

Facial Emotion based Song Recommendation

Armaan Khan

Department of Computer Science and Engineering, MIT
School of engineering, MIT-ADT University, Pune,
412201, India

Ankit Kumar

Department of Computer Science and Engineering, MIT
School of engineering, MIT-ADT University, Pune,
412201, India

Abhishek Jagtap

Department of Computer Science and Engineering, MIT
School of engineering, MIT-ADT University, Pune,
412201, India

Dr. Mohd.Shafi Pathan

Department of Computer Science and Engineering, MIT
School of engineering, MIT-ADT University, Pune,
412201, India

Abstract— Larger part of individuals of any age likes music, and we accept that music players should be able to do significantly more than just playing melodies and permitting clients to make playlists. Therefore, paying attention to music (premium promotion free or with ads) to loosen up after work can work on one's wellbeing. For this, a music player has to be smart and answer the decisions of the client. A music player must help clients in consequently sorting out and playing tunes without requiring a lot of exertion in melody determination and revamping. An Emotion-Based Music Player gives all music audience members a superior stage and guarantees melody determination robotization. It is frequently puzzling for an individual to choose which music he/she need to tune in from a huge variety of existing choices. There have been a few idea structures accessible for issues like music, eating, and shopping relying on the state of mind of client. The primary target of our music proposal framework is to give ideas to the clients that fit the client's inclinations. The examination of the look/client feeling might prompt figure out the ongoing profound or mental condition of the client. Music and recordings are one of the areas where there is a huge opportunity to recommend overflowing decisions to clients considering their inclinations and furthermore recorded data. The fact that humans utilize makes it prominent looks to communicate all the more obviously what they need to say and the setting in which they implied their words

Keywords—: Convolutional Neural Networks; Deep Learning; Face Recognition, Song Recommendation

I. INTRODUCTION

Individuals will quite often communicate their feelings, mostly by their expressions. Music has forever been known to adjust the state of mind of a person. The project plans to catch the feeling communicated by an individual through expressions. A music player is intended to catch human feeling through the web camera interface accessible on figuring frameworks. The product catches the picture of the client and afterward with the assistance of picture division and picture handling procedures extricates highlights from the essence of an objective individual and attempts to distinguish the feeling that the individual is attempting to communicate. The undertaking means to ease up the mind-set of the client, by playing tunes that match the necessities of the client by catching the picture of the client. Since old-fashioned times the best type of appearance examination known to humanity is look acknowledgment. The most ideal way in which individuals will more often than not break down or close the inclination or the inclination or the

observations that another individual is attempting to communicate is by look. Sometimes, mind-set change may likewise help in beating circumstances like unhappiness and sadness. With the guide of speech investigation, numerous comfort dangers can be kept away from, and furthermore there can be steps taken that help carries the mind-set of a client to a superior stage.

II. LITERATURE SURVEY

S Metilda Florence and M Uma [1] (2020) proposed a paper "Facial Detection and Music Recommendation System in view of User Facial Expression" where the proposed framework can identify the looks of the client and in view of his/her looks extricate the facial milestones, which would then be characterized to get a specific feeling of the client. When the inclination has been ordered the melodies matching the client's feelings would be displayed to the client. It could help a client to pursue a choice in regards to which music one ought to pay attention to assisting the client with decreasing his/her pressure levels. The client wouldn't need to burn through any time in looking or to gaze upward for tunes. In spite of the fact that it had a few restrictions like the proposed framework couldn't record every one of the feelings accurately because of the less accessibility of the pictures in the picture dataset being utilized. The picture that is taken care of into the classifier ought to be taken in a sufficiently bright air for the classifier to give precise outcomes. The nature of the picture ought to be basically higher than 320p for the classifier to anticipate the feeling of the client precisely. High quality includes frequently need enough generalizability in nature settings.

H. Immanuel James, J. James Anto Arnold, J. Maria Masilla Ruban, M. Tamilarasan [2] (2019) proposed "Feeling Based Music Recommendation" which targets filtering and deciphering the facial feelings and making a playlist likewise. The drawn-out assignment of physically Segregating or gathering melodies into various records is decreased by creating a fitting playlist in view of a person's close to home highlights. The proposed framework centres around distinguishing human feelings for creating feeling-based music players. Direct classifier is utilized for face location. A facial milestone guide of a given face picture is made in light of the pixel's force values recorded of each point utilizing relapse trees prepared with an inclination supporting calculation. A multiclass SVM Classifier is utilized to group feelings Emotions are arranged as Happy, Angry, Sad or

Surprise. The restrictions are simply the proposed framework is as yet not ready to record every one of the feelings accurately because of the less accessibility of the pictures in the picture dataset being utilized. Different feelings are not found.

Arto L. and Jukka H. et al [3] created a system which recommends songs by having user interact with a set of images and based on user interaction with these images a song is recommended. This system uses textual meta data for describing genre of song

Bruce F. and Markus S. et al [4] proposed that we can use personality traits and emotional states to improve the music recommendation. They believe that by including these psychological factors we can improve the accuracy of recommendation system. Vinay. P et al. [5] proposed system using the extraction of appearance and facial geometric features from image of a face. They used svm and got an accuracy of around 90% on real-time images and around 98% for test images.

Deny John Samuel. et al. [6], proposed a system which uses svm and OpenCV. OpenCV was used to extract features from images and svm to predict the emotion from those extracted features. they were able to recognize 4 emotions Happy, Anger, Surprise, and Neutral. A. S. Bhat et al. [7] proposed a system which classifies mood and the tone of the music.it used Thayer's model to classify the songs.it classified the songs with an accuracy of 94.4%. In [8] Zheng modified the OpenCV based AdaBoost algorithm. They used timer and dual thread methods for face detection. And concluded that the dual thread method was fast and efficient

III. PROPOSED ARCHITECTURE

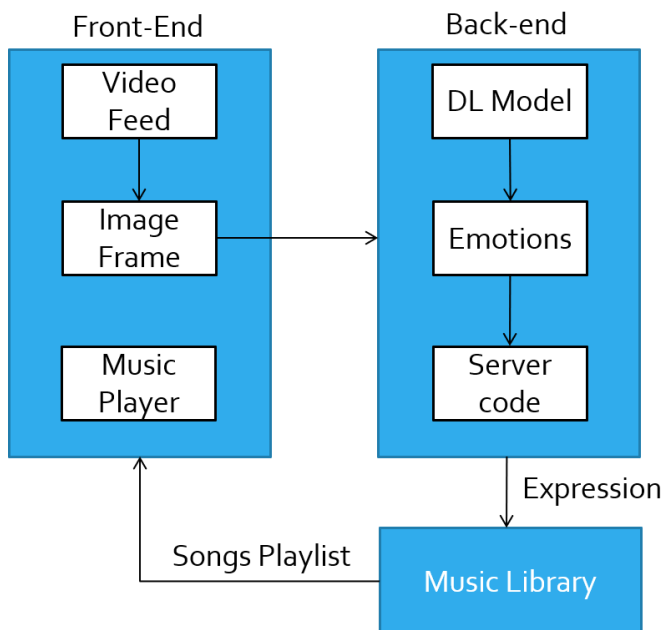


Fig. 1. Proposed Architecture

Our methodology is to utilize Deep Neural Networks (DNN) to learn the most suitable element deliberations straightforwardly from the information taken in an uncontrolled climate and handle the constraints of handmade

highlights. DNNs have been a new fruitful methodology in visual item acknowledgment, human posture assessment, face check and some more. CNNs are extremely powerful in lessening the quantity of boundaries without losing on the nature of models. The proposed framework can recognize the looks of the client and in view of person's looks utilizing CNN model. When the inclination has been characterized the tune matching the client's feelings would be played.

We would be capturing the input image using a webcam at real time. Using the webcam we will be getting the video feed so we have to convert the captured video to image so that we can pass it to the model, so we need to capture a single frame from the video that would be image frame which will be processed for prediction. Then the feature mapping will come into picture and based that the features of the input image frame will be captured. The image frame goes to the backend where we have our machine learning model and based on the features of image frame, we get out output as an emotion that is recognised by the machine learning model. Based on the emotion a song related to that emotion will be played for the user in the frontend.

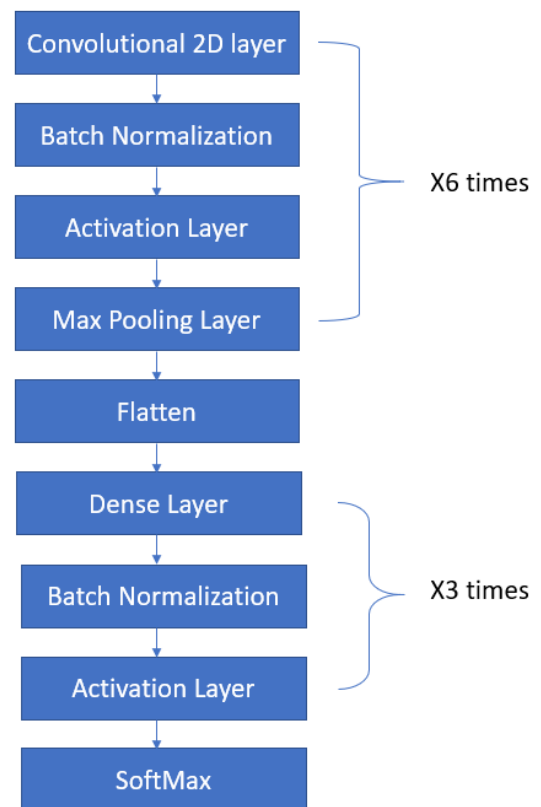


Fig. 2. Model Architecture

Nowadays the Convolutional Neural Networks models are almost accurate in identifying the objects and have come very close to the human performance. We have built a CNN model which contains various CNN layers stack as given in below architecture

Total Number of Parameters for the model are 2,137,991. There are total 6 Convolutional layers and 4 dense layers. We are using Adam as the optimizer and the values of its parameters are as follows:

Learning rate – 0.01

Beta_1 – 0.9

Beat_2 – 0.999

Epsilon – $1e^{-7}$

IV. FEATURES OF THE PROJECT

- It will be a website that will Capture the User's facial expressions.
- It will use that facial expression for determining the user's current mood.
- After determining those expressions it will recommend songs based on those expressions
- Website will contain a functioning Song Player so that users can play songs on the website

V. RESULT

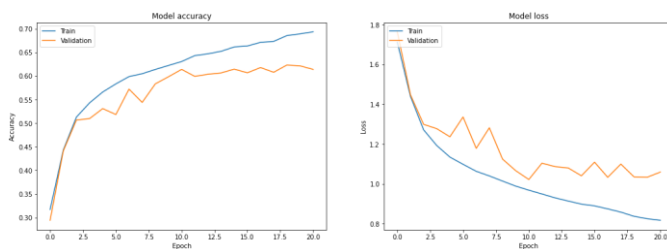


Fig. 3. Accuracy and loss over number of epochs

The final test accuracy that we are getting using our model is 62.22%.

The proposed system detects emotion in 7 different categories those are happy, angry, disgust, sad, neutral, fear and disgust respectively.

VI. FUTURE WORK

There are many research going on face recognition recently as the technology is growing with this the computational power is also increasing which have made it easier to perform high computation computer vision tasks.

As for future work, the accuracy can be increased as we train the model on more data and better algorithms. It can be used for face detection also which can be used for invigilation and monitoring purposes and further a mobile can be made which will recommend songs based on emotions on the mobile only.

VII. CONCLUSION

Focusing on different elements, like specific setting, individual boundaries, sentiments and feelings, is exceptionally vital to a dynamic course of proposals. Contemporary music proposal frameworks face the hole in Change, human sentiments, relevant inclinations and profound elements while recommending music. In this paper, we

proposed emotion-driven recommendation system with respect to personalized preferences and particular life and activity contexts. The approach presented in this study is targeted to provide maximum benefits for people from the music listening experience. It is important to make the system aware of how it is doing the recommendations, to continuously improve the music selection. By feeding the data from various sources, the system is aimed to listen to each particular user and understand their purposes of listening, feelings and contextual preferences to select the best-suited music pieces for them. We observed what kind of data is needed for the recommendation system and how it can be fetched. Main data processing tools are clarified in the scope of this paper and the experimental prototype has been elaborated. However, to achieve maximum accuracy in predictions and make them more or less relevant, machine learning systems require a large amount of the data to train the models. At this moment the data collection is in active process. At the same time this kind of system requires significant clinical research and collaboration with psychologists to tune and test the model for real recommendations and reduce possible associated risks. Further work on the implementation and testing of the recommendation engine, empirical experiments and impact evaluations are considered for the next step when the appropriate amount of the data will be collected. Music creation by artificially intelligent systems with particular music attributes to move states of human emotions can be considered as the further elaboration work in this context. spend that time listening to music.

VIII. REFERENCES

- [1] Florence, S. Metilda, and M. Uma. "Emotional Detection and Music Recommendation System based on User Facial Expression." IOP Conference Series: Materials Science and Engineering. Vol. 912. No. 6. IOP Publishing, 2020.
- [2] James, H. L., Arnold, J. J. A., Ruban, J. M. M., Tamilarasan, M., & Saranya, R. (2019). Emotion based music recommendation system. *Emotion*, 6(03).
- [3] Arto Lehtiniemi and Jukka Holm, "Using Animated Mood Pictures in Music Recommendation", 2012 16th International Conference on Information Visualisation.
- [4] Bruce Ferwerda and Markus Schedl "Enhancing Music Recommender Systems with Personality Information and Emotional States": A Proposal: 2014
- [5] Vinay p, Raj p, Bhargav S.K., et al. "Facial Expression Based Music Recommendation System" 2021 International Journal of Advanced Research in Computer and Communication Engineering, DOI: 10.17148/IJARCCCE.2021.10682
- [6] Deny John Samuvel, B. Perumal and Muthukumar Elangovan, "Music rec-ommendation system based on facial emotion recognition", 2020
- [7] A. S. Bhat, V. S. Amith, N. S. Prasad and D. M. Mohan, "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, 2014, pp. 359-364, doi: 10.1109/ICSIP.2014.63.
- [8] Xianghua Fan, Fuyou Zhang, Haixia Wang and Xiao Lu, "The system of face detection based on OpenCV," 2012 24th Chinese Control and Decision Confer-ence (CCDC), 2012, pp. 648-651, doi: 10.1109/CCDC.2012.6242980