Indian Institute of Technology, Jodhpur

॥ त्वं ज्ञानमयो विज्ञानमयोऽसि ॥

Project Report

Artificial Intelligence (CS7610)

Project: Music recommendation system based on facial emotion recognition.

Submitted By:

Bikash Dutta (D22CS051)

Prashant Gautam (M22CS057)

Deepak Kumar (M22MA004)

1. Abstract

We offer a novel method for automatically playing music based on facial emotion. The majority of known options involve manually playing music, using wearable computing devices, or classifying based on auditory attributes. Instead, we propose that the manual sorting and playing be changed. For emotion identification, we used a Convolutional Neural Network. An integrated camera records facial expressions. On input facial photos, feature extraction is used to determine emotions such as happy, sad, angry, and neutral. The user's current emotion is used to produce an automatic music playlist.

Artificial intelligence, an extensive, prominent and imperative domain that has attracted a lot of researchers and programs in recent times. This particular domain has taken over the world in very short notice. It is incorporated in daily life in the form of chatbots, digital assistants like Siriand several other technology-based systems. One of the most prominent powers up of artificial intelligence is face recognition techniques. The basic example of its usage is the grouping of Google Photos of a particular person.There are many existing systems that could recognize facial emotions. On the other hand, there are systems that recommend music. Bringing together a system which will recommend music by recognizing the mood of the user from facial emotions is the overall concept described in the paper. Emotion recognition would have larger scope in the near future in fields like robotics for efficient sentimental analysis without the involvement of another human.

Many of the studies in recent years admit that humans reply and react to music and this music has a high impression on the activity of the human brain. In one examination of the explanations why people hear music, researchers discovered that mu-sic played a crucial role in relating arousal and mood. Two of the most important functions of music are its ability is partici-pants rated to help them achieve a good mood and become more self-aware.

2. Objective

To build a song recommendation application which recommends songs based on the mood of the user. Also to provide the users with a better way to enjoy music. Music is one of the best ways to relax and plays a vital role in everyone's life. This project will give a new and better way to explore music.

3. Literature Review

The review is done to get insights into the methods, their shortcomings which we can overcome. A literature review, a literature survey is a text of a scholarly paper, which includes the current understanding along with great findings, as well as theo-retical and methodological contributions to a particular topic. The latent qualities of humans that can provide inputs to any system in various ways have brought the attention of several learners, scientists, engineers, etc. from all over the world. The current mental state of the person is provided by facial expressions. Most of the time we use nonverbal clues like hand gestures, facial expressions, and tone of voice to express feelings in interpersonal communication.

Yusuf Yaslan et al. proposed an emotion-based music recommendation system that learns the user's emotion from signals obtained through wearable computing devices that are integrated with galvanic skin response (GSR) and photoplethysmog-raphy (PPG) physiological sensors in their paper.

Ayush Guidel et al stated that human being’s state of mind and current emotional mood can be easily observed through their facial expressions.The paper proposed by Ramya Ramanathan conveyed the intelligent music player using emotion recognition.The paper specifically makes a M. Athavle et al. ISSN (Online) : 2582-7006 International Conference on Artificial Intelligence (ICAI-2021)

Journal of Informatics Electrical and Electronics Engineering (JIEEE) A2Z Journals specialty of the methodologies available for detecting human emotions for developing emotion-based music players, the approach a music player follows to detect human emotions, and the way it is ideal to apply the proposed system for emotion detection. It additionally offers a brief idea about our systems working, playlist generation, and emotion classification.CH Radhika et al [8] advised manual segregation of a playlist and annotation of songs, following the current emotional state of a user, as a labour-intensive and time-consuming task.The paper presents an algorithm that automatically does the process of generating a playlist of audio, based on the facial expressions of a person, for rendering salvage of time as well as labour, invested in performing this process manually.

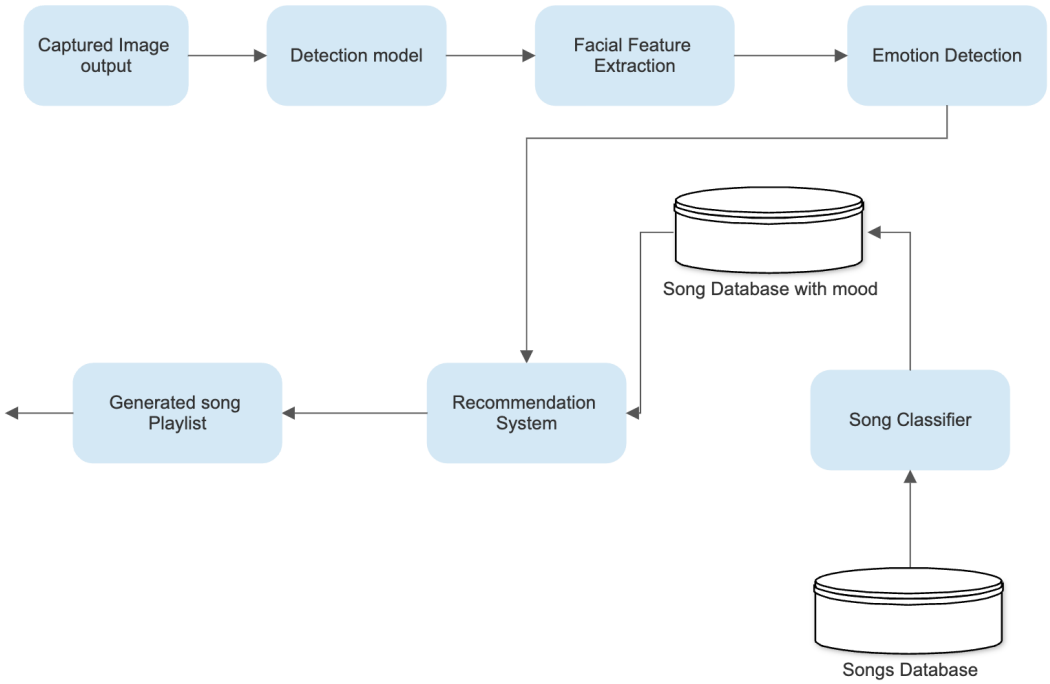
4. Motivation

As a passionate music enthusiast we often face issues of random music recommendation over several platforms. Apps like Gaana, Spotify or YouTube Music also struggle to keep up with the mood of the listener when it comes to recommendation based on the song played currently. This generally leads to frustration amongst the user's. The reason for the existing algorithms being less accurate is that they are missing a very key feature which could change the game of recommendation that is using facial expression as a heuristic. This led us to the idea of working on this project.

The purpose of this project was to explore facial expression recognition for implementation of an emotion-based music player. The manual analysis of faces by people was completely replaced by reasonable computer programming. Apart from providing theoretical background, this study provides approaches to outline and execute emotion-based music players with a wide variety of image processing techniques.In the proposed system, facial images are processed and basic emotions are recognized, and then music is played based on the user's emotions, and also suggested music that enhances mood. We would like to improve our system's ability to recognize emotions in the future and also recognize more different emotions.

5. Methodology

The proposed system benefits us to present interaction between the user and the music player. The system's goal is to adequately capture the face using the camera. Convolutional Neural Network, which forecasts emotion, is fed captured images. The playlist of songs is then generated using the emotion gleaned from the taken image. Music recommendation based on facial mood contains five modules. Image capture, Emotion recognition, Song classification, Music recommendation.



Block diagram of the system

5.1. Emotion Detection Module

A system known as facial emotion recognition can analyse a person's feelings from several sources, including images and videos. It is a member of the group of technologies known as "affective computing," which is a multidisciplinary field of study on how well computers can recognise and understand human emotions and affective states. Affective computing frequently builds on AI technologies. Human emotions can be inferred from facial expressions, which are a form of non-verbal communication. Face detection, facial expression recognition, and classification of facial expressions to an emotional state make up the three processes of emotion detection. Emotion detection is based on the analysis of facial landmark positions (e.g. end of nose, eyebrows). Depending on the algorithm, facial expressions can be classified to basic emotions (e.g. anger, disgust, happiness, sadness, and surprise). The algorithms for face recognition are usually model based.

Regular convolutional neural networks are used in face recognition models, and they are in charge of representing faces as vectors. A face pair of the same person should be more similar than a face pair of two distinct people, according to our expectations. Similarity could be calculated by different metrics such as Cosine Similarity, Euclidean Distance.
We've used several models like SVM, CNN and ended up with DeepFace facial attribute detection system. Deepface is a lightweight face recognition and facial attribute analysis framework for python. Deepface also comes with a strong facial attribute analysis module including age, gender, race and facial expression predictions. It gives best accuracy among other models and is fairly easy to deploy in our application. Deepface serves an API. The model runs on an external server. An image is sent by the application to that server on which the model analyses the image and returns back the prediction to the application.

5.2. Song Classification

It's crucial to categorise songs according to various moods so that the recommender system can recognise each song and deliver the desired result. The majority of how songs are organised today is based on an artist's overall genre rather than the emotion a song evokes. Although it is difficult, trying to classify music using engineering methods may be able to reduce the differences in listeners' classification preferences. Automatically determining a piece's mood would be incredibly helpful for organising huge digital music libraries like those on Apple Music or Spotify. The classification is done through breaking down each of the songs into quantifiable acoustic components such as tempo, rhythm, harmony, valence, tone, resonance etc.
Faster tempos are associated with high-energy songs, "happy mood", and slower tempos with lower energy, "sadder songs". Loudness, or intensity of a song can be connected with "anger", while softer songs would suggest tenderness, sadness, or fear. Higher overall pitch can be an indicator of happiness, carefree and light moods within a song, while lower pitch implies a sad and serious tone.
We've performed mood classification using a decision tree on Kaggle Spotify dataset which contains song information and its attributes. It has around 15000 songs and its features like tempo, key, valence, liveness, danceability, acoustic-ness etc.

5.3. Music Recommendation System

A recommender system, often known as a recommendation system, is a type of information filtering system that aims to anticipate the "rating" or "preferred" a user would assign to a certain item. They are mainly applied in commercial settings. More formally, the recommender problem can be interpreted as determining the mapping (c, i) → R where c denotes a user, i denotes an item, and R is the utility of the user being recommended with the item. The top N items are then recommended to the user once the items have been sorted by utility. The recommendation system searches the songs in the dataset and groups the songs according to their mood making playlists. The newer songs that are added to the database go through song classifiers and then the recommender system adds them to already existing playlists by calculating the cosine similarity.

5.4. Mobile Application Overview

We benefit from the proposed system by presenting interaction between the user and the music player. The system's goal is to adequately capture the face using the camera. Images are captured and sent into a Convolutional Neural Network, which predicts emotion. The emotion produced from the taken image is then used to generate a song playlist. The result is the recommended playlist for the user in the UI of the music player by showing captions according to detected emotions. We have used Flutter to create a mobile application that integrates the emotion detection and song recommendation system. The app opens a front camera module on startup to capture a photo of the user. This image is then sent to a server running the Emotion detection model. The server sends back the detected feature as a string. Recommendation system receives the string and then makes a playlist the user might like. The music player module plays the playlist. Variables like playlist, songstatus, and root are used for storing the name of all songs, storing the status of currently active songs, and for the main GUI window respectively.

6. Conclusion

As the strength and benefits of AI-powered applications gain popularity, our project will be a cutting-edge trending technological application. We present an explanation of how music can affect the user's mood and how to choose the correct music tracks to improve the user's moods in this system. The system in place can identify the user's emotions. The system can detect happy, sad, angry, neutral, or shocked emotions. Following the determination of the user's emotion, the suggested system presented the user with a playlist including music matches that detected the mood. Processing a large dataset is both memory and CPU expensive. This will make development more difficult and appealing. The motive is to create this application in the cheapest possible way and also to create it under a standardised device. Our music recommendation system based on facial emotion recognition will reduce the efforts of users in creating and managing playlists.

6.2 Limitations

The existing algorithm shall work perfectly if the listener is ecstatic or joyous but 6t would rather recommend sad songs if the person is a bit upset. It could be better if the algorithm would work in a way that changes the dull mindset into an excited mindset.
It's challenging to fetch high resolution features with consistency at all the times which could mislead the recommendation system due to less stable internet connection and low quality webcams. People are generally neutral when it comes to expressing thus the facial expressions can't always be relied upon for the prediction.

6.3 Future Scope/Future Enhancements

This system, although completely functioning, does have scope for improvement in the future. There are various aspects of the application that can be modified to produce better results and a smoother overall experience for the user. Some of these are an alternative method, based on additional emotions which are excluded in our system as disgust and fear. This emotion included supporting the playing of music automatically. The future scope within the system would style a mechanism that might be helpful in music therapy treatment and help the music therapist to treat the patients suffering from mental stress, anxiety, acute depression, and trauma. The current system does not perform well in extremely bad light conditions and poor camera resolution thereby provides an opportunity to add some functionality as a solution in the future.

7. References

- Bhat, A. S., Amith, V. S., Prasad, N. S., & Mohan, M. (2014). An Efficient Classification Algorithm For Music Mood Detection In Western and Hindi Music Using Audio Feature Extraction. 2014 Fifth International Conference on Signal and Image Processing, 359-364. https://doi.org/10.1109/ICSIP.2014.63
- An Accurate Algorithm for Generating a Music playlist based on Facial Expressions. International Journal of Computer Applications, 100(9). https://pdfs.semanticscholar.org/312b/2566e315dd6e65bd42cfcbe4d919159de8a1.pdf
- Ahlam Alrihail, Alaa Alsaedi, Kholood Albalawi, Liyakathunisa Syed, Music recommender system for users based on emotion detection through facial features, Department of Computer Science Taibah University, (DeSE), 2019. https://doi.org/10.1109/DeSE.2019.00188
- CH. sadhvika, Gutta.Abhigna, P. Srinivas Reddy, Emotion-based music recommendation system, Sreenidhi Institute of Science and Technology, Yamnampet, Hyderabad; International Journal of Emerging Technologies and Innovative Research (JETIR) Volume 7, Issue 4, April 2020.
- Other resources
    - https://volpato.dev/posts/simple-flutter-music-player-app/
    - https://ytmusicapi.readthedocs.io/en/stable/reference.html
    - https://www.sciencedirect.com/science/article/pii/S1877050919310646#:~:text=By%20using%20music%20recommender%20system,that%20has%20been%20heard%20previously.
    - https://www.section.io/engineering-education/building-spotify-recommendation-engine/#types-of-recommendation-systems

8. Roles of the team members

Bikash Dutta: Mobile application making. Both front and back end including model integration.
Prashant Gautam: Model selection, Songs classification, Report making.
Deepak Kumar: Songs classification, Recommendation systems, Report making.