e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:05/May-2024 Impact Factor- 7.868 www.irjmets.com

3WAY EMOTION DETECTION

Srinidhi kulkarni*1, Gnanesh CS*2, Punith SY*3, Sanjana A Hiremath*4, Sanjay HP*5

*1Assistant Professor, Department Of Computer Science And Engineering, Jyothy Institute Of Technology, Bengaluru, Karnataka, India.

*2,3,4,5Under Graduate Student, Department Of Computer Science And Engineering, Jyothy Institute Of Technology, Bengaluru, Karnataka, India.

ABSTRACT

This project aims to develop a comprehensive emotion detector utilizing three distinct modalities: facial expressions, voice characteristics, and text analysis. The first component employs Convolutional Neural Networks (CNN) to analyze facial expressions and recognize emotions from images. The second facet utilizes the librosa library for audio processing, employing CNNs and Mel- Frequency Cepstral Coefficients (MFCC) concepts to extract emotional cues from voice data. The third component leverages Natural Language Processing (NLP) and machine learning techniques to analyze textual content for emotional sentiment. By combining these three modalities, the system provides a holistic understanding of the user's emotional state.

Upon emotion detection, the system offers personalized recommendations to enhance the user's well- being. Recommendations include nearby healthcare professionals or therapists for emotional support based on detected emotions, curated movie and podcast suggestions to align with the user's current mood, and tailored news articles to keep the user informed without causing undue stress. Furthermore, the system integrates with YouTube to provide emotion-aware video recommendations, creating a user-centric platform that enhances their emotional and mental well being by offering relevant and empathetic content based on their emotional state. This multifaceted approach not only showcases the integration of diverse technologies but also demonstrates the potential for a more nuanced and user- specific recommendation system in various domains.

I. INTRODUCTION

In the fast-paced and interconnected world of today, understanding and responding to human emotions have become pivotal for enhancing user experience and well-being. This project embarks on the development of an advanced emotion detector that harnesses the power of three key modalities: facial expressions, voice characteristics, and textual content. Recognizing that emotions manifest in multifaceted ways, we employ Convolutional Neural Networks (CNNs) to analyze facial expressions, librosa and CNNs for voice emotion detection using Mel-Frequency Cepstral Coefficients (MFCC), and natural language processing (NLP) coupled with machine learning for extracting sentiments from textual data. By synergistically combining these technologies, we aim to create a more holistic and accurate representation of the user's emotional state.

The primary goal of this project extends beyond mere emotion detection; it endeavors to leverage the detected emotional states to provide personalized recommendations for the user. These recommendations are designed to cater to the user's emotional needs, offering suggestions such as nearby healthcare professionals or therapists based on detected emotions, curated movie and podcast selections aligned with the user's current emotional state, and tailored news articles to maintain a balanced and informed perspective. In addition, the project integrates with YouTube, offering emotion-aware video recommendations, thereby creating a unique and empathetic user experience that contributes positively to their mental and emotional health. Through this project, we aspire to showcase the potential of integrating diverse technologies to create a sophisticated, emotion-aware system that enhances user engagement and satisfaction in a variety of domains.

In an increasingly interconnected world, the importance of mental health and emotional well- being cannot be overstated. Despite advancements in technology and heightened awareness of mental health issues, many individuals still struggle to effectively manage their emotions and seek appropriate support. Traditional methods of emotion detection and support often lack the depth and personalization required to address the complex emotional needs of individuals.

Recognizing this gap, there is a pressing need for more sophisticated and user- centric approaches to emotional well-being.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:05/May-2024 Impact Factor- 7.868 www.irjmets.com

The motivation behind this project lies in addressing the shortcomings of existing emotion detection systems and providing a comprehensive solution that integrates cutting-edge technologies. By harnessing the power of facial expression analysis, voice emotion detection, and textual sentiment analysis, we aim to develop a holistic understanding of the user's emotional state. This multifaceted approach not only captures a wider range of emotional cues but also enables more accurate and personalized support for individuals navigating their emotional landscape.

Furthermore, this project is motivated by the potential of technology to bridge the gap between individuals and the support they need. By leveraging machine learning algorithms and natural language processing techniques, we aspire to create a supportive ecosystem that empowers individuals to better understand and manage their emotions. Through personalized recommendations for emotional support resources, curated media content, and empathetic engagement, we aim to foster emotional resilience and flourishing in individuals across diverse communities.

Moreover, this project aligns with the broader societal shift towards prioritizing mental health and well-being. As conversations around mental health become increasingly normalized, there is a growing demand for innovative solutions that cater to individual needs and preferences. By developing an emotion detection system that combines advanced technology with empathetic design principles, we seek to contribute to the advancement of user-centric technology solutions that prioritize emotional well-being.

Ultimately, the motivation behind this project stems from a shared vision of creating a more inclusive and supportive digital landscape that fosters emotional well-being for all. By leveraging the power of technology to gain deeper insights into human emotions and provide personalized support, we aspire to make a meaningful impact on the lives of individuals, empowering them to lead healthier and more fulfilling lives.

II. METHODOLOGY

The proposed methodology for this groundbreaking project involves a multi-modal approach to emotion detection, transcending the limitations of existing single-modal systems. We seamlessly integrate advanced technologies, including Convolutional Neural Networks (CNN) for facial analysis, Librosa and CNN for audio processing, and Long Short-Term Memory (LSTM) networks coupled with Natural Language Processing (NLP) for text sentiment analysis.

This comprehensive integration forms a dynamic platform capable of interpreting emotions across diverse modalities, enhancing accuracy and introducing adaptability crucial for real-time interactions.

Existing System

Current emotion detection systems often rely on isolated modalities, such as facial expression analysis or voice emotion detection, resulting in a limited understanding of the complex nature of human emotions. These systems typically lack integration across modalities and fail to provide personalized recommendations based on the detected emotions. Additionally, many emotion detectors lack the sophistication to interpret emotions from textual content, overlooking a crucial dimension of user emotion. The absence of a comprehensive and integrated approach in the existing systems hinders their ability to deliver a nuanced understanding of user emotions and, consequently, limits their effectiveness in providing tailored recommendations for user engagement and well-being.

DRAWBACKS Limited scope due to unimodal focus, neglecting the holistic nature of human emotions . Lack of integration across modalities hampers accuracy in emotion detection.Absence of personalized recommendations based on detected emotions limits user engagement and well-being enhancement.

The proposed system aims to overcome the limitations of existing emotion detection frameworks by integrating facial expression analysis, voice emotion detection, and text sentiment analysis into a unified and sophisticated model. Leveraging Convolutional Neural Networks (CNNs) for facial expressions, librosa for voice features, and Natural Language Processing (NLP) for textual content, the system ensures a comprehensive understanding of user emotions. By seamlessly combining these modalities, our approach provides a holistic and nuanced representation of emotional states. Moreover, the proposed system goes beyond detection, offering personalized recommendations based on the user's emotions, including suggestions for nearby healthcare professionals, tailored movie and podcast selections, curated news articles, and emotion-aware

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:05/May-2024 Impact Factor- 7.868 www.irjmets.com

YouTube content. This holistic and user centric approach is designed to significantly enhance the overall user experience, fostering emotional well-being through empathetic and contextually relevant content recommendations.

III. MODELING AND ANALYSIS

Modalities Integration Feasibility: The feasibility of integrating three distinct modalities facial expressions analysis, voice characteristics assessment, and text sentiment analysis is assessed. This involves evaluating the compatibility and interoperability of different technologies, such as Convolutional Neural Networks (CNNs), librosa library for audio processing, and Natural Language Processing (NLP) techniques.

Technological Viability: The viability of employing CNNs for facial expression recognition, utilizing Mel-Frequency Cepstral Coefficients (MFCC) for voice data analysis, and applying NLP and machine learning techniques for textual sentiment analysis is examined. This includes assessing the availability of relevant libraries, tools, and datasets for each modality.

Emotion Detection Accuracy: The accuracy and reliability of emotion detection using each modality individually and in combination are investigated. This involves conducting preliminary experiments and feasibility tests to determine the effectiveness of Convolutional Neural Networks (CNNs), MFCC concepts, and NLP algorithms in accurately detecting and interpreting emotions from facial expressions, voice characteristics, and textual content.

Recommendation System Viability: The feasibility of implementing a recommendation system based on detected emotions is explored. This includes assessing the availability of relevant data sources for generating personalized recommendations, such as healthcare professional databases, movie and podcast libraries, news articles, and YouTube video recommendations.

User Acceptance and Ethical Considerations: The potential acceptance of the proposed emotion detection and recommendation system by users is evaluated. Additionally, ethical considerations related to user privacy, data security, and the responsible use of emotional data are addressed to ensure the system's ethical integrity and compliance with regulations.

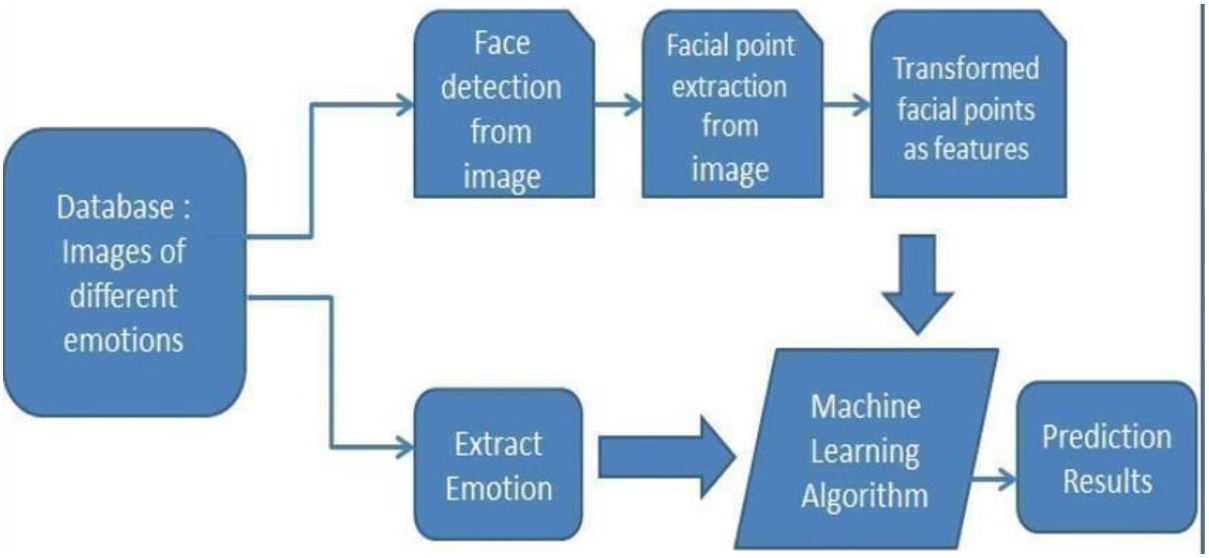

Figure 1: Facial Emotion Recognition Using Machine Learning.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:05/May-2024 Impact Factor- 7.868 www.irjmets.com

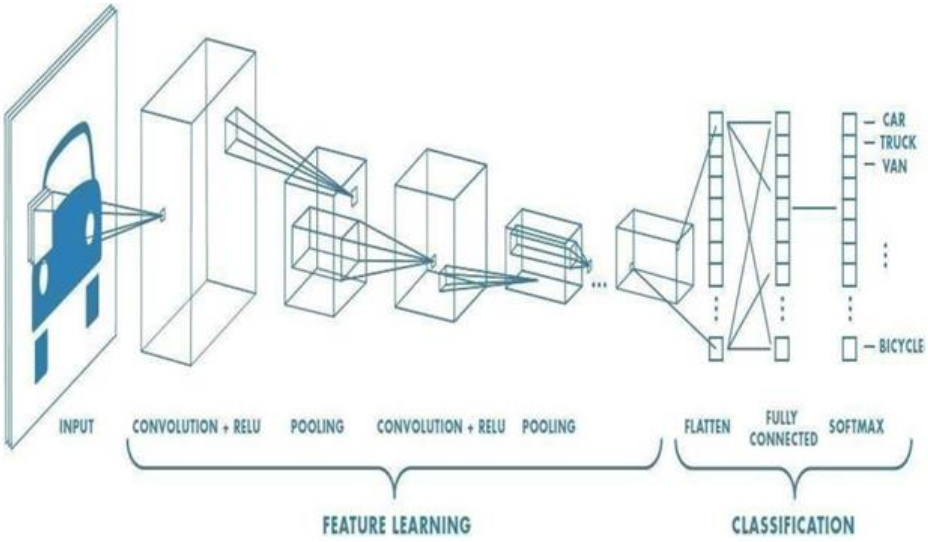

Figure 2: Neural network with many convolutional layers

IV. RESULTS AND DISCUSSION

The working of emotion based music player is demonstrated in the following content with the help of snapshots.

Figure 3: Login page

Figure 4: User Interface

The above figure shows the interface that will be displayed to the user on entering the application. It provides three entry prompts to the user. The first on is the camera which allow users to record the facial emotion. The

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:05/May-2024 Impact Factor- 7.868 www.irjmets.com

second is recorder which is used to record user's voice to detect the emotion. The third is the text which is used to detect the emotion of a person using the text as input.

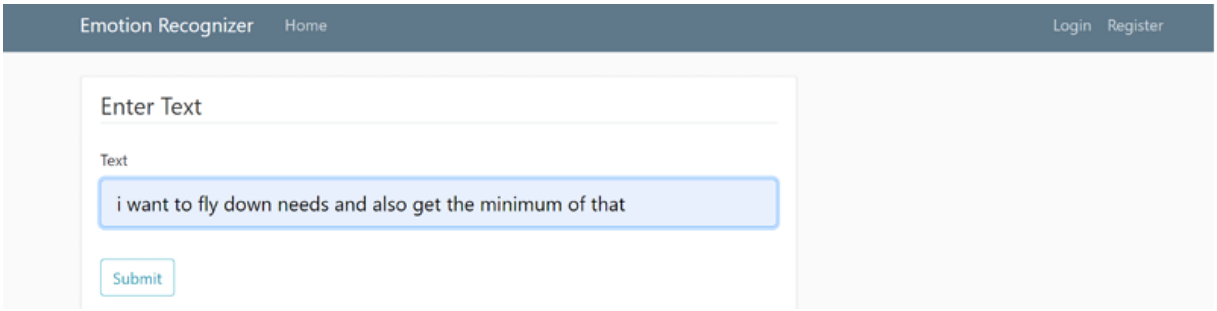

Figure 5: Text-Based Emotion Detection.

Figure 6: Device camera captures person's face and detects the emotion .

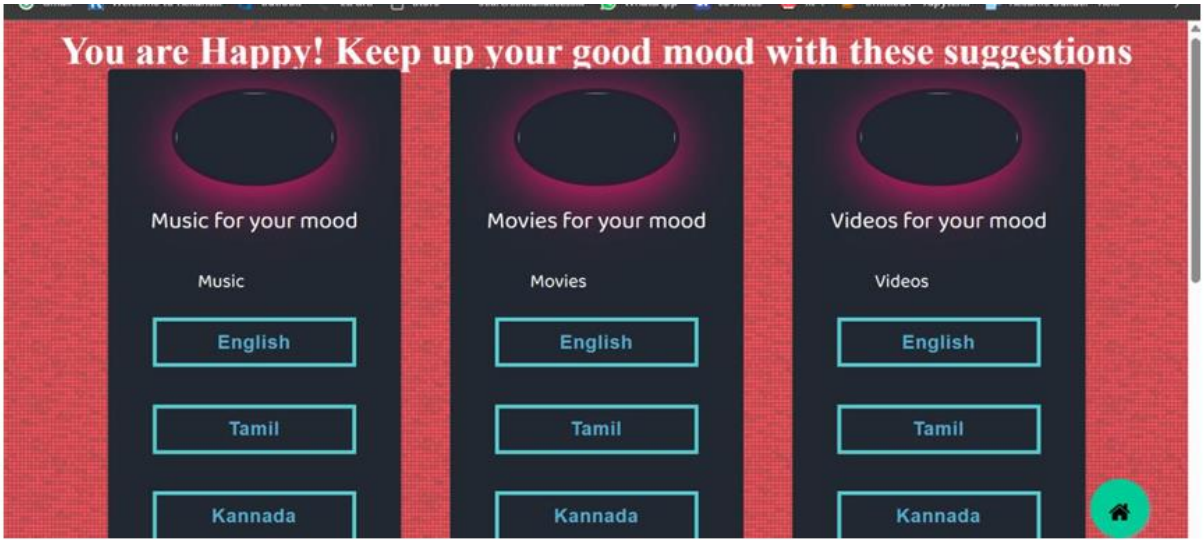

Figure 7: User is happy, so the page will be redirected to happy site .

V. CONCLUSION

In this project, we presented a music recommendation system based on emotion detected. The system uses a two-layer convolution network model for facial emotion recognition. The model classifies 7 different facial emotions from the image dataset. The model has comparable training accuracy and validation accuracy which convey that the model is having the best fit and is generalized to the data. We also recognize the room for improvement. It would be interesting to analyze how the system performs when additional emotions are taken

into consideration. User preferences can be collected to improve the overall system using collaborative filtering. We plan to address these issues in future work.

VI. REFERENCES

[1] X. Zhang, G. Qi, X. Fu and N. Zhang, "Robust Emotion Recognition Across Diverse Scenes: A Deep Neural Network Approach Integrating Contextual Cues," in IEEE Access, vol. 11, pp. 73959-73970, 2023, doi: 10. 1109/ACCESS.2023.3296316.

[2] A. R. Jafari, G. Li, P. Rajapaksha, R. Farahbakhsh and N. Crespi, "Fine-Grained Emotions Influence on Implicit Hate Speech Detection," in IEEE Access, vol. 11, pp. 105330-105343, 2023, doi: 10. 1109/ACCESS. 2023.3318863. in

[3] A. M. Yildiz et al., "FF-BTP Model for Novel Sound-Based Community Emotion Detection," IEEE Access, 10. 1109/ACCESS.2023.3318751. vol. 11, pp. 108705-108715, 2023, doi:

[4] F. Alrowais et al., "Modified Earthworm Optimization With Deep Learning Assisted Emotion Recognition for Human Computer Interface," in IEEE Access, vol. 11, pp. 35089 35096, 2023.

[5] M. Al-Hashedi, L. -K. Soon, H. -N. Goh, A. H. L. Lim and E. -G. Siew, "Cyberbullying Detection Based on Emotion," in IEEE Access, vol. 11, pp. 53907-53918, 2023, doi: 10.1109/ACCESS.2023.3280556.