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

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 4, April 2023

DOI: 10.17148/IJARCCE.2023.124144

MUSIC RECOMMENDATION BASED ON FACIAL EMOTION RECOGNITION

Mr Chakrapani D S[1], Sidrath Iram[2] Suchitra R Bhat Agni[3], Supritha L[4], Leelavathi S[5]

Assistant Professor, Department of Computer Science, JNNCE, Shivamogga, India[1]

Students, Department of Computer Science, JNNCE, Shivamogga, India[2-5]

Abstract: The development of a Music Recommendation System involved the utilization of the FER-2013 and Age, Gender (Facial Data) datasets. The system utilizes the CNN architecture, commonly employed for such purposes, to train three separate models: Emotion, Gender, and Age. To enhance the models' performance, additional layers are incorporated into the training phase. These models are subsequently employed as classifiers. To predict the user's mood, age, and gender, a snapshot of the user captured through the camera is forwarded to the trained models. Based on the outcomes of these classifiers, various playlists sourced from a database are suggested to the user. The goal is to create a functional and user-friendly environment for music selection. Once the playlists are proposed, the user can select their desired playlist and begin listening to the recommended music.

Keywords: Deep Learning, CNN, Emotion, Age, Gender, Music Recommendation System.

I. INTRODUCTION

The modern landscape of multimedia technology has given rise to a wide array of music players, boasting an assortment of capabilities. These include fast forward, variable playback speed, reverse playback, local and streaming options, volume adjustment, and genre categorization, all aimed at providing a customizable and immersive music listening experience. These developments cater to the diverse needs and preferences of music enthusiasts and listeners. Moreover, deep learning-based facial expression identification has paved the way for accurately identifying human emotion states, including anger, fear, neutrality, happiness, disgust, sadness, and surprise. This technique relies on automatically detecting facial expressions by capturing face that have been labelled from a facial expression collection. Through the implementation of Convolutional Neural Network (CNN) models, the system trains itself to determine the specific facial expression being captured. The proposed system consists of three essential modules: The Recommendation Module, the Music Classification Module, and the Emotion Module. These modules work in synergy to provide a comprehensive music recommendation experience, taking into account the user's emotional state and facial expressions detected by the CNN model. This integration aims to enhance the music selection process and cater to the individual's emotional needs based on their age, gender and preferences.

The music player's dataset was created using user defined song dataset. To offer personalized and suitable music recommendations, a Convolutional Neural Network (CNN) was used to detect facial emotions. This approach achieved an impressive accuracy of about 95.14% [1]. By accurately identifying facial emotions, the music player was able to provide users with music recommendations that matched their emotional state.

II. LITERATURE SURVEY

Various approaches have been devised for recognizing human emotions [2], [4]. For instance, one approach involved extracting features from electroencephalograph (EEG) signals to identify emotions during music listening [2]. In another study [4], emotions of users utilizing a music recommendation system were classified using wearable devices equipped with galvanic skin response and photo plethysmography physiological sensors. However, these methods may not be practical for individuals without access to such specialized devices.

A more practical approach to determine human emotion classes involves analyzing speech and video data captured by microphones and web cameras [6]. In the field of music recommendation research, webcam images are commonly used to detect facial expressions and analyses facial emotion expression [3], [5], [7]. These studies leverage the information captured from webcam images to understand the user's emotional state, which plays a crucial role in recommending suitable music tracks. By employing facial emotion analysis techniques, researchers aim to provide more personalized and relevant music recommendations to users. Renuka R Londhe et al. [8] proposed a paper which focused on the study

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

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 4, April 2023

DOI: 10.17148/IJARCCE.2023.124144

of changes in the curvatures of the face and the intensities of the corresponding pixels. The author used Artificial Neural Networks(ANN), which was used to classify the emotions and proposed with diverse techniques for a playlist.

It is important to note that recommending the same playlist solely based on emotions may not be ideal, as the preferences and requirements vary between individuals, such as children and adults with similar emotions. Furthermore, recommendations might differ based on cultural backgrounds or nationalities. Hence, it is crucial to consider additional factors beyond emotions in order to provide more personalized and suitable music recommendations.

III. PROBLEM DEFINITION

Our goal is to create a music player system that utilizes machine learning algorithms to recommend music based on the user's real-time mood captured through a web camera. The system aims to analyse the user's facial expressions and determine their emotional state, which will then be used to suggest music that aligns with their mood. By leveraging machine learning techniques, aims to provide a seamless and enjoyable music listening experience tailored to the user's current emotional state.

IV. PROPOSED SYSTEM OVERVIEW

The method is to employ Deep Neural Networks (DNN) and learn the best suitable element deliberations straightforwardly from the details taken in and handle the constraints of handmade highlights. CNNs are exceptionally powerful in lessening the mass of boundaries without overlooking on the nature of models. The proposed framework can concede the looks of the client and in view of individual's looks utilizing CNN model. When the inclination has been set apart the tune matching the client's feelings would be played. The camera captures the input image using a webcam instantaneously. Using the webcam the system shall get the image feed so it has to convert the captured video to image so that it shall be passed to the model, so it needs to capture a single frame from the video that would be image frame which shall be processed for prediction. Then comes role of feature mapping and based on the received input the property of the input image frame will be captured. The image frame now goes to the backend where there shall be machine learning model and based on the features of image frames taken, The output is received as an emotion that is acknowledged by the machine learning model. Based on emotion interpreted a song related to that emotion is played for the user in the frontend.

currently the Convolutional Neural Networks models are pretty much accurate in associating the objects and have come very adjacent to the human performance. A CNN model which contains varied CNN layers stack is built.

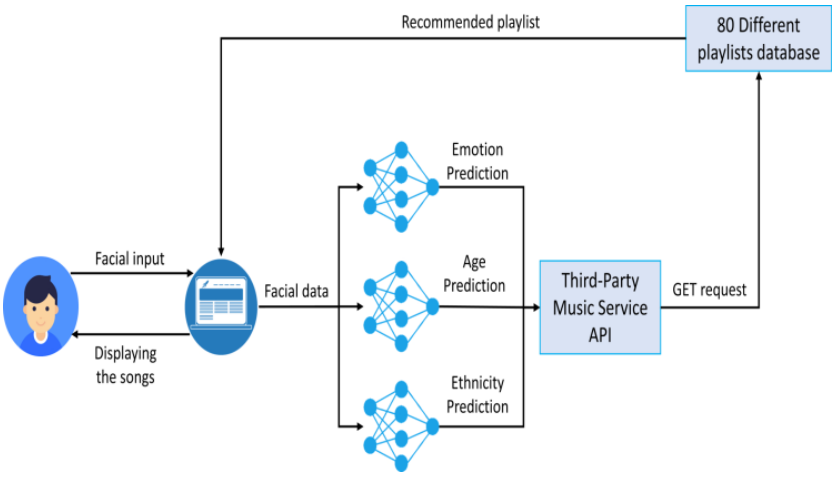

Fig1: Block diagram of the proposed system

,

The project consists of two major parts that work together to enhance the user's experience. In the first part, we implement three deep learning methods based on Convolutional Neural Networks (CNNs) to analyze the image captured from the user's camera. These methods focus on detecting the user's face and extracting valuable information such as their emotion, age, and gender. This analysis is performed simultaneously and in parallel.

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

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 4, April 2023

DOI: 10.17148/IJARCCE.2023.124144

Moving on to the second part, once the user's emotion, age, and gender are determined, the system utilizes this information to provide a personalized playlist suggestion. By considering the user's unique characteristics, we can tailor the playlist to their mood, preferences, and demographic profile. This ensures that the recommended playlist aligns with their current emotional state, age group, and gender.

To give you an overview of the system's operation, Figure 1 illustrates the Block diagram of the proposed System flow. The user starts by accessing the application and activating their camera to take an instant snapshot. The image obtained from the camera is then processed using the three CNN models running in parallel. These models work simultaneously to detect the user's emotion, age, and gender within the captured frame. Based on these findings, an appropriate playlist is selected and suggested to the user, offering a personalized music experience.

V. PROPOSED CNN ARCHITECTURE

CNNs are a popular deep learning approach utilized in a wide range of applications. Typically, CNN architectures consist of various combinations of five fundamental layer types: convolutional (Conv), max-pooling (MP), batch normalization (BN), dropout (DR), and fully connected (FC) layers.

In the Conv layer, diverse filters are applied to the image using convolutional operations to extract essential features. This process aims to capture meaningful patterns present in the image. The MP layer reduces the image size in a specific manner, thereby reducing computational load and filtering out unnecessary information. Excessive features can lead to overfitting, which can be mitigated by incorporating DR layers into CNN architectures.

The FC layer, serving as the final layer, performs the classification task by transforming the image into a vector representation and adjusting its weights accordingly. This vector is then fed as input to the FC layer for the classification process.

The proposed CNN architecture incorporates all these layer types to effectively extract and learn features from the input data.

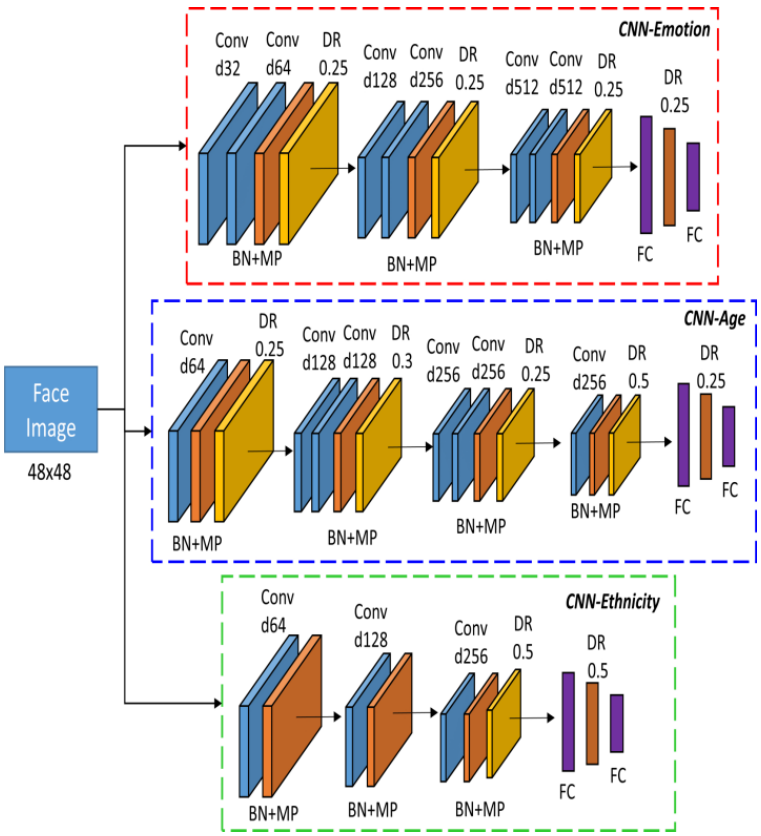

Fig 2.1: CNN architecture

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

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 4, April 2023

DOI: 10.17148/IJARCCE.2023.124144

VI. RESULTS AND ANALYSIS

In this project, an automated music playlist recommendation system has been developed that utilizes images to determine the user's emotion, ethnicity, and age. Initially, transfer learning methods were considered, but a CNN-based architecture was deemed more suitable based on performance analysis. Trained models were saved and integrated into the system, and act as classifiers that use a snapshot of the user. Based on the output of the three models, a playlist is proposed to create a pleasing user experience. The result of the proposed system is shown in the below figure.

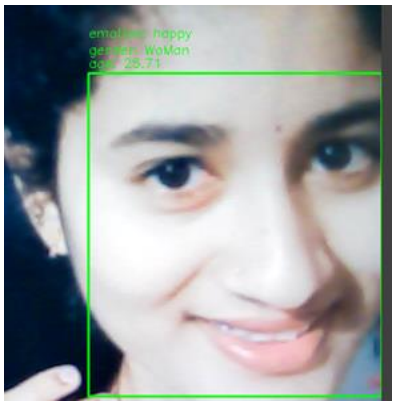

Fig :6.1 Results for happy

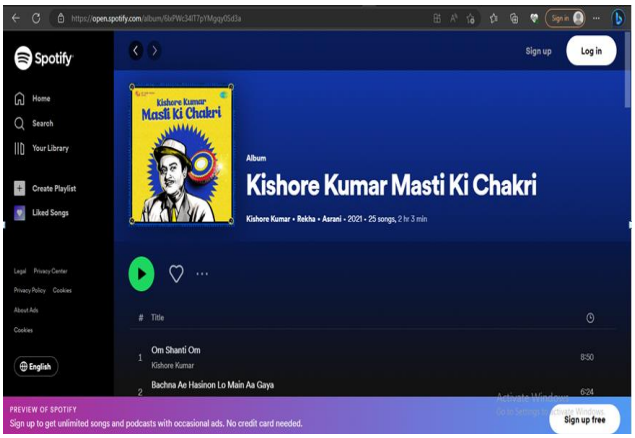

Fig:6.2 Playlist for Happy mood

VII. CONCLUSION

The project focuses on leveraging trending technology to develop a state-of-the-art music recommendation system. The aim is to explore the impact of music on the user's mood and select suitable tracks to improve their emotional state. The system can detect emotions such as happiness, sadness, anger, neutrality, and surprise. Based on the user's emotion, the system provides a curated playlist of music that matches the detected mood. This approach adds a challenging and appealing aspect to the development process. The objective is to create an affordable and standardized application that utilizes facial emotion recognition for music recommendations. By incorporating this technology, the system simplifies the process of playlist creation and management for users.

VIII. FUTURE SCOPE

While the current system is operational, there is room for future enhancements and improvements to enhance the overall user experience. One area of future scope involves developing a mechanism within the system that can aid music therapy treatments. This would enable music therapists to effectively treat patients who are dealing with mental stress, anxiety, acute depression, and trauma. By leveraging the system's capabilities, therapists can utilize music recommendations to create personalized playlists that cater to the specific needs and emotions of their patients, enhancing the therapeutic process.

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 4, April 2023
DOI: 10.17148/IJARCCE.2023.124144

Additionally, there is room for improvement in the system's performance under challenging conditions such as extremely bad lighting or poor camera resolution. Future development can focus on implementing solutions to overcome these limitations. This could involve incorporating advanced algorithms or image processing techniques to enhance the system's ability to detect and analyses facial expressions accurately, even in adverse conditions. By addressing these technical challenges, the system can provide a more robust and reliable music recommendation experience for users, regardless of the environmental conditions.

Future possibilities for the music recommendation system include integration with wearable devices for real-time recommendations, improved accuracy of facial emotion and gender recognition algorithms, integration with social media platforms for personalized recommendations, personalized playlists based on user characteristics, and expansion to other domains like movies or books.

REFERENCES

[1] Shlok Gilda, Husain Zafar, Chintan Soni, Kshitija Waghurdekar, Smart music player integrating facial emotion recognition and music mood recommendation, Department of Computer Engineering, Pune Institute of Computer Technology, Pune, India, (IEEE),2017
[2] Lin, Y. P., Wang, C. H., Jung, T. P., Wu, T. L., Jeng, S. K, Duan, J. R., Chen,J. H.: EEG-Based Emotion Recognition in Music Listening. IEEE Transactions on Biomedical Engineering, 57, 7, 1798-1806 (2010)
[3] Athavle, M., Mudale, D., Shrivastav, U., Gupta, M.: Music Recommendation Based on Face Emotion Recognition. Journal of Informatics Electrical and Electronics Engineering. 2, 2, 1-11 (2021)
[4] Ayata, D., Yaslan, Y., Kamasak, M. E.: Emotion Based Music Recommendation System Using Wearable Physiological Sensors. IEEE Transactions on Consumer Electronics. 64, 2, 196-203 (2018)
[5] Shantashalini, K., Jaichandran, R., Leelavathy, S., Raviraghul, R., Ranjitha, J., Sar- avanakumar, N.: Emotion Based Music Recommendation System Using Wearable Physiological Sensors. Turkish Journal of Computer and Mathematics Education, 12, 1, 912-917 (2021)
[6] Singh, L., Singh, S., Aggarwal, N., Singh, R., Singla, G.: An Efficient Temporal Feature Aggregation of Audio-Video Signals for Human Emotion Recognition. 6th IEEE International Conference on Signal Processing, Computing and Control. India (2021)
[7] Mahadik, A., Milgir, S., Patel, J., Jagan, V. B., Kavathekar, V.: Mood based music recommendation system. International Journal of Engineering Research and Technology. 10 (2021)
[8] Londhe RR and Pawar DV 2012 Analysis of facial expression and recognition based on statistical approach International Journal of Soft Computing and Engineering