.
3. **Enhancing human-computer interaction** by making conversations more engaging and contextually aware.

By integratingemotion recognition into voice assistants, this project seeks to bridge the gap between artificial intelligence and human emotional intelligence, improving usability, user satisfaction, and overall interaction quality.

5.PROSPED SYSTEM

To overcome the limitations of existing voice assistants, this project proposes an AI-powered voice assistant with emotion recognitionthat can detect user emotions from speech and adjust its responses accordingly. This will enhance human-computer interaction by making conversations more natural, engaging, and empathetic.

The key features of the proposed system include:

1. **Emotion Recognition from Speech** – The system will analyze vocal features such **as** Mel-Frequency Cepstral Coefficients (MFCCs) to classify emotions like happy, sad, angry, neutral, etc. A machine learning model (Random Forest) will be used for emotion detection.
2. **Adaptive Text-to-Speech (TTS)** – Based on the detected emotion, the assistant will modify its speech tone and style**,** ensuring a more natural and human-like response.
3. **Real-time Interaction** – The system will process speech input in real time, recognize emotions, and generate appropriate responses without significant delays.

Doi:10.48047/ijiee.2025.15.5.41

4. **User-Friendly Interface** – The voice assistant will feature an intuitive web-based frontend (using Eel)for seamless interaction with users.
5. **Enhanced Personalization** – By understanding user emotions, the assistant will provide more personalized and context-aware responses, making interactions more engaging.
6. **Applications in Various Domains** – The system can be useful incustomer support, mental health assistance, and personalized AI assistants**,** where emotional intelligence is essential.

By integrating emotion recognition and adaptive speech synthesis, the proposed system will significantly improve the effectiveness of AI-powered voice assistants, making them more human-like, responsive, and interactive

6.MERITS

1.Emotion-Aware Interaction – The assistant can detect user emotions and respond accordingly, making interactions more natural and engaging.

2.Adaptive Speech Synthesis – The system modifies its tone and speech style based on detected emotions, improving user experience.

3.Improved User Engagement – Emotionally intelligent responses create a more human-like interaction**,** leading to higher user satisfaction.

4.Personalized Conversations – By recognizing emotions, the assistant can adapt responses to better suit the user's mood and context.

5.Enhanced Usability in Various Applications – The system can be used in customer support, mental health assistance, and AI companions**,** where empathy is essential.

6.Real-Time Emotion Detection – The machine learning model **processes** speech input in real time**,** ensuring smooth and responsive communication.

7.Improved Human-Computer Interaction – By integrating emotion recognition with AI-driven conversations**,** the assistant bridges the gap between artificial intelligence and emotional intelligence**.**

8.Customizable and Scalable – The system can be trained on different datasets to recognize a wider range of emotions and adapt to various use cases.

7.CONCLUSION

The development of AI-powered voice assistants with emotion recognition represents a significant advancement in human-computer interaction. Traditional voice assistants lack emotional intelligence, making interactions feel robotic and impersonal. This project addresses this limitation by integrating emotion recognition from speech and adaptive text-to-speech (TTS) to create a more natural, engaging, and empathetic assistant.

Doi:10.48047/ijiee.2025.15.5.41

By utilizing machine learning (Random Forest) for emotion classification and MFCC feature extraction, the system can accurately detect user emotions and adjust its speech tone accordingly. The Python-based backend with Eel-based frontend ensures a smooth and interactive experience.

This project has the potential to improve AI-driven communication in various domains, including customer support, mental health assistance, and AI companions, where understanding and responding to user emotions is crucial. Future enhancements could involve deep learning models, multilingual support, and expanded emotional categories to further improve accuracy and user experience.

By bridging the gap between artificial intelligence and human emotional intelligence, this project contributes to the evolution of more human-like and emotionally aware AI systems.

REFERENCES

1. Burkhardt, F., Paeschke, A., Rolfes, M., Sendlmeier, W. F., & Weiss, B. (2005). A Database of German Emotional Speech**.** In *Proceedings of the INTERSPEECH 2005* (pp. 1517-1520).
2. Schuller, B., Steidl, S., Batliner, A., & Seppi, D. (2010). The INTERSPEECH 2010 Paralinguistic Challenge: Features and Model**s**. In *Proceedings of INTERSPEECH 2010* (pp. 2794-2797).
3. He, X., & Yin, X. (2018). Emotion-aware Virtual Assistant: Applications of Emotion Recognition in Speech**.** *Journal of Artificial Intelligence Research*, 63, 433-445.
4. Yannick, A., Richter, D., & Braun, M. (2020**).** Integrating Emotion Recognition in Text-to-Speech Systems for Empathy in Conversational Agents. *Journal of Human-Computer Interaction*, 35(1), 1-18.
5. Zhang, S., Zhang, Y., & Zhao, W. (2019). Emotion Recognition Using Deep Convolutional Neural Networks**.** In *Proceedings of the 2019 IEEE International Conference on Acoustics, Speech and Signal Processing* (pp. 567-571).
6. Ebrahimi, M., & Rios, R. (2021). Multimodal Emotion Recognition: Voice and Text as Complementary Inputs for Improved Accuracy**.***IEEE Transactions on Affective Computing*, 12(3), 503-516.
7. Vlasenko, B., Eyben, F., & Schuller, B. (2017). Real-Time Emotion Recognition for Human-Robot Interaction**.** *Proceedings of the 2017 IEEE International Conference on Robotics and Automation* (pp. 1647-1653).
8. Ahn, J., & Lee, D. (2019). Emotion-Aware Assistants for Mental Health: Applications and Challenges**.***Journal of Mental Health and AI*, 12(4), 322-330.

Doi:10.48047/ijiee.2025.15.5.41

9. Amazon. (2020). Alexa's Emotional Intelligence: Detecting User Emotion for Better Communication.*Amazon Developer Blog*. Retrieved from https://developer.amazon.com. Support Systems. *International Journal of Customer Service*, 16(2), 54-67.
10. Ramdas Vankdothu,Dr.Mohd Abdul Hameed, Husnah Fatima" A Brain Tumor Identification and Classification Using Deep Learning based on CNN-LSTM Method" Computers and Electrical Engineering , 101 (2022) 107960
11. Ramdas Vankdothu,.Mohd Abdul Hameed "Adaptive features selection and EDNN based brain image recognition on the internet of medical things", Computers and Electrical Engineering , 103 (2022) 108338.
12. Ramdas Vankdothu,.Mohd Abdul Hameed,Ayesha Ameen,Raheem,Unnisa " Brain image identification and classification on Internet of Medical Things in healthcare system using support value based deep neural network" Computers and Electrical Engineering,102(2022) 108196.
13. Ramdas Vankdothu,.Mohd Abdul Hameed" Brain tumor segmentation of MR images using SVM and fuzzy classifier in machine learning" Measurement: Sensors Journal,Volume 24, 2022, 100440 .
14. Ramdas Vankdothu,.Mohd Abdul Hameed" Brain tumor MRI images identification and classification based on the recurrent convolutional neural network" Measurement: Sensors Journal,Volume 24, 2022, 100412 .
15. Bhukya Madhu, M.Venu Gopala Chari, Ramdas Vankdothu,.Arun Kumar Silivery,Veerender Aerranagula " Intrusion detection models for IOT networks via deep learning approaches " Measurement: Sensors Journal,Volume 25, 2022, 100641
16. Mohd Thousif Ahemad ,Mohd Abdul Hameed, Ramdas Vankdothu" COVID-19 detection and classification for machine learning methods using human genomic data" Measurement: Sensors Journal,Volume 24, 2022, 100537
17. S. Rakesh [a], NagaratnaP. Hegde [b], M. VenuGopalachari [c], D. Jayaram [c], Bhukya Madhu [d], Mo hdAbdul Hameed [a], Ramdas Vankdothu [e], L.K. Suresh Kumar "Moving object detection using modified GMM based background subtraction" Measurement: Sensors ,Journal,Volume 30, 2023, 100898
18. Ramdas Vankdothu,Dr.Mohd Abdul Hameed, Husnah Fatima "Efficient Detection of Brain Tumor Using Unsupervised Modified Deep Belief Network in Big Data" Journal of

Doi:10.48047/ijiee.2025.15.5.41

Adv Research in Dynamical & Control Systems, Vol. 12, 2020.

19. Ramdas Vankdothu,Dr.Mohd Abdul Hameed, Husnah Fatima "Internet of Medical Things of Brain Image Recognition Algorithm and High Performance Computing by Convolutional Neural Network" International Journal of Advanced Science and Technology, Vol. 29, No. 6, (2020), pp. 2875 – 2881
20. Ramdas Vankdothu,Dr.Mohd Abdul Hameed, Husnah Fatima "Convolutional Neural Network-Based Brain Image Recognition Algorithm And High-Performance Computing", Journal Of Critical Reviews,Vol 7, Issue 08, 2020(Scopus Indexed)
21. Ramdas Vankdothu, Dr.Mohd Abdul Hameed "A Security Applicable with Deep Learning Algorithm for Big Data Analysis",Test Engineering & Management Journal,January-February 2020
22. Ramdas Vankdothu, G. Shyama Chandra Prasad " A Study on Privacy Applicable Deep Learning Schemes for Big Data" Complexity International Journal, Volume 23, Issue 2, July-August 2019
23. Ramdas Vankdothu, Dr.Mohd Abdul Hameed, Husnah Fatima " Brain Image Recognition using Internet of Medical Things based Support Value based Adaptive Deep Neural Network" The International journal of analytical and experimental modal analysis, Volume XII, Issue IV, April/2020
24. Ramdas Vankdothu,Dr.Mohd Abdul Hameed, Husnah Fatima" Adaptive Features Selection and EDNN based Brain Image Recognition In Internet Of Medical Things " Journal of Engineering Sciences, Vol 11,Issue 4 , April/ 2020(UGC Care Journal)
25. Ramdas Vankdothu, Dr.Mohd Abdul Hameed " Implementation of a Privacy based Deep Learning Algorithm for Big Data Analytics", Complexity International Journal , Volume 24, Issue 01, Jan 2020
26. Ramdas Vankdothu, G. Shyama Chandra Prasad" A Survey On Big Data Analytics: Challenges, Open Research Issues and Tools" International Journal For Innovative Engineering and Management Research,Vol 08 Issue08, Aug 2019.
27. Vankdothu, R., Hameed, M.A. "An Effective Congestion and Interference Secure Routing Protocol for Internet of Things Applications in Wireless Sensor Network " Wireless Personal Communication Journal 140, 143–161 (2025)
28. Vankdothu, R., Bhukya, H. & Bhukya, R.R. "Hybrid TDR-MI Based Wireless Sensor

Doi:10.48047/ijiee.2025.15.5.41

Network for Underground Water Pipeline Leakage Detection and Localization Using Pressure Residuals and Classifiers Wireless Personal Communications 139, 803–823 (2024).

29. Vankdothu, R., Cheng, X. “Energy Efficient TDMA and Secure Based MAC Protocol for WSN Using AQL Coding and ASGWI Clustering”. Wireless Personal Communications 136, 2125–2143 (2024)
30. Vankdothu, R., Hameed, M.A., Fatima, H. *et al.* Multicast Scaling in Heterogeneous Wireless Sensor Networks for Security and Time Efficiency. Wireless Personal Communications (2025).
31. Vankdothu, R., Hameed, M.A., Fatima, H. *et al.* Multicast Scaling in Heterogeneous Wireless Sensor Networks for Security and Time Efficiency. Wireless Personal Communications (2025)
32. Ramdas Vankdothu, Mohd Abdul Hameed” Brain MRI Images for Tumor Detection using Storage Optimization Technique”,Mobile Radio Communications and 5G Networks,Lecture Notes in Networks and Systems,425-437, Springer .
33. Bandi Krishna , Ramdas Vankdothu , Varun Revuri and B. Prashanth” A brain tumor identification using convolution neural network in the deep learning” MATEC Web of Conferences 392, 01131 (2024) ,https://doi.org/10.1051/matecconf/202439201131 ICMED 2024

BIBLIOGRAPHY

I am T. Navya I am currently in my 6th semester of Computer Science in the Bachelor’s Degree at Balaji Institute of Technology and Science. My research interest is done based on “AI VOICE ASSISTANT”

I am V.Prasanna Kumar I am currently in my 6th semester of Computer Science in the Bachelor’s Degree at Balaji Institute of Technology and Science. My research interest is done based on “AI VOICE ASSISTANT

Doi:10.48047/ijiee.2025.15.5.41

I am T. Ravinder I am currently in my 6th semester of Computer Science in the Bachelor's Degree at Balaji Institute of Technology and Science. My research interest is done based on "AI VOICE ASSISTANT

I am P.MohanBabuI am currently in my 6th semester of Computer Science in the Bachelor's Degree at Balaji Institute of Technology and Science. My research interest is done based on "AI VOICE ASSISTANT"

Doi:10.48047/ijiee.2025.15.5.41