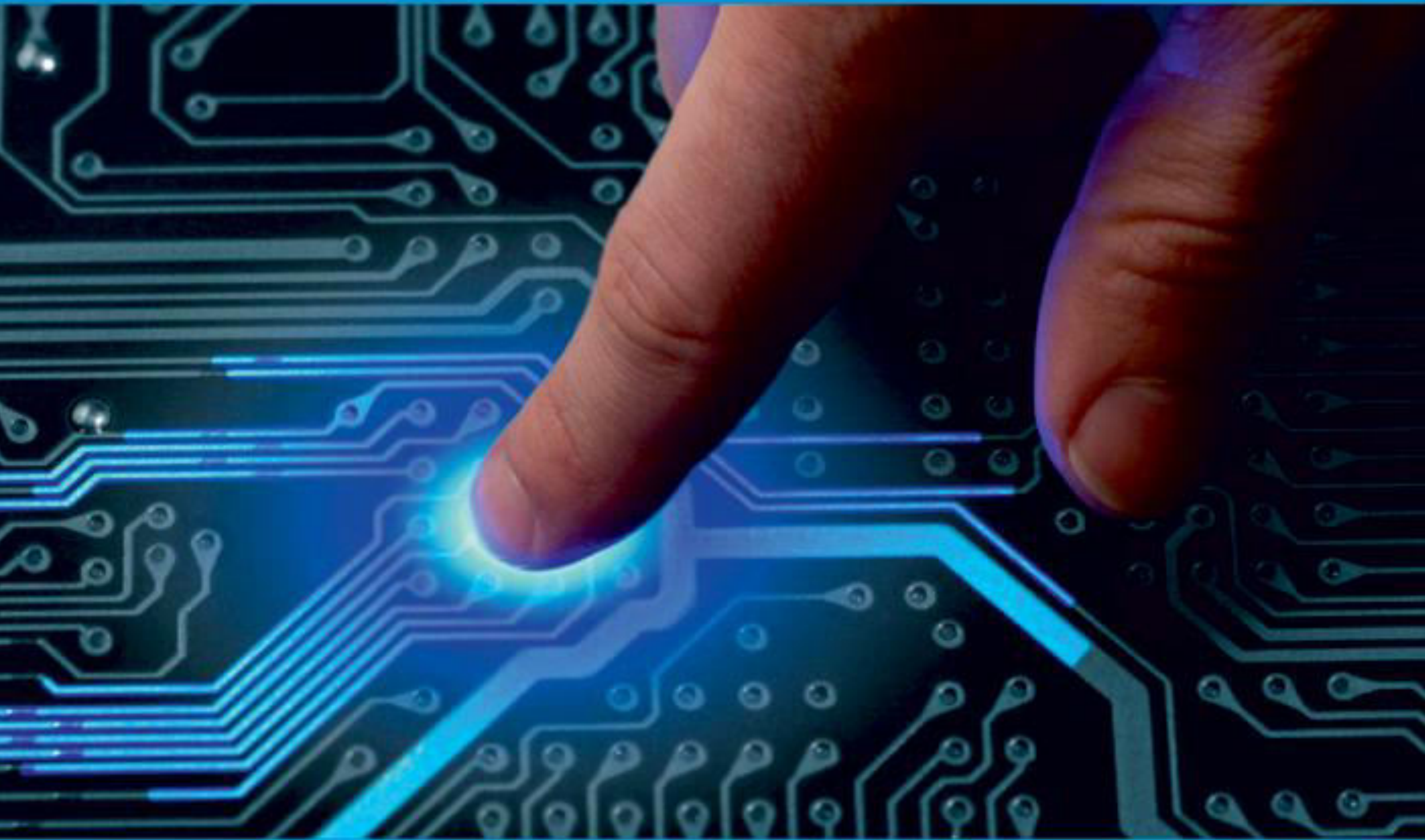
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 3, March 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

9940 572 462 6381 907 438 ijircce@gmail.com www.ijircce.com

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 3, March 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1203097 |

Emotion Based Music Player

Dr. O. Aruna[1] , Bhukya Venkata Sai Naik[2] , B.Chaturved[3], CH.Sambashiva Rao[4], A.Poojitha[5]

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

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

ABSTRACT: Music plays a significant role in people's daily lives, catering to their individual tastes and moods. However, manually browsing for music and creating playlists can be time-consuming. To address this issue, a proposed project aims to efficiently generate playlists based on users' current moods, using facial expressions as indicators. By capturing facial expressions through a webcam and analyzing them using a learning algorithm, the system identifies the predominant emotion. Subsequently, it suggests a playlist tailored to that emotion, thereby saving users time. The accuracy of this approach is evaluated using performance metrics to assess its effectiveness.

KEYWORDS: Haarcascade, CNN, Emotion Detection, Pandas, Scikit-Learn, Spotify API.

I. INTRODUCTION

The integration of emotion-based music players represents a significant advancement in how humans interact with and experience music. By leveraging technology to analyze and respond to users' emotional states, these systems offer a personalized and immersive music listening experience. The impact of emotion-based music players on humans is multifaceted and profound.

Moreover, emotion-based music players can have positive effects on mental health and well-being. Music has long been recognized for its therapeutic properties, and when coupled with the ability to adapt to the user's emotions in real-time, it can serve as a powerful tool for mood regulation and stress relief. For instance, during times of stress or sadness, the music player can intelligently select calming or uplifting tracks to help improve the user's mood.

The objective of this project is to suggest songs for users based on their mood by capturing facial expressions. Facial expressions are captured through webcam and such expressions are fed into learning algorithm which gives most probable emotion. Once the emotion is recognized, the system suggests a play-list for that emotion, thus saves a lot of time for a user. Once the emotion is detected by then the model generates a playlist according the emotion of the user.

Overall, emotion-based music players have the potential to significantly enhance the way we engage with and benefit from music. By recognizing and responding to our emotional cues, these systems can create personalized and impactful music experiences that positively impact our mood, well-being, and productivity.

II. LITERATURE REVIEW

Convolutional Neural Networks (CNNs) have emerged as powerful tools for various computer vision tasks, including image classification, object detection, and facial expression recognition. A plethora of research has been conducted to explore the capabilities and advancements of CNNs in different domains.[5]

One seminal work that popularized CNNs is the AlexNet architecture proposed by Krizhevsky et al. (2012). AlexNet demonstrated superior performance in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) by significantly outperforming traditional methods. This architecture consisted of multiple convolutional layers followed by max-pooling layers, leading to breakthroughs in image classification tasks.

Overall, CNNs have revolutionized the field of computer vision by providing robust and scalable solutions to various tasks. From image classification to object detection and facial expression recognition, CNNs continue to drive advancements in artificial intelligence and shape the future of visual perception.

The original Haar Cascade approach proposed by Viola and Jones utilized integral images and a boosting algorithm to efficiently search for object features. This technique allowed for rapid detection of objects by sequentially applying a set of weak classifiers to sub-regions of an image.[7]

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 3, March 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1203097 |

In summary, Haar Cascade classifiers represent a fundamental and influential technique in computer vision for object detection. Through continuous refinement and adaptation, they have become essential tools for various real-world applications, offering a balance of accuracy, efficiency, and ease of implementation.

III. PROPOSED METHODOLOGY

The methodologies used in the Emotion based music Recommendation System are:

A.Emotion Detection Using CNN:

A Convolutional Neural Network (CNN) is a type of deep learning algorithm commonly used for tasks involving image recognition, classification, and segmentation. CNNs are inspired by the organization of the visual cortex in animals and have been highly successful in various computer vision tasks due to their ability to automatically learn features from raw data.

CNNs consist of multiple layers, including convolutional layers, pooling layers, and fully connected layers. Here's a brief overview of each component:

Convolutional Layers: These layers apply convolutional operations to input images, using learnable filters (also known as kernels) to extract features such as edges, textures, and patterns. Convolutional layers preserve the spatial relationship between pixels by sliding the filters across the input image and computing dot products to produce feature maps.

Pooling Layers: Pooling layers reduce the spatial dimensions of feature maps while retaining important information. Common pooling operations include max pooling, which selects the maximum value from each region of the feature map, and average pooling, which computes the average value. Pooling helps to make the representation invariant to small translations and distortions in the input.

Activation Functions: Activation functions introduce non-linearity into the network, allowing it to learn complex mappings between input and output. Common activation functions used in CNNs include ReLU (Rectified Linear Unit), sigmoid, and tanh.

Fully Connected Layers: These layers connect every neuron from one layer to every neuron in the next layer, allowing the network to learn high-level features and make predictions. Fully connected layers are typically used in the final stages of the CNN architecture for classification tasks.

Dropout: Dropout is a regularization technique commonly used in CNNs to prevent overfitting. During training, dropout randomly deactivates a fraction of neurons in the network, forcing the model to learn more robust features and reducing reliance on specific neurons.

Fig1 .Face Recognization

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 3, March 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1203097 |

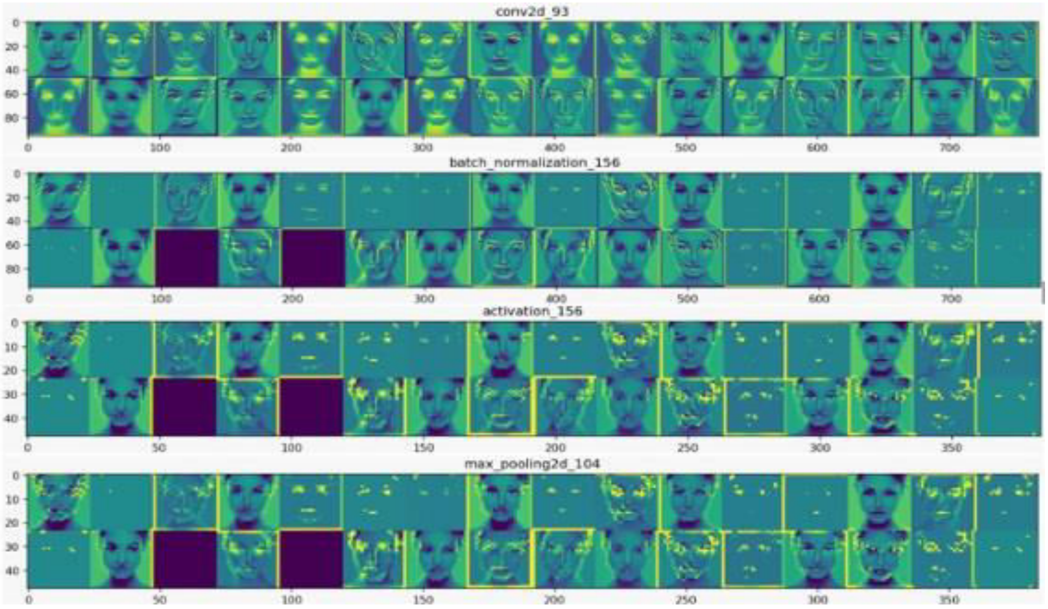


Fig 2: Feature Extraction

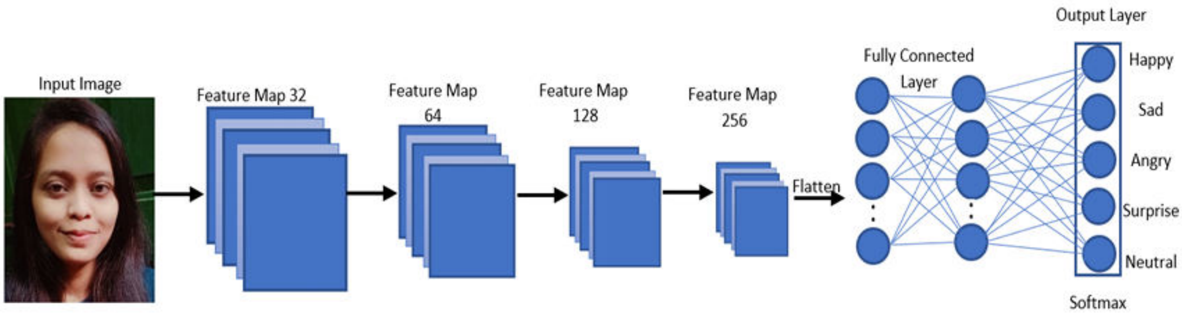


Fig 3: Convolution neural network Architecture

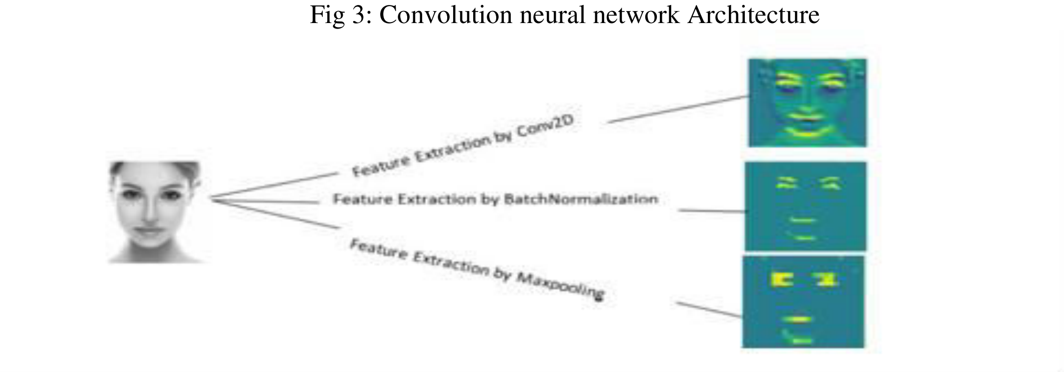


Fig 4: Feature Extraction by each layer in Convolutional Neural Network

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 3, March 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1203097 |

Fig 5: Results of Emotion Detection

B.Haar Cascade Classifier for Face Detection: The Haar Cascade classifier is employed for detecting faces in the webcam feed. It is initialized with the pre-trained frontal face detection XML file provided by OpenCV.

C.Detecting Emotions in Real-Time: Once the webcam starts capturing frames, the code processes each frame to detect faces using the Haar Cascade classifier. For each detected face, the emotion detection model predicts the predominant emotion based on the facial expression.

D.Customizing Song Recommendations: Users can select a genre and optionally specify an emotion for customizing song recommendations. The code provides preset values for audio features corresponding to different emotions. Users can also adjust the year range and slider values for acousticness, danceability, energy, instrumentalness, valence, and tempo to further customize recommendations.

E.Nearest Neighbors for Song Recommendation: The code uses a nearest neighbors algorithm to recommend songs similar to the selected genre and emotional context. It calculates the nearest neighbors based on the specified audio features and retrieves the corresponding Spotify track URIs.

F.Displaying Song Recommendations: The recommended songs are displayed as embedded Spotify track players within the Streamlit app. Additionally, a polar plot is provided for each recommended song, showing the values of the selected audio features.

G.Handling User Interactions: Users can interact with the app by selecting different genres, emotions, and adjusting slider values. They can also click on a button to request more song recommendations, which updates the displayed recommendations dynamically.[9]

IV. RESULTS & DISCUSSIONS

Here's a breakdown of the implementation steps:
Import Libraries: The first step is to import the necessary libraries. These include Streamlit for building the web application, OpenCV (cv2) for computer vision tasks, NumPy for numerical operations, and Keras for loading the pre-trained emotion detection model.
Load Pre-Trained Model: The pre-trained emotion detection model (facialemotionmodel.h5) is loaded using the load_model function from Keras. This model is used to detect emotions from facial expressions captured through the webcam.
Define Emotion Classes: A list of emotion classes is defined, representing the emotions that the model can recognize. These include 'Angry', 'Disgust', 'Fear', 'Happy', 'Sad', 'Surprise', and 'Neutral'.
Load Face Detector: The Haar Cascade classifier for face detection is loaded using cv2.CascadeClassifier. This classifier is used to detect faces in the webcam feed.
Define Emotion Detection Function: A function detect_emotion is defined to detect the emotion from a given face image. This function takes the face image, resizes it, and feeds it into the emotion detection model to predict the dominant emotion.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 3, March 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1203097 |

Main Function: The main() function is defined, which serves as the entry point of the Streamlit web application. It creates a button ('Start Camera') that initiates the webcam feed and detects emotions in real-time.
Start Camera Button: When the 'Start Camera' button is clicked, the webcam feed is initiated using cv2.VideoCapture(0). The webcam frames are continuously read, and faces are detected using the Haar Cascade classifier.

Detect Emotions: For each detected face, the detect_emotion function is called to predict the emotion. The predicted emotion is overlaid on the webcam frame using bounding boxes and text annotations.
Display Webcam Feed: The processed webcam frames with detected emotions are displayed in the Streamlit application using the placeholder.image function
Display Detected Emotion: Once an emotion is detected, it is displayed below the webcam feed using st.write.
Streamlit Interface: The Streamlit interface allows users to choose a genre and customize the emotion manually or through facial expression detection. The chosen features are used to recommend songs from Spotify based on similarity.
Recommend Songs: The system recommends songs from Spotify based on the selected genre and emotion. The recommendations are updated dynamically based on user inputs.

The results of the project will be as follows:

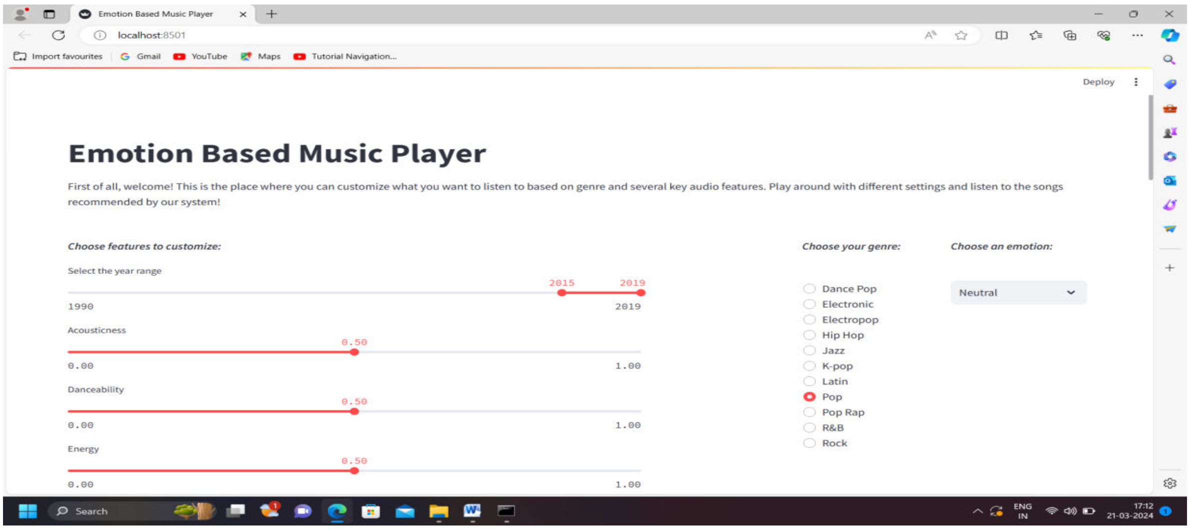


Fig 6: User Interface

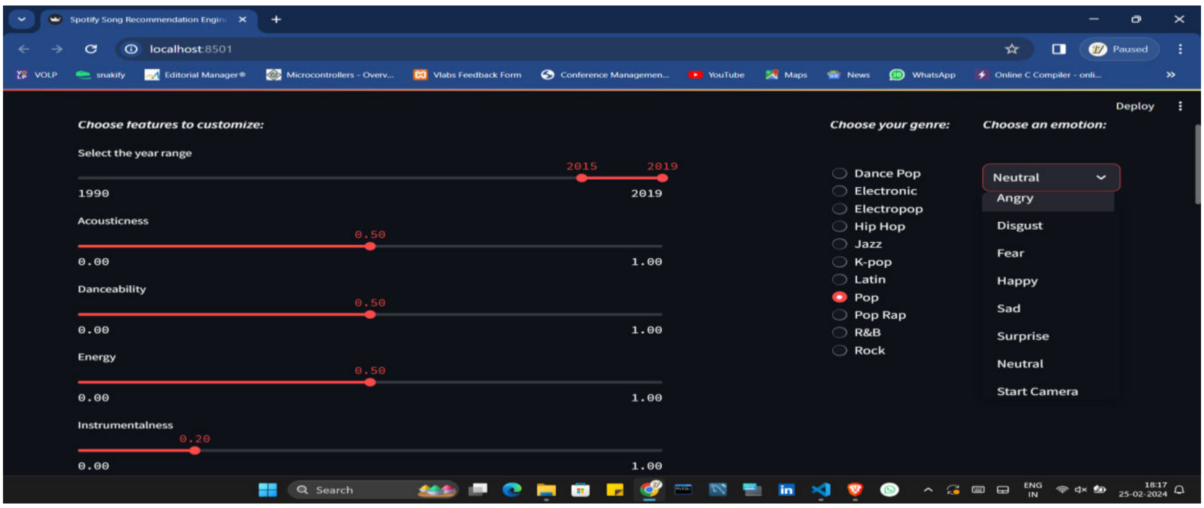


Fig 7: Emotion Capturing

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 3, March 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1203097 |

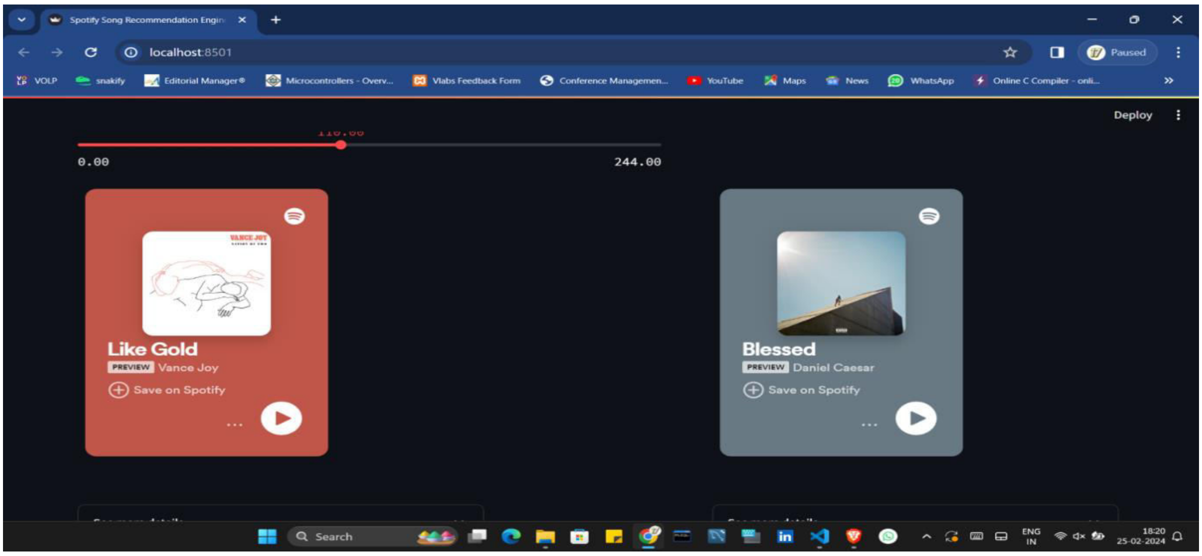


Fig 8: Playlist Suggestion based on emotion

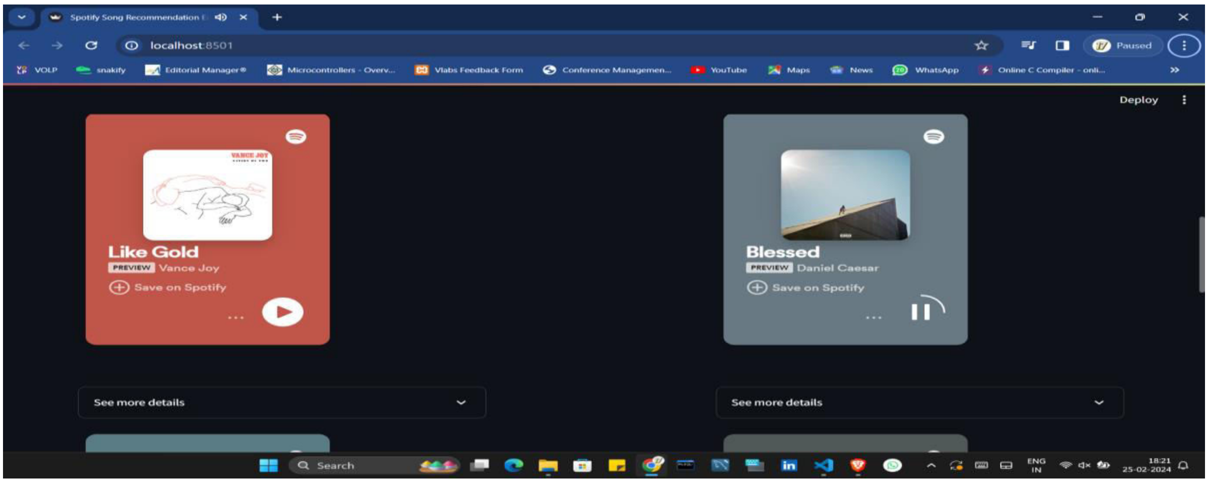


Fig 9: Playing Songs

V. CONCLUSION

In conclusion, this integrated system offers an interactive and intuitive platform for users to discover music that resonates with their emotions and preferences. It demonstrates the potential of combining computer vision techniques with machine learning algorithms to create personalized and engaging user experiences in music recommendation applications.

VI. LIMITATIONS & FUTURE WORK

A. Accuracy of Emotion Detection**:**

The accuracy of the emotion detection system heavily relies on the performance of the underlying facial expression recognition model. While the Haar Cascade classifier used for face detection is a well-established technique, the accuracy of emotion recognition may vary depending on factors such as lighting conditions, facial expressions, and individual differences in facial features. Improving the robustness and accuracy of emotion detection algorithms would enhance the reliability of the music recommendation system.

B. Limited Emotion Classes:

The project currently supports a limited set of emotion classes, including Angry, Disgust, Fear, Happy, Sad, Surprise, and Neutral. However, human emotions are complex and nuanced, and categorizing them into discrete classes may

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 3, March 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1203097 |

oversimplify the range of emotions experienced by users. Expanding the emotion classification system to include a more comprehensive set of emotions could provide more personalized and accurate music recommendations.

C. Limited Music Dataset:
The music recommendation system relies on a pre-existing dataset of music tracks for generating recommendations. The effectiveness of the recommendations may be limited by the size and diversity of the dataset, as well as the availability of metadata such as genre, tempo, and mood. Incorporating a larger and more diverse collection of music tracks, as well as leveraging user feedback to refine recommendations over time, could improve the quality and relevance of music suggestions.

The future work as follows:
A. Enhanced Emotion Detection**:**
Improve the accuracy and robustness of emotion detection by exploring advanced deep learning models or ensemble techniques. Fine-tuning the existing model on larger and more diverse datasets can help capture subtle nuances in facial expressions and emotions.

B.Interactive User Interface**:**
Design an intuitive and user-friendly interface that allows users to interactively customize and refine their music preferences based on detected emotions. Incorporate features such as drag-and-drop playlist editing, mood sliders, and visualizations to enhance user engagement and satisfaction.

C. Integration with Music Streaming Platforms:
Collaborate with popular music streaming platforms (e.g., Spotify, Apple Music) to integrate the emotion-based recommendation system as a feature or extension. This integration can enhance the reach and adoption of the system among a wider audience of music enthusiasts.

VII. ACKNOWLEGMENT

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 , SpotifyAPI and the FER 2013 dataset creators. We are deeply thankful to our Project Coordinator **Dr N. Lakshmi Prasanna** 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] Ramanathan, R., Kumaran, R., Rohan, R. R., Gupta, R., & Prabhu, V. (2017, December). An intelligent music player based on emotion recognition. In *2017 2nd International Conference on Computational Systems and Information Technology for Sustainable Solution (CSITSS)* (pp. 1-5). IEEE.
[2] Gilda, S., Zafar, H., Soni, C., & Waghurdekar, K. (2017, March). Smart music player integrating facial emotion recognition and music mood recommendation. In *2017 international conference on wireless communications, signal processing and networking (wispnet)* (pp. 154-158). IEEE.
[3] Alrihaili, A., Alsaedi, A., Albalawi, K., & Syed, L. (2019, October). Music recommender system for users based on emotion detection through facial features. In *2019 12th International Conference on Developments in eSystems Engineering (DeSE)* (pp. 1014-1019). IEEE.
[4] Li, S., & Deng, W. (2020). Deep facial expression recognition: A survey. *IEEE transactions on affective computing*, *13*(3), 1195-1215.
[5] Preema, J. S., Rajashree, S. M., Savitri, H., & Shruthi, S. J. (2018). Review on facial expression-based music player. *International Journal of Engineering Research & Technology (IJERT) ICRTT*, *6*(15).

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 | Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 3, March 2024 ||

| DOI: 10.15680/IJIRCCE.2024.1203097 |

[6] Vijaysinh, G. U., Shivaji, G. S., Ashok, M. R., & HrutvikShahaji, J. Emotion based Music Recommendation System.

[7] Qin, Z., Yu, F., Liu, C., & Chen, X. (2018). How convolutional neural network see the world-A survey of convolutional neural network visualization methods. *arXiv preprint arXiv:1804.11191*..

[8] Nordén, F., & von Reis Marlevi, F. (2019). A Comparative Analysis of Machine Learning Algorithms in Binary Facial Expression Recognition.

[9] Singhal, P., Singh, P., & Vidyarthi, A. (2020). Interpretation and localization of Thorax diseases using DCNN in Chest X-Ray. *Journal of Informatics Electrical and Electronics Engineering (JIEEE)*, *1*(1), 1-7.

[10] Vinny, M., & Singh, P. (2020). Review on the Artificial Brain Technology: BlueBrain. *Journal of Informatics Electrical and Electronics Engineering (JIEEE)*, *1*(1), 1-11.

[11] Singh, A., & Singh, P. (2021). Face mask detection using deep learning to manage pandemic guidelines. *Journal of Management and Service Science (JMSS)*, *1*(2), 1-21.

[12] A. Singh and P. Singh (2021) License Plate Recognition. Journal of Management and Service Science, 1(2), 1, pp. 1-14.

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

[14] 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.

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING