IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁒ Peer-reviewed & Refereed journal ⁒ Vol. 13, Issue 2, February 2024

DOI: 10.17148/IJARCCE.2024.13248

Enhanced AI Bot with Facial Emotion Detection

Dr. O. Aruna[1], Medikondu Mukesh[2], Neelam Sai Satwik[3], Mallela Chaitanya Krishna[4], Muppuri Naga Vamsi Kiran[5]

Associate Professor, Computer Science and Engineering, VVIT, Guntur, India[1]

Undergraduate, Computer Science and Engineering, VVIT, Guntur, India[2-5]

Abstract: One Important application of natural language processing (NLP) is the recognition of emotions. The main objective of this project is to develop a bot that could talk based on the current emotional situation of the user. This bot can detect the emotion of the user by fetching their facial expression and analysing them based on previously trained models using a dataset. Usually, the chatbots or any other bots doesn't consider the user emotion in any ways. If they would like to consider the emotion the bot would just ask the user emotion then the user should specify the emotion, and this may manipulate the original emotion of the user. In this project the emotion is captured from the facial expression of the user, and the user could communicate to the system in any language (mentioned in the project), and the user could expect the response in the same language in different slangs.

Keywords: CNN, NLP, NLTK, Django, Speech Recognition, Speech Synthesis.

I. INTRODUCTION

In today's ever-evolving digital landscape, human-computer interaction has reached new heights, thanks to advancements in Natural Language Processing (NLP) and computer vision. As we navigate through a world increasingly interconnected through technology, there is a growing need for more empathetic and responsive virtual companions. One critical aspect of human interaction that has often been overlooked by artificial intelligence systems is the recognition of human emotions.

This project aims to bridge that gap by introducing an innovative solution – a chatbot capable of engaging in conversations tailored to the user's current emotional state. Unlike traditional chatbots that rely on explicit user input for emotional context, our chatbot harnesses the power of facial expression analysis and deep learning techniques to detect and adapt to the user's emotion in real-time.

This groundbreaking approach not only eliminates the need for users to manually specify their emotions but also offers a more natural and immersive interaction experience.

Imagine a virtual companion that can understand your emotions through your webcam's lens, allowing you to communicate in any language and even various slangs, while receiving responses that match your mood. Whether you're feeling happy, sad, excited, or any emotion in between, this chatbot will be your empathetic and supportive companion, providing you with engaging and personalized conversations.

In this introduction, we will delve into the project's objectives, methods, and potential impact, showcasing how this innovative application of NLP and computer vision technology can revolutionize the way we interact with artificial intelligence systems, making them not to just intelligent but emotionally aware.

Join us on a journey to explore the future of human-computer interaction, where your emotions are no longer hidden from the virtual companion your emotion-aware chatbot.

II. LITERATURE REVIEW

Natural Language Processing (NLP) has seen significant advancements in recent years, with one important application being the recognition of emotions to enhance human-computer interaction. Emotion recognition enables systems to adapt their responses based on the user's emotional state, thereby improving the overall user experience. Traditional approaches often rely on explicit user input or text analysis, but recent developments have explored the use of facial expression analysis for more accurate emotion detection.

DOI: 10.17148/IJARCCE.2024.13248

Advancements in computer vision and machine learning have greatly contributed to the development of emotion recognition systems. Bartlett et al. (2006) utilized machine learning techniques to automatically classify facial expressions into basic emotions, achieving high accuracy rates. Similarly, Liu et al. (2018) proposed a deep learning-based approach for facial expression recognition, which outperformed traditional methods by capturing more subtle facial cues.

In summary, the integration of facial expression analysis with NLP holds great potential for creating more empathetic and responsive human-computer interfaces. Future research directions may focus on improving the accuracy and scalability of emotion recognition models, as well as addressing ethical and privacy concerns to ensure the responsible deployment of these technologies.

III. PROPOSED METHODOLOGY

The proposed method for our Enhanced AI Bot with Facial Emotion Detection consists of several key components:

A. Emotion Recognition using Convolutional Neural Network

Convolutional Neural Networks (CNNs) are employed to analyse and detect emotions from real-time webcam data. This component enables the chatbot to continuously assess the user's emotional state during interactions.

A CNN is a DL algorithm which takes an input image, assigns importance (learnable weights and biases) to various aspects/objects in the image and is able to differentiate between images. The preprocessing required in a CNN is much lower than other classification algorithms. Figure 1 shows the CNN operations.

The architecture of a CNN is analogous to that of the connectivity pattern of neurons in the human brain and was inspired by the organization of the visual cortex [32]. One of role of a CNN is to reduce images into a form which is easier to process without losing features that are critical for good prediction.

This is important when designing an architecture which is not only good at learning features but also is scalable to massive datasets. The main CNN operations are convolution, pooling, flatten, dense and dropout which are described below.

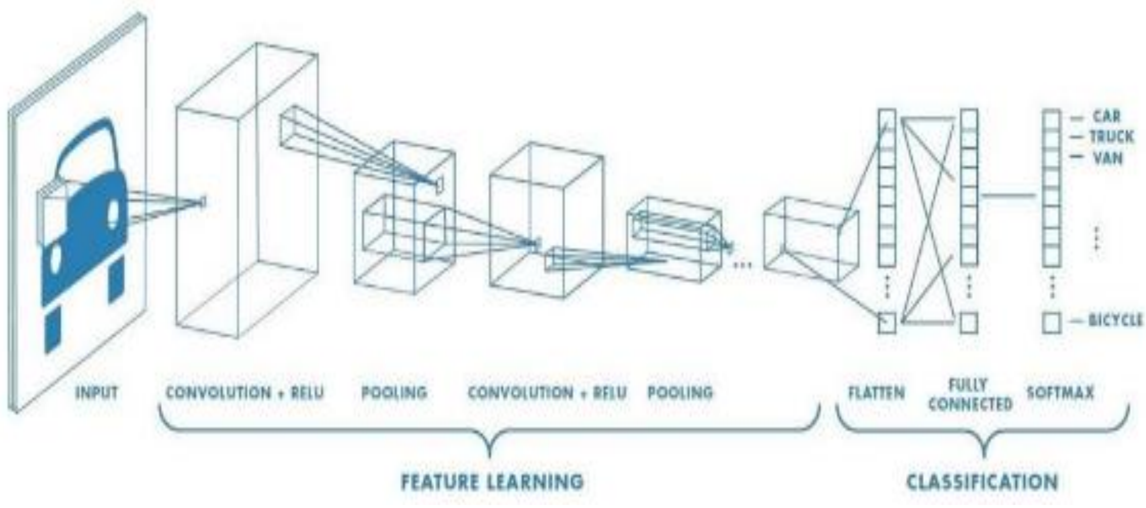

Fig. 1 Operations of Convolutional Neural Network

1) Conv2D: This layer performs 2-D convolutions, which are essentials for feature extraction from input data like images. It applies a set of learnable filters to the input data to create feature maps.

ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⫶ Peer-reviewed & Refereed journal ⫶ Vol. 13, Issue 2, February 2024
DOI: 10.17148/IJARCCE.2024.13248

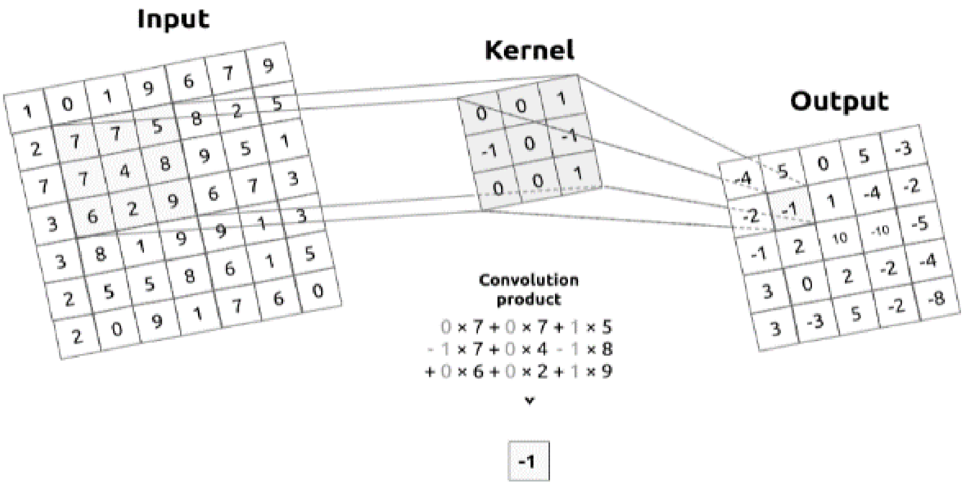

Fig. 2 Process of Convolution Layer

2) *MaxPooling2D:* This reduces the spatial dimensions of feature maps by selecting the maximum value within each local region. This helps in downscaling and reducing the computational load.

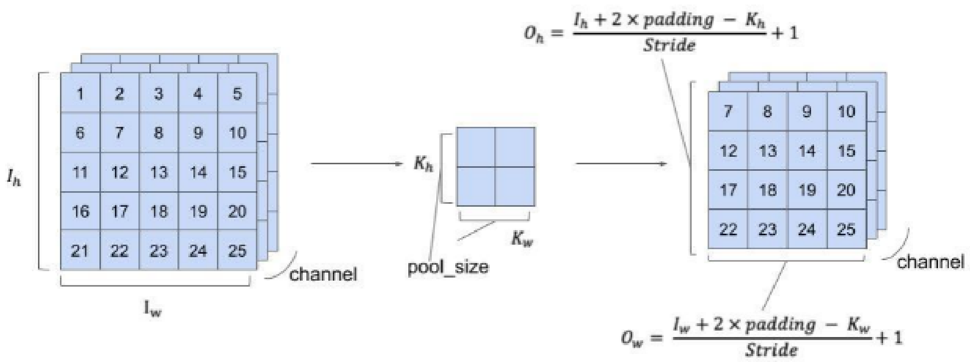

Fig. 3 Process under Pooling Layer

3) *Dropout (0.1):* This layer randomly deactivates a specified portion of neurons during training (e.g. 10% in this case) to prevent overfitting and improve model generalization.

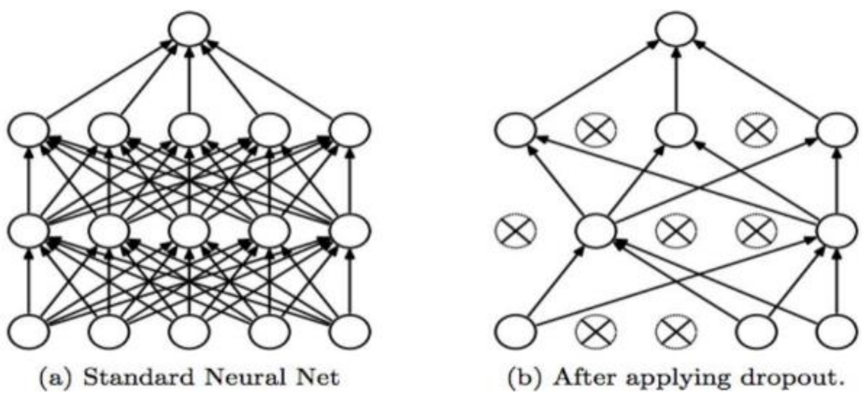

Fig. 4 Neural Network before and after applying the Dropout Layer

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ≋ Peer-reviewed & Refereed journal ≋ Vol. 13, Issue 2, February 2024

DOI: 10.17148/IJARCCE.2024.13248

4) *Flatten:* This layer reshapes the 2D feature maps into 1D vector, which is required before feeding the data into fully connected layers.

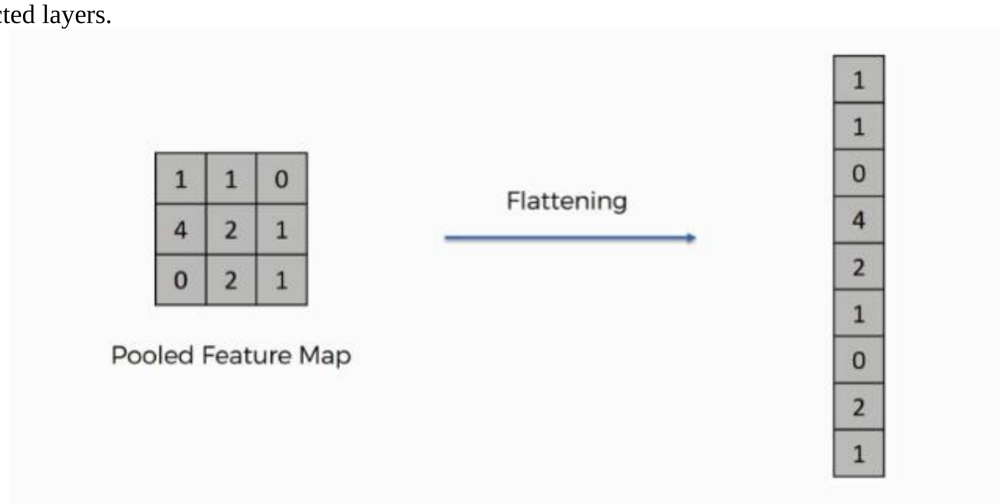

Fig. 5 Working of Flatten Layer

5) *Dense:* This layer connects every neuron from the previous layer to every neuron in the current layer. It is typically used in the final layers for classification or regression tasks

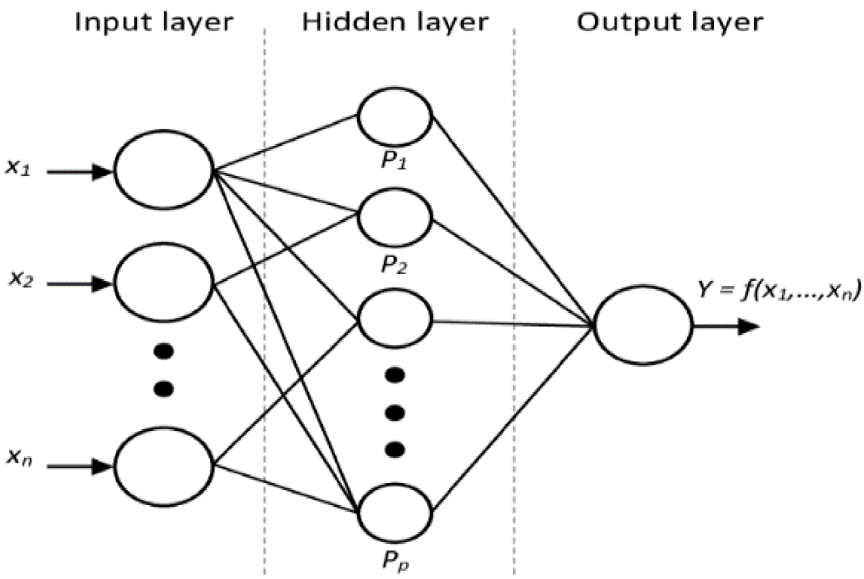

Fig. 6 Working of Dense Layer

B. *Transformer Models for Response Generation*

Transformer – based models such as GPT (Generative Pre-trained Transformer) or BERT (Bidirectional Encoder Representations from Transformers), serve as the core of our chatbot's response generation system. These models generate contextually relevant responses based on the conversation context and the detected emotional state of the user.

C. Language Translation

Language translation capabilities are integrated to ensure that users can communicate with the chatbot in their preferred language. This feature enhances inclusivity and user engagement, accommodating a diverse range of languages.

D. Text-to-Voice Conversation

Text-to-voice conversation functionality allows the chatbot to not only respond in text but also audibly communicate with users. This feature adds an additional layer of interactivity and accessibility to the chatbot's interactions.

DOI: 10.17148/IJARCCE.2024.13248

Text-to-voice conversation functionality empowers the chatbot to respond in text and audibly communicate with users. This feature adds a layer of interactivity and accessibility to the chatbot's interactions, enhancing the overall user experience.

E. Voice-to-Text Conversation

Conversely, voice-to-text conversation functionality enables the chatbot to transcribe spoken words into text, providing users with an additional means of input. This feature enhances accessibility and allows users to interact with the chatbot through speech, further diversifying the avenues of communication available.

IV. IMPLEMENTATION

The implementation of the project involves several components, including a Django backend for handling user requests, HTML pages for user interaction, an OpenAI integration for generating responses, and a Convolutional Neural Network (CNN) model for facial emotion detection using the FER 2013 dataset.

Below is an overview of the implementation steps:

A. Setting up Django Backend

- Create a Django project and set up the necessary configurations.
- Define Django models for storing user data.
- Implement Django views and URL routing to handle user requests and responses.

B. Integrating HTML Pages

- Design HTML pages for user interaction, including a user input form and a display area for bot responses.
- Use CSS for styling and JavaScript for any dynamic interactions if needed.
- Ensure proper communication between HTML pages and Django views through form submissions or AJAX requests.

C. Implementing Facial Emotion Detection

- Train a CNN model using the FER 2013 dataset for facial emotion detection.
- Preprocess the input facial images (e.g., resize, normalize) before feeding them into the model.
- Integrate the trained model into the Django backend to analyze facial expressions captured from user-provided images.
- Implement error handling and validation to handle cases where facial expression analysis fails or returns inconclusive results.

D. Integrating OpenAI for Response Generation

- Set up an OpenAI API account and obtain the necessary credentials for accessing the GPT-3 model.
- Integrate the OpenAI API into the Django backend to generate responses based on user input.
- Ensure proper error handling and rate limiting to comply with OpenAI's usage guidelines.

E. Handling Multilingual Support and Slang

- Implement language detection to identify the language of user input.
- Utilize libraries or APIs for language translation if needed to process user input and generate responses in multiple languages.
- Develop mechanisms to handle slang and informal language in user input and bot responses to enhance user engagement.

F. Testing and Deployment

- Test the integrated system thoroughly to ensure functionality and performance.
- Deploy the Django application to a web server or cloud platform for public access.
- Monitor system performance and user feedback for continuous improvement.

V. RESULTS

The results of the project will be as follows:

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 2, February 2024

DOI: 10.17148/IJARCCE.2024.13248

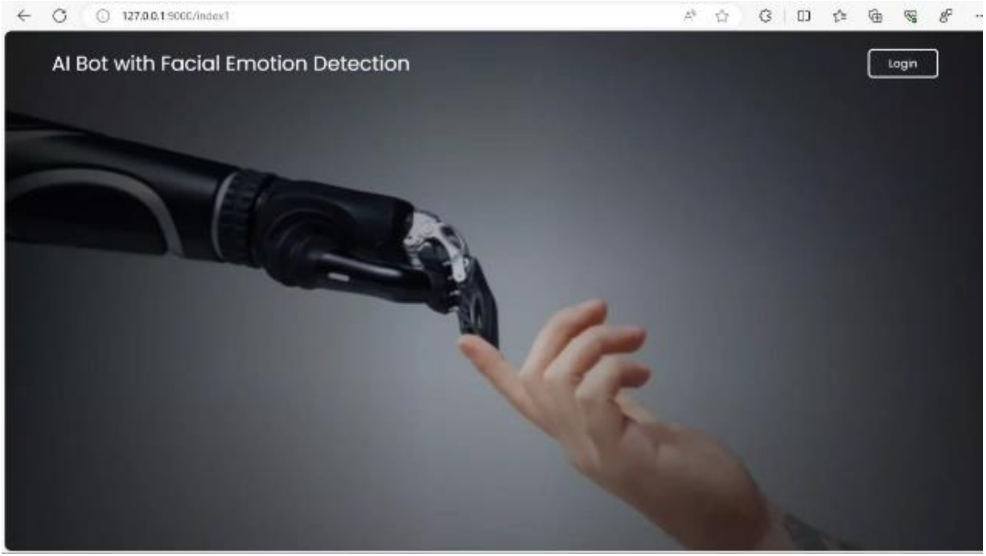

Fig.7 User Interface

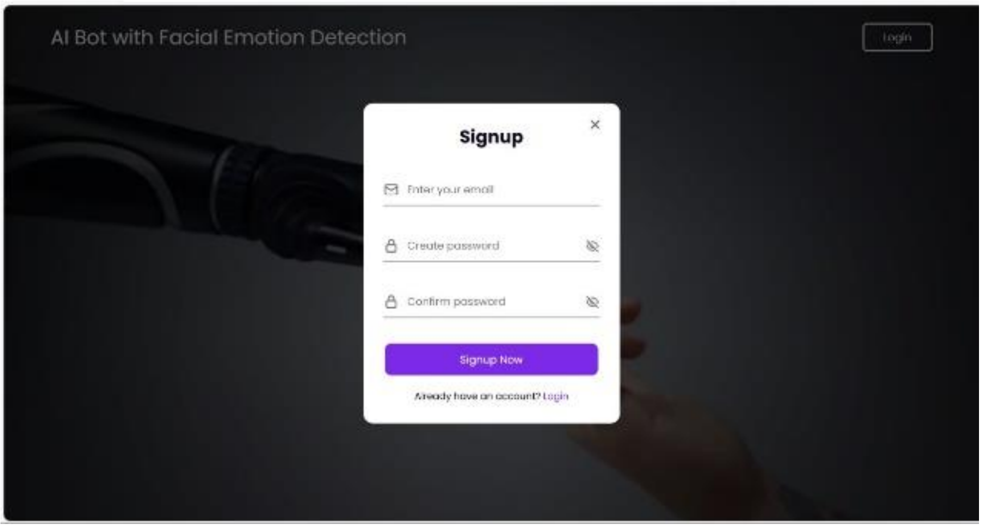

Fig.8 Sign-up page

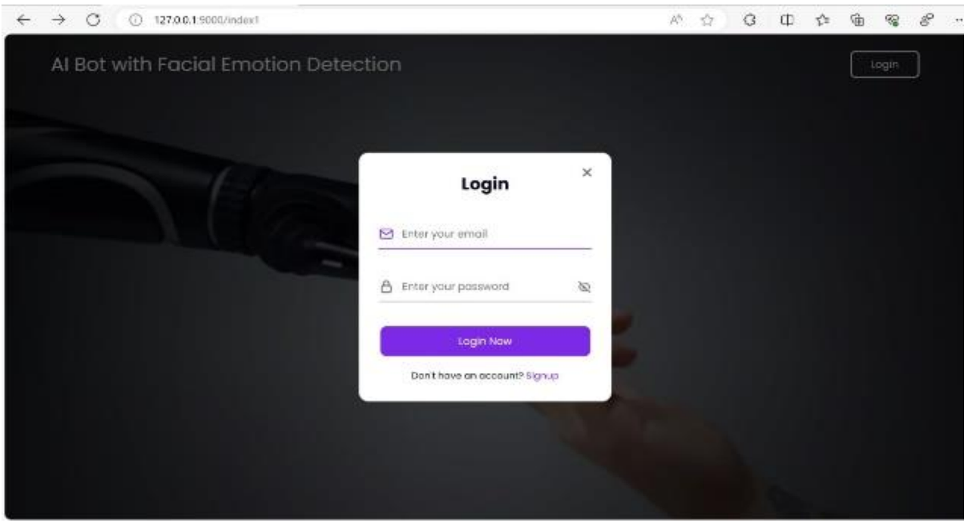

Fig.9 Login page

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 2, February 2024

DOI: 10.17148/IJARCCE.2024.13248

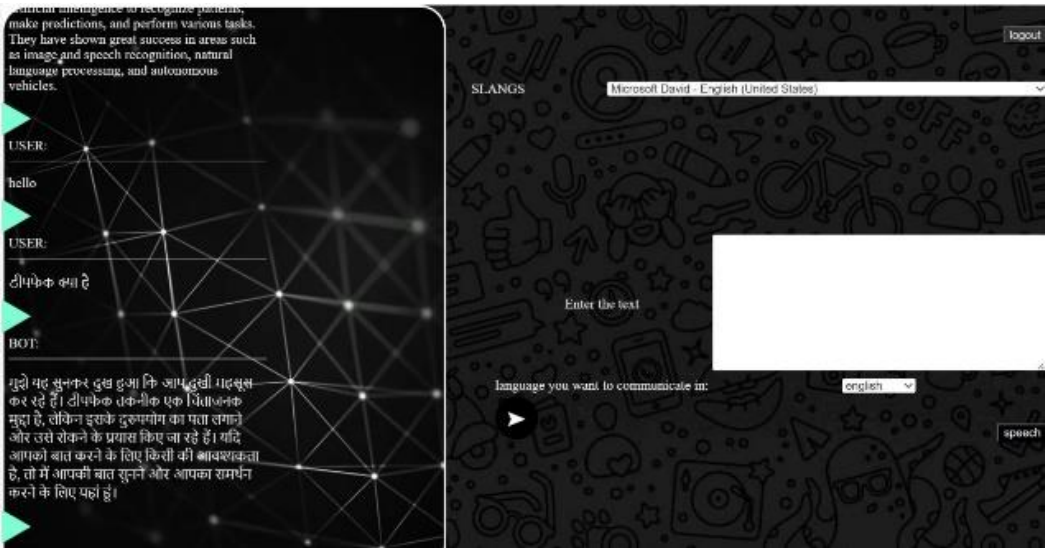

Fig.10 Chat-Bot page

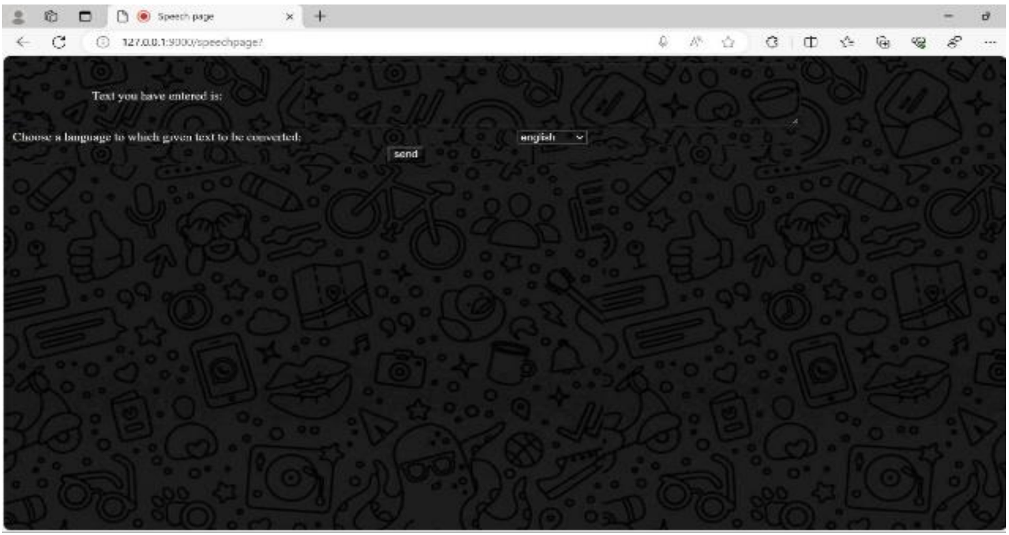

Fig.11 Speech page

VI. CONCLUSION

This Emotion-Aware Chatbot project has successfully implemented a Convolutional Neural Network for a real-time emotion detection, allowing the chatbot to dynamically adapt its responses based on user emotions. Utilizing Django for the interactive interface, the chatbot significantly enhances user experience and engagement. This application may be incorporated in any industries, with particular promise in mental health support. Overall, the Emotion-Aware chatbot presents a pioneering step in human-computer interaction, show casting the potential of AI to understand and respond to human emotions.

VII. LIMITATIONS

The Limitations for the project may be as follows:

A. Limited Emotion Recognition Accuracy

While using a CNN model with the FER 2013 dataset for facial emotion detection provides a reasonable level of accuracy, it may still struggle with certain facial expressions, lighting conditions, or cultural nuances, leading to inaccuracies in emotion recognition.

B. Dependency on Facial Expressions

The project heavily relies on capturing facial expressions for emotion detection. This approach may not be suitable for scenarios where facial data is unavailable or where users prefer not to share facial images due to privacy concerns.

C. Response Generation Quality

Although OpenAI's GPT-3 model excels at generating human-like text, it may still produce nonsensical or inappropriate responses in certain contexts, leading to a degradation in user experience or even offensive output.

D. Computational Resource Requirements

Training and deploying the CNN model for facial emotion detection, as well as accessing the OpenAI API for response generation, require significant computational resources. This may limit the scalability of the system, especially for large user bases or high traffic scenarios.

VIII. FUTURE WORK

The Future work for the project may be as follows:

A. *Enhanced Emotion Recognition Models*

Develop and integrate more advanced machine learning or deep learning models for emotion recognition, capable of capturing subtle facial cues and improving accuracy across diverse populations and cultural contexts. Explore techniques such as multi-modal emotion recognition combining facial expressions with other modalities like voice or text for more robust emotion detection.

B. *Commercialization and Deployment*

Explore opportunities for commercialization and deployment of the emotion-aware bot in various domains such as customer service, mental health support, education, or entertainment. Collaborate with industry partners to refine the system for specific use cases and market segments, ensuring scalability, reliability, and regulatory compliance.

ACKNOWLEDGEMENTS

We would like to express our sincere gratitude to everyone who contributed to the successful completion of this project. First and foremost, we extend our heartfelt appreciation to the developers and contributors of the open-source tools and libraries that formed the backbone of our implementation, including Django, OpenAI, and the FER 2013 dataset creators. We are deeply thankful to our Project Coordinator **Dr N. Sri Hari** and Project Guide **Dr O. Aruna** for their invaluable guidance, support, and insightful feedback throughout the project's development. Their expertise and encouragement have been instrumental in shaping our approach and overcoming challenges.

My classmates and coworkers have helped me polish concepts, overcome obstacles, and meet project goals by providing assistance, suggestions, and teamwork throughout the duration of the project.

This project would not have been possible without the collective efforts and support of everyone mentioned above We are truly grateful for their contributions and collaboration.

REFERENCES

[1]. Azizan, I., & Khalid, F. (2018, December). Facial emotion recognition: a brief review. In *International Conference on Sustainable Engineering, Technology and Management*.

[2]. Srivastava, V., & Shyam, R. (2021). Enhanced object detection with deep convolutional neural networks. *International Journal of All Research Education and Scientific Methods (IJARESM)*, *9*(7), 27-36.

[3]. Noroozi, F., Marjanovic, M., Njegus, A., Escalera, S., & Anbarjafari, G. (2017). Audio-visual emotion recognition in video clips. *IEEE Transactions on Affective Computing*, *10*(1), 60-75.

[4]. Fathallah, A., Abdi, L., & Douik, A. (2017, October). Facial expression recognition via deep learning. In *2017 IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA)* (pp. 745-750). IEEE.

[5]. Wiskott, L., & Von Der Malsburg, C. (1996). Recognizing faces by dynamic link matching. *Neuroimage*, *4*(3), S14-S18.

[6]. Rouast, P. V., Adam, M. T., & Chiong, R. (2019). Deep learning for human affect recognition: Insights and new developments. *IEEE Transactions on Affective Computing*, *12*(2), 524-543.

DOI: 10.17148/IJARCCE.2024.13248

[7]. Chandra, Y. W., & Suyanto, S. (2019). Indonesian chatbot of university admission using a question answering system based on sequence-to-sequence model. *Procedia Computer Science*, *157*, 367-374.

[8]. Akhand, M. A. H., Roy, S., Siddique, N., Kamal, M. A. S., & Shimamura, T. (2021). Facial emotion recognition using transfer learning in the deep CNN. *Electronics*, *10*(9), 1036.

[9]. Jabeen, S., Mehmood, Z., Mahmood, T., Saba, T., Rehman, A., & Mahmood, M. T. (2018). An effective content-based image retrieval technique for image visuals representation based on the bag-of-visual-words model. *PloS one*, *13*(4), e0194526.

[10]. Saleem, S., Ghani Khan, M. U., Saba, T., Abunadi, I., Rehman, A., & Bahaj, S. A. (2022). Efficient Facial Recognition Authentication Using Edge and Density Variant Sketch Generator. *Computers, Materials & Continua*, *70*(1).

[11]. Baez, M., Daniel, F., & Casati, F. (2019, November). Conversational web interaction: proposal of a dialog-based natural language interaction paradigm for the web. In *International Workshop on Chatbot Research and Design* (pp. 94-110). Cham: Springer International Publishing.

[12]. Baxter, J., & Bartlett, P. L. (2001). Infinite-horizon policy-gradient estimation. *journal of artificial intelligence research*, *15*, 319-350.

[13]. Li, B. Y., Mian, A. S., Liu, W., & Krishna, A. (2013, January). Using kinect for face recognition under varying poses, expressions, illumination and disguise. In *2013 IEEE workshop on applications of computer vision (WACV)* (pp. 186-192). IEEE.

[14]. Adam, M., Wessel, M., & Benlian, A. (2021). AI-based chatbots in customer service and their effects on user compliance. *Electronic Markets*, *31*(2), 427-445.

[15]. Shivam, K., Saud, K., Sharma, M., Vashishth, S., & Patil, S. (2018). Chatbot for college website. *Int J Comp Technol*, *5*(6), 74-77.

[16]. Babu, E., & Wilson, G. (2021). Chatbot for college enquiry. *International Journal of Creative Research Thoughts*.

[17]. Colace, F., De Santo, M., Lombardi, M., Pascale, F., Pietrosanto, A., & Lemma, S. (2018). Chatbot for e-learning: A case of study. *International Journal of Mechanical Engineering and Robotics Research*, *7*(5), 528-533.

[18]. Ekman, P. (1973). Cross-cultural studies of facial expression. *Darwin and facial expression: A century of research in review*, *169222*(1), 45-60.