

Melomaniac- Emotion Based Music Recommendation System

Ishwar More¹, Vedant Shirpurkar², Yudhish Gautam³, Natasha Singh⁴

^{1,2,3,4}U.G Students of Computer Engineering, Sinhgad Institute of Technology, Maharashtra, India

ABSTRACT

An innovative approach that generates a playlist for the user according to his /her mood. In today's modern world music has become an integral and crucial part of human life and advanced technologies. Hard part of listening a song according to our mood is to find the apt song which can be overcome by using advanced CNN techniques which precisely detects users emotions. Problems like detection and location of faces in a cluttered image, facial feature extraction and expression classification should be detected by Facial Expression Recognition system. The model after training precisely classifies the emotions in the category of angry, happy, sad, neutral.

KEYWORDS – Emotion Recognition, Song Fetching, Activation Function(Relu), Adam Optimizer, Feature Extraction Module, CNN, Max Pooling, etc.

INTRODUCTION

This system eliminates the extra pain that the user has to go through to search a song that reflects his emotions by predicting users emotions and suggesting him/her song accordingly. Face capture and emotion recognition gives an edge by predicting as to how the user is feeling with the help of deep learning. The best way to detect a human's emotion is by observing his/her facial expression. Since every human face has a unique demographics it becomes easy for the model to compute and detect any emotion and thus computation time also decreases. Thus face emotion happens to be the most efficient way to predict emotion thus helping us to suggest music accordingly.

Motivation:

We are all melophile thus we named this software accordingly. As music can be a great antidepressant, we are creating such software for those who need help and cannot express their emotions. Each piece of music has its own emotion which we try to mend with person's emotion for greater effect. This software would be a revelation as it would not only help in listening music but also it will help in expressing one's true emotion.

Objectives and Goals:

- 1) To implement Convolution Neural Networks for classification of facial expressions.
- 2) User facial expressions based recommended music.
- 3) System correctly suggests a playlist to the user by precisely extracting and detecting the emotion of the user.
- 4) To enhance user experience.
- 5) To link song listening experience with growing and evolving technologies.

RELATED WORK OR LITERATURE SURVEY

“Emotion Based Music Player”

Author: Aditya Gupte, Arjun Naganarayanan, Manish Krishnan

In today's world stress has become a major reason behind deteriorating mental health because of various factors such as economy, health, expense, pollution etc. Music plays a vital role by distracting one's mind from the difficulties of day to day life. Music works best when it is in alignment of the emotion the user is feeling. To contribute to this cause this paper proposes music player which solely operates on user's various emotion such as happy, angry, neutral and sad.

An Intelligent Music Player based on Emotion Recognition

Author: ramya.ramanathan,rohan

This paper focuses on using python for detection of emotion through which music is recommended. Pattern clustering techniques are also mentioned through which emotion classification is done. Lyrics and melody of the song is used to categorize music according to predicted emotion.

Emotion Based Music Recommendation System Using Wearable Physiological Sensors

Author: De`ger Ayata, Yusuf Yaslan and Mustafa E.Kamasak

Most of the existing music recommendation system content based and collaborative recommendation engines. System stores previous results and recommendations so that they can be used for future recommendations. However this feature is not satisfactory enough to recommend music as emotion too plays a vital role in music suggestion. This paper focuses on the methodology of learning the emotion of user from the signals fetched from the Wearable Physiological Sensors.

PROPOSED SYSTEM

One's mood can be very precisely detected from his/her face. Camera is used to capture the human face which acts as an input. After capturing the image with the help of certain algorithms we extract the features which eventually helps us deduce the emotion or mood of an individual. Predicted emotion can then help in recommending a playlist to the user which is created beforehand for every emotion. Emotion based music recommender captures the face and extracts features and trains itself in order to learn the pattern so that it correctly identifies emotion for different faces. Since music is being used in alignment of the emotion or mood it becomes extremely important for the system to detect the emotion as precise as possible

Project is implemented in two phases:

1. Emotion detection software detects the emotion using facial expression with the help of python
2. Integrating the code with web service in order to play the recommend music based on facial expression.



ANGRY



HAPPY



NEUTRAL



SAD

ADVANCED SYSTEM ARCHITECTURE

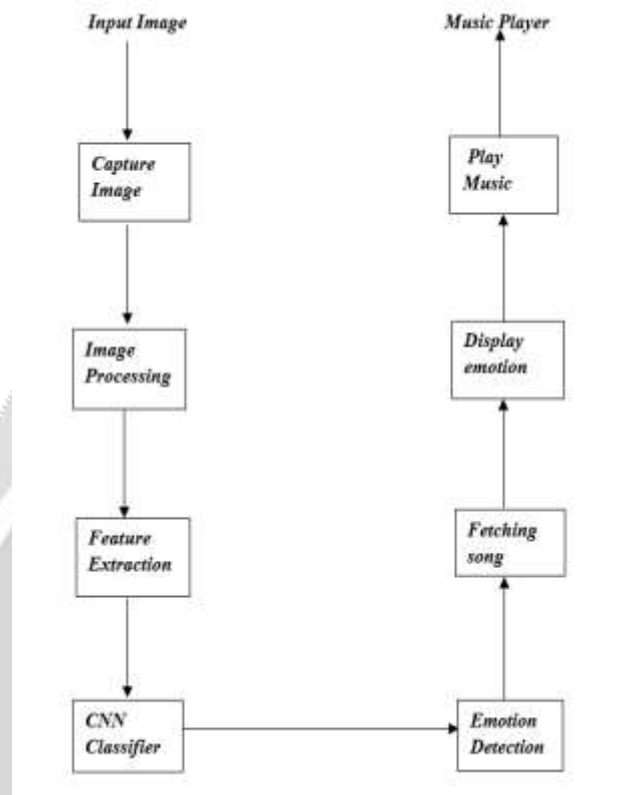


Figure: Advance System Architecture

Capture Image: The main goal is to detect an emotion from real time face image and recommend the music accordingly. In the beginning image is captured using a library called cv2 (opencv).

Process Image: By using several functions (facecascade) of cv2 we tried to detect and grab the face from real time image.

Feature Extraction: In this phase the task is to identify emotion among several classified emotions. This is the phase of pattern recognition in model. Greyscale and RGB transformation techniques are used to normalize image dataset that is further used for feature extraction.

CNN classifier: Convolutional neural network is a multilayered neural network to detect complex features of data. Several modules like Conv 2d, Max pooling is used to play with the images.

Emotion: In this phase the image from image dataset is classified. Tensorflow library is used for fast numerical computations and specifically to work with arrays where keras is acting as an interface for tensors (Tensors = mathematical object describe physical properties like scalars and vectors Scalar (rank 0

tensor) vector(rank 1)). Flatten , dense , dropout etc. are used as the layers of sequential model building for pattern recognition in image .

Fetching local music data : In this phase music playlist is manually classified according to various emotions .

Music feature Extraction : Every music is classified in accordance with the specific emotion .

Cluster of music : From that cluster of music a song is played randomly.

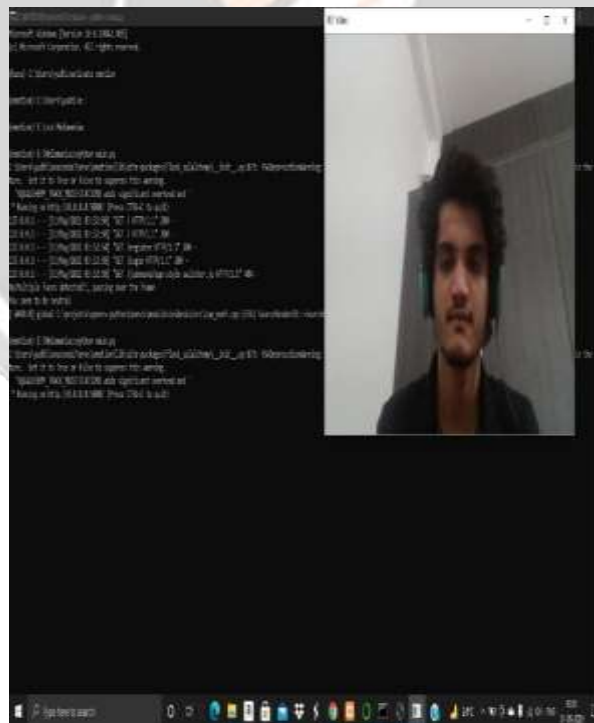
Methodology and Results

Methodology is divided into three phases:

1. Face Recognition
2. Emotion Classification
3. Music Recommendation

1. FACE DETECTION

The main goal of face detection technique is to recognise the face in the frame by eliminating the external noises in real time. Image is captured using OpenCV library. Along with cv2 usage of FisherFaces techniques increase the feature extraction efficiency as FisherFaces uses Principal Component Analysis(PCA) and Linear Discriminant Analysis(LDA). Its also better than EigenFaces as it doesn't consider illumination as a separate feature. Once the image is captured it is converted to gray scale and then feature vector is created by calculating Euclidean distance.



2. EMOTION CLASSIFICATION

For emotion classification we have used deep learning's CNN model. Convolutional neural network is a multilayered neural network to detect complex features of data. CNN model applies various filters to the image which helps the model to understand the pattern of the images passed. In this system to train the model we pass number of images for each of the 4 emotions so that the model with the help of CNN layers becomes trained enough to predict the emotions of the test data. In CNN to increase the efficiency of the model we do MaxPooling which brings in non-linearity making the model capable enough to deal with a variation of data.



3. MUSIC RECOMMENDATION:

We have manually parted the music using JavaScript as arrays and assigned them to each of the 4 emotions which is why whenever an emotion is detected the playlist automatically suggests a particular song.



Advantages:

1. Extremely fast feature computation.
2. Efficient feature selection.
3. Instead of scaling the image itself (e.g. pyramid-filters), webscale the features.

Disadvantages:

1. Accuracy of the model is the major concern in case of emotion detection ,due to which there are cases where predicted emotion happens to be wrong
2. Humans have a wide range of emotions, which the proposed system does not consider
3. Music categorization is done manually, which is a exhausting and a relative way for classifying music.

CONCLUSION

In this paper, we proposed an algorithm for web cam based emotion recognition with no manual design of features using a CNN. Using the extracted emotion ,music is recommended.

FUTURE SCOPE

Future work should attempt to combine our technique with other modalities such as audio modality, including working with other datasets. This domain itself holds countless possibilities. In future a wide range of emotions can be taken into account which will eventually increase the precision of predicting the right emotion. Even the model can be optimized with the help of additional features which might give users an option to travel to places and listen to songs accordingly based on the users emotion.

Even the words of the song can be used to detect emotion and accordingly suggest user the song that is apt for the emotion he is feeling.

REFERENCES

Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556 (2016)

- [1] Alizadeh, Shima, and Azar Fazel. "Convolutional Neural Networks for Facial Expression Recognition." Stanford University, 2017
- [2] Raghuvanshi, Arushi, and Vivek Choksi. "Facial Expression Recognition with Convolutional Neural Networks." Stanford University, 2018
- [3] Healthcare Department, "Healthcare Department news," 2 February 2018. [Online].
- [4] I. Cohen, A. Garg and T. S. Huang, *Emotion Recognition from Facial Expressions using Multilevel HMM*, 2015.
- [5] McFee, Brian, Colin Raffel, Dawen Liang, Daniel PW Ellis, Matt McVicar, Eric Battenberg, and Oriol Nieto. "librosa: Audio and music signal analysis in python." In Proceedings of the 14th python in science conference, pp. 18-25. 2015. [9] Giannakopoulos, Theodoros. "