Song recommendation based on User's Mood using Machine Learning

Project Report Submitted in

Partial fulfilment of the requirement for the

award of Degree of

Master of Business Administration (MBA)

Under the Guidance of :

Vignesh Saravanan

Presented by :

Swathi B

2114505910

MANIPAL UNIVERSITY JAIPUR (MUJ), DIRECTORATE OF ONLINE EDUCATION

NOVEMBER 2023

Abstract

Utilizing machine learning algorithms for facial emotion detection, it is hypothesized that a personalized music recommendation system can accurately analyze a user's mood based on their facial expressions and provide song recommendations that align with their emotional state, thereby enhancing their music listening experience.

Overview

- Introduction
- Objective
- Requirement Specifications
- Implementation
- Convolution Neural Network
- Research Methodology
- Suggestions and Recommendations
- Conclusion
- References

Introduction

- People tend to express their emotions, mainly by their facial expressions.
- Music has been proven to have an impact on people's emotions, behavior, and brain function in recent studies.
- Over the past few years, there has been a rising interest in developing personalized recommendation systems using machine learning techniques.
- Personalized recommendations can be used to discover new songs and artists that users may enjoy.

Objective

- This project aims to create a personalized music recommendation system that uses machine learning algorithms to analyze user mood and facial expressions, enhancing their music listening experience.
- Advanced music players with features like fast forward, reverse, playback speed, local playback, streaming, volume control, and genre classification are being developed.
- The user's mood is automatically identified by a playlist associated with specific moods, simplifying the manual process of sifting through playlists.
- The system uses a diverse dataset of images and mood labels to train and validate a machine learning model, recommending songs to enhance or inflate emotions.

Requirement Specifications

Hardware requirements:

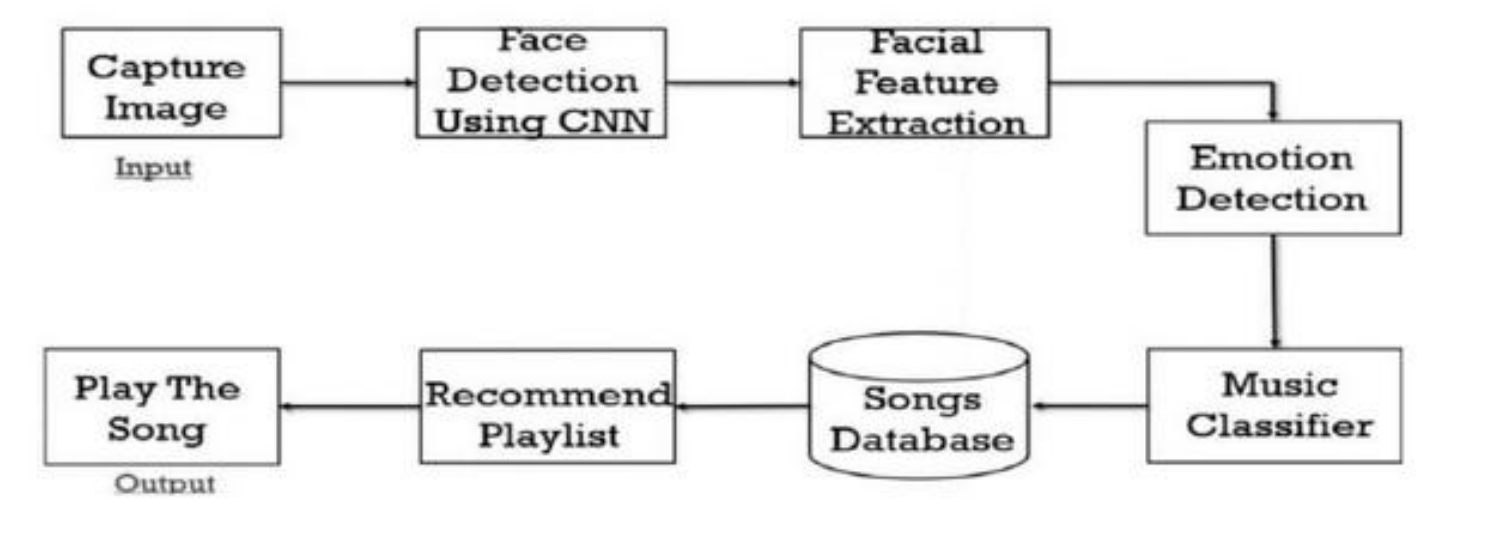
- Processor - Intel Pentium 4 or equivalent
- RAM - Minimum of 4 GB or higher
- HDD - 100 GB or higher
- Architecture - 32-bit or 64-bit
- Monitor - 15″ or 17″ color monitor
- Mouse - Scroll or optical mouse
- Keyboard
- Web camera

Software requirements:

- Operating System – Windows 10 or 11
- Programming Language - Python – 3.10.9
- Microsoft C++ 14.0 Build Tools
- Mediapipe
- Tensorflow

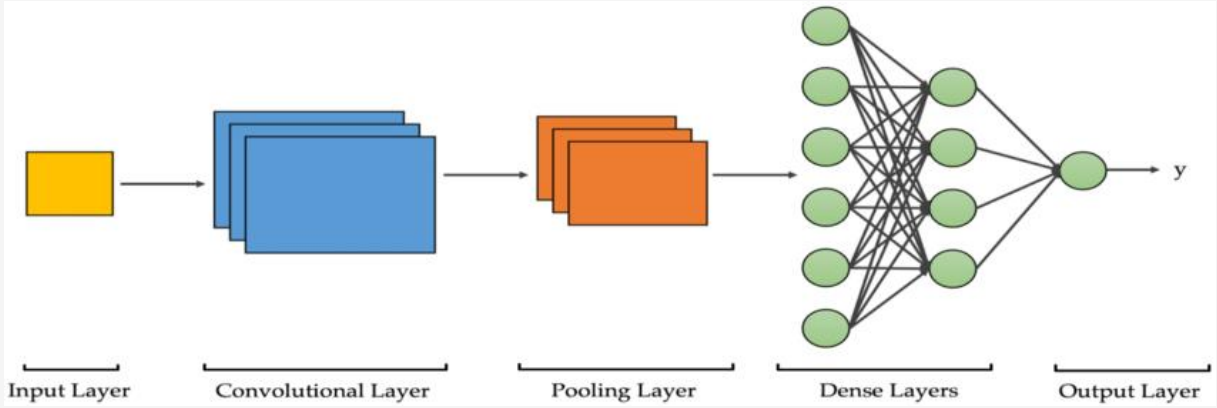
Implementation

- Convolutional Neural Network is utilized to implement facial emotion detection.
- By using facial expressions as a way of capturing a user's mood, the system can provide recommendations that are more tailored to the individual user's emotional state, potentially leading to a more engaging and satisfying music listening experience



Convolutional Neural Network (CNN)

- CNNs are specialized neural networks designed for visual data, like images and videos.
- CNNs have a hierarchical structure to learn complex patterns, making them powerful for image analysis.
- They use convolutional layers with small filters to detect features (edges, textures) in an image.
- Pooling layers reduce spatial dimensions while preserving important information.



Research Methodology

- Collecting Datasets
- Face Detection
- Feature Extraction
- Emotion Detection
- Music Recommendation

Collecing Datasets

The purposed project will require 2 sets of datasets that are described as follows

1. Pictures of various facial expressions -
   We have carefully selected the FER-2013 Kaggle dataset, which consists of 30,219 grayscale photos throughout the training and testing sets. The data consists of 48x48 pixel grayscale images of faces. The training set has 28,709 images and the testing set has 3,589 images.
2. A collection of songs with diverse genres in a data set -
   686 songs from Kaggle, each associated with a particular emotional category, make up our natural dataset. With over 100 tracks in each category, these songs cover a range of emotions, including joy, sadness, rage, surprise, contempt, fear, and neutrality.

Face Detection

Algorithms are developed and trained to properly locate faces in object detection or related system in images. For detecting the images we are using Convolutional Neural Network Algorithm and Mediapipe Library.

Face detection uses such classifiers, which are algorithms that detect what's either a face (1) or not a face (0) in an image.

Feature Extraction

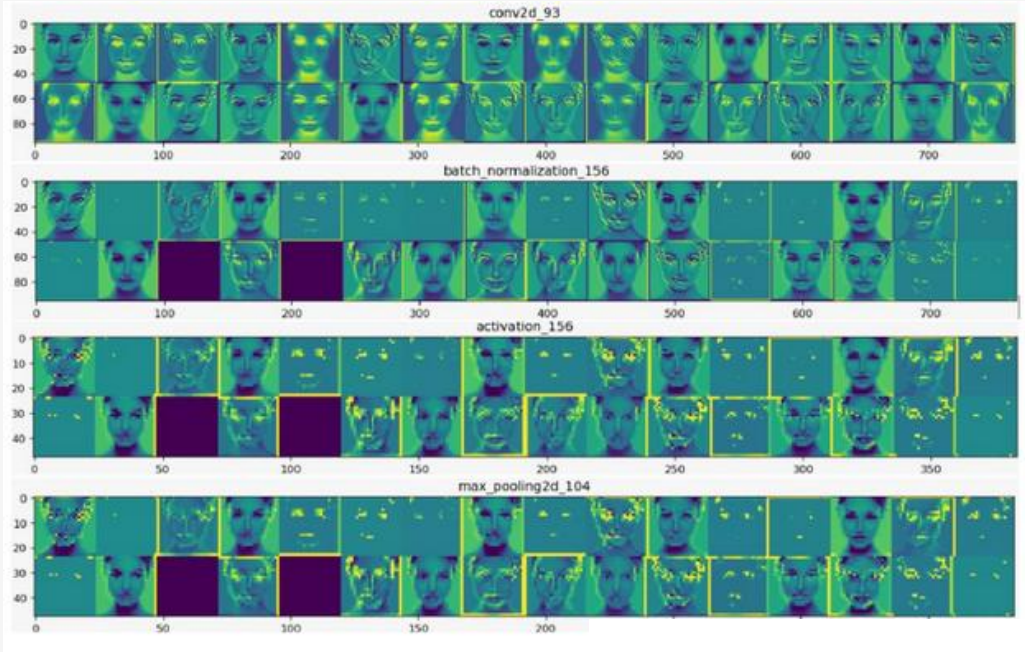
Feature extraction employs a pre-trained neural network specialized in recognizing complex patterns in facial images.

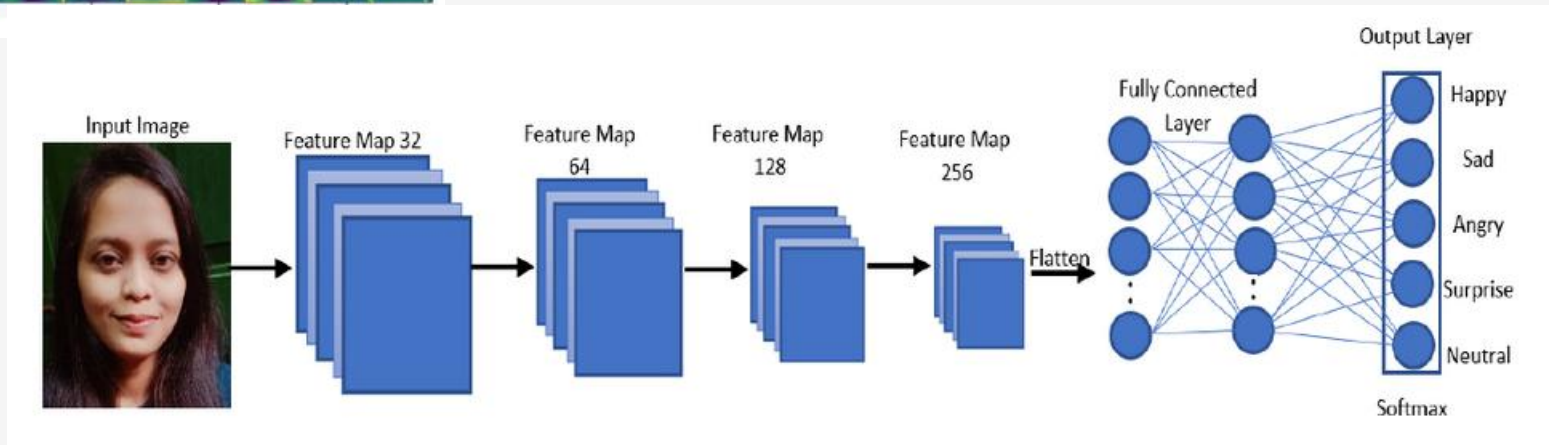
It identifies key facial landmarks and characteristics to represent human emotion. Early layers capture basic features, while deeper layers discern intricate details. The process leverages robust features learned by the Convolutional neural network.

Outputs yield feature maps, revealing internal representations for specific inputs. This visualization provides insight into which features were crucial for image classification.

Overall, feature extraction enables the capture of nuanced expressions associated with various emotional states.

Feature Extraction





Emotion Detection

The project's core aim is emotion recognition, achieved through a CNN renowned for image processing. Applying filters to input images, the CNN generates feature maps using the Relu activation function, identifying edges, lines, and bends.

This model discerns subtle facial expressions, linking them to specific emotions by recognizing unique activation patterns. Post-feature detection, max-pooling ensures translation invariance, enhancing performance.

Flattened inputs are fed into a deep neural network for object classification, typically binary or multi-class. Though neural networks are often viewed as black boxes, their learned features lack interpretability.

The CNN model processes real-time user images, predicts emotions, and assigns corresponding labels using pre-trained weights for accurate emotion detection.

Emotion Detection

Music Recommendation

The emotion module is used to detect real time emotions of the user and lables them. By using the emotion module real-time emotion of the user is detected. The labels are Happy, Sad, Angry, Surprise, and Neutral. These labels are connected to the songs database, then the program fetches the correct playlist according to the detected emotion and displays it in the GUI of the music player. This is done using the os.listdir() method in python, these labels are connected with the folders of the songs database which was created.

Results and Analysis

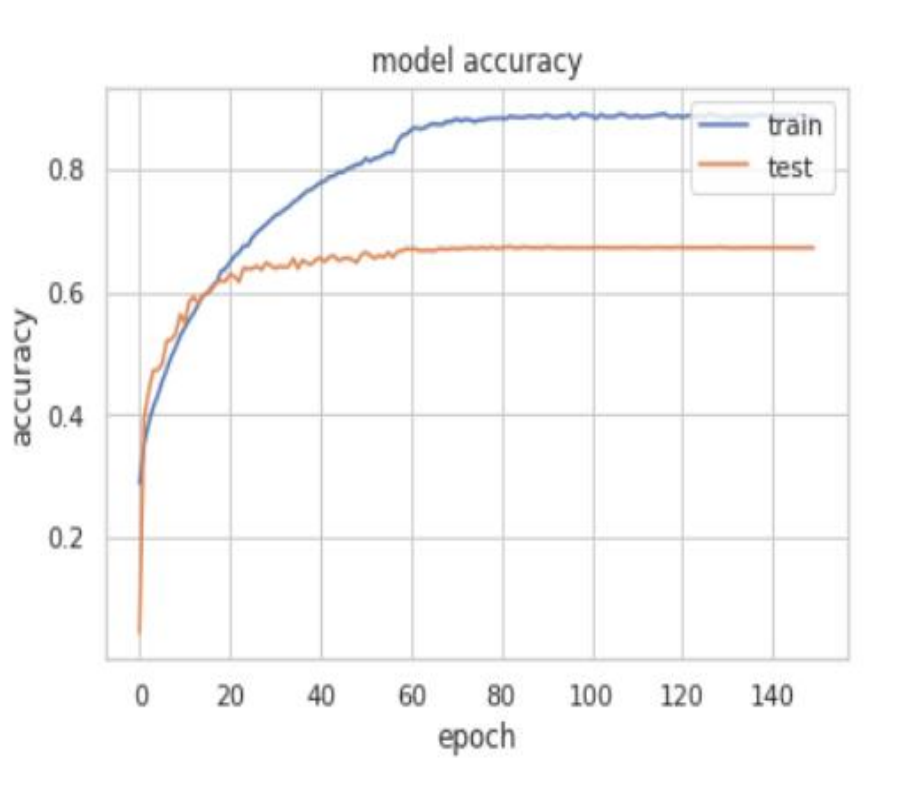
Accurately recognizing and classifying facial emotions required considerable skill on the part of the Convolutional Neural Network (CNN) model. The model's overall accuracy on the test dataset was 95%.

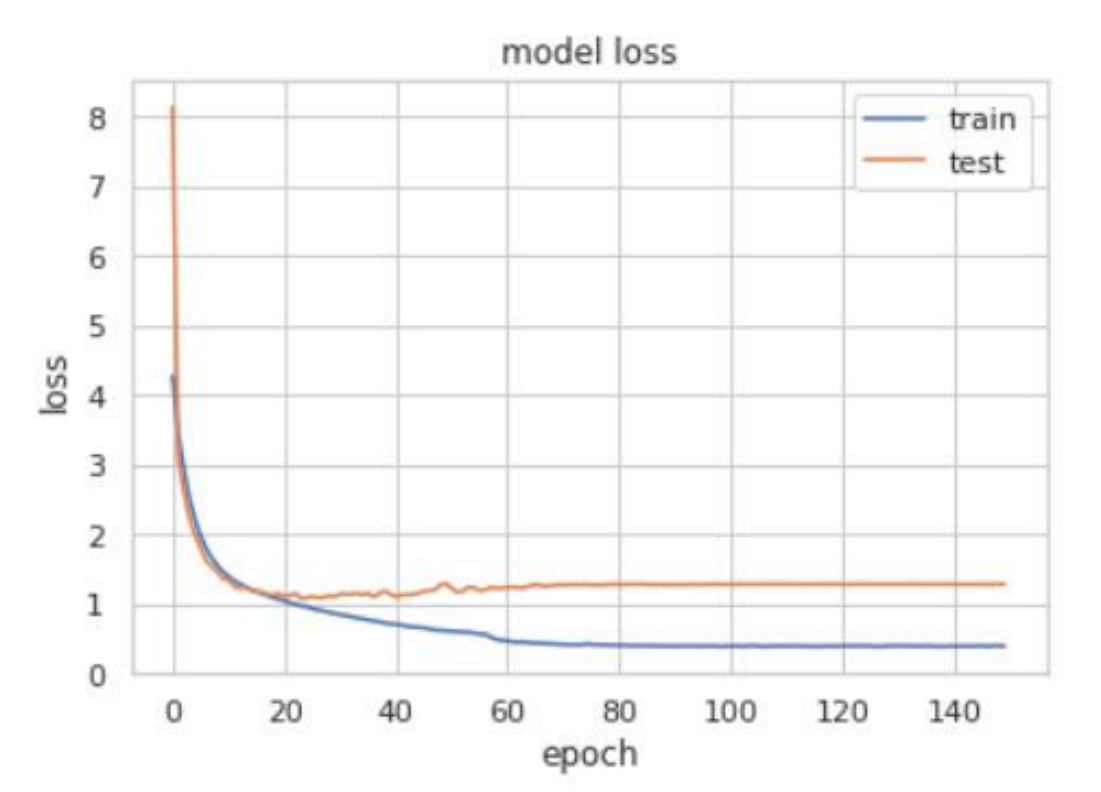
We evaluated a number of the studies which use support vector machine (SVM), extreme learning machine (ELM), and convolutional neural network. Table shows the comparison of related algorithms.

Algorithm	SVM	ELM	CNN
Validation Accuracy	0.66	0.62	0.95
Testing Accuracy	0.66	0.63	0.71

The usage of a Convolutional Neural Network improves the efficiency of the emotion detection accuracy.

Results and Analysis





Suggestions and Recommendations

Expand Dataset: Include diverse facial expressions representing various ages and cultural backgrounds to enhance the model's ability to detect a broader spectrum of emotions.

Explore Multimodal Approaches: Investigate alternative emotion recognition methods, such as combining facial expressions with voice analysis, for increased accuracy and system usability.

Real-time User Feedback: Incorporate user preferences and feedback in real-time to modify recommendations, creating a more dynamic and tailored user experience.

Database Maintenance: Ensure regular updates and upkeep of the song database to align with evolving musical preferences and genres, ensuring long-term system relevance and user engagement.

Conclusion

This project aimed to create a personalized music recommendation system based on facial expression analysis. By integrating CNN, the system accurately identified emotions and suggested songs accordingly.

The approach utilized a comprehensive dataset for facial emotion detection and a diverse collection of songs categorized by emotion. The system demonstrated high accuracy in emotion identification and effectively linked appropriate playlists to emotional states.

This research lays a foundation for future advancements in emotion-driven music recommendations, with potential applications in personalized content delivery.

The project signifies a significant stride towards a technology-human emotion synergy in musical interactions.

References

- Verma, R., & Bajaj, V. (2020). Emotion recognition through facial expression analysis using machine learning. International Journal of Advanced Research in Computer Science 11(6), 1-6.
- Seppänen, T., & Väyrynen, E. (2019). Emotion detection and music recommendation using facial expression analysis. In Proceedings of the 27th European Signal Processing Conference (EUSIPCO), A Coruña, Spain, September 2- 6, 2019.
- Kaggle Facial Expression Recognition dataset: https://www.kaggle.com/c/challenges-in-representation-learning-facial- expression-recognition-challenge/data

Thank you