Intermediate Report

Music Therapy Treatment based on Emotion Detection

Elisha Maria K Orlina
A20520119

Raj Shah
A20524266

Abhishek Jaiswal
A20380004

1. Introduction

Facial expressions play a critical role in revealing a person's emotions [2]. The future of computer vision systems may rely heavily on dynamic user interactions, such as those found in present-day computer systems. Facial emotions have a significant impact on security, entertainment, and human-machine interaction (HMI), and they can be conveyed through a person's lips and eyes.

In today's world, having a vast music playlist is commonplace, as music is a crucial source of entertainment and can even serve as a therapeutic tool in some cases. Music can uplift one's mood, boost happiness, and help reduce anxiety, and it has been an integral part of the human experience for ages [4]. The correlation between music and mood is so apparent that its impact is often known to improve one's emotional and psychological well-being.

This project proposes a system that detects a person's mood through facial emotion recognition using convolutional neural networks and generates a music playlist based on the detected mood. Unlike humans, machines find detecting facial expressions to be a challenging task. However, with advancements in technology, Convolutional Neural Networks (CNNs) have proven to be an efficient solution for facial emotion recognition. The model will be trained on the FER Dataset, which contains grayscale facial images that are 48 pixels in height and width.

The training set comprises 28,709 examples, and the public test set comprises 3,589 examples. Keras tuner will be used to optimize the model and select the optimal number of convolutional layers, flattened layers, and dense layers to achieve the highest accuracy.

Once the user's face has been detected using a Haar-cascade classifier in the image captured through the webcam, the image is passed on to the trained and optimized CNN model to predict the user's mood. A playlist is then generated from the music database based on the predicted mood. The songs are pre-fed to the system as a dictionary which is determined to soothe the predicted mood of the user and can act as the therapy treatment. Songs are randomly selected from the respective mood category, and the system generates the selected songs through external storage and displays them to the user to enhance their mood.

This project aims to provide an alternative method of suggesting songs to users based on their facial emotions and to create a high-accuracy CNN model for detecting facial emotions.

2. Problem Statement

We are trying to solve the problems mentioned below-

1. Helping people to overcome mental stress, anxiety, and acute depression, including other mood fluctuations like- gloom, worry, and fear (*untouched yet*). Thus it will be more like a music therapy system rather than a mere music recommendation system.
2. We are also focussing on people who drive unaccompanied and are going through mental issues. Our model will detect the mood and play as a positive distraction, maybe by lifting the mood.

3. Description of Data

The dataset is called the "Challenges in Representation Learning: Facial Expression Recognition Challenge" dataset. It was created for the ICML Workshop on Challenges in Representation Learning. This dataset comprises images of human faces displaying different facial expressions, which are categorized into seven classes such as angry, worried, fearful, happy, sad, surprised, and neutral. These images were collected from Google Images and were then labeled by human annotators using the Amazon Mechanical Turk platform.

The dataset contains two files, "fer2013.csv" and "test.csv". The "fer2013.csv" file contains 28,709 training examples and 3,589 public test examples, while the "test.csv" file contains an additional 3,589 private test examples. Each example in the dataset includes a 48x48 grayscale image of a human face and a corresponding label that indicates the expression displayed in the image.

Here's an in-depth interpretation and insights of the dataset-

<table>
<tr>
<td>

```
df[' Usage'].value_counts()
```

```
Training       28709
PublicTest      3589
PrivateTest     3589
Name:  Usage, dtype: int64
```

</td>
<td>

Observations-

1. Training data count- 28709
2. Test data count- 3589

</td>
</tr>
<tr>
<td>

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 35887 entries, 0 to 35886
Data columns (total 3 columns):
 #   Column   Non-Null Count  Dtype
---  ------   --------------  -----
 0   emotion  35887 non-null  int64
 1    Usage   35887 non-null  object
 2    pixels  35887 non-null  object
dtypes: int64(1), object(2)
memory usage: 841.2+ KB
```

</td>
<td>

Observations -

1. There are no null values.
2. We have mainly two data types- int and object

</td>
</tr>
</table>

4. Project Progress

4.1 Survey of the Background Research & Developments

Through a thorough review of the existing literature, we have identified various approaches to implementing a music recommender system, and we have leveraged this knowledge to develop our system's objectives. By focusing on the correlation between music and mood, our system offers a novel way of improving users' moods through music. In the current scenario, most of the recommendation systems focus on moods like - happy, sad, neutral, or surprised.

However, our idea aims to revolutionize the music industry by leveraging the power of AI-powered applications to create a state-of-the-art music recommender system. Our system offers a unique approach to music recommendations by utilizing facial emotion recognition, especially facial emotions that are not explicitly targeted yet, such as worry, fear, gloom, etc. Therefore, through this new dimension, we can determine the user's mood more accurately and provide them with a curated playlist that matches their current emotional state.

Additionally, our model will try to lessen users' efforts in creating and managing playlists. Furthermore, our music recommendation system based on facial emotion recognition offers a **user-friendly** and **innovative** solution to the traditional approach of music recommendations, with the ability to detect certain un-explored emotions such as worry, fear, and gloom, among others.

In the future, we envision expanding the scope of our system to include a mechanism that would be beneficial in **music therapy treatment** in the wider application of healthcare, particularly for patients experiencing mental stress, anxiety, acute depression, and trauma.

4.2 Tasks | Completed

The preliminary plan follows the below steps-

4.2.1. Data Collection- We will be collecting adequate data through open-source datasets available on the Kaggle, & UCI ML repo.

4.2.2. Exploratory Data Analysis (EDA) - To get insights from the data we will do the EDA after data preprocessing.

4.3 Tasks | In progress

4.3.1. Model Training & Hyperparameter Tuning - After getting the data and performing resizing and rescaling we will separate the data into training, testing, and validation sets. Thereafter we will be using the Keras to train the model.

4.3.2. Documentation - This is an ongoing process that is related to other tasks.

4.4 Tasks | To Do

4.4.1. Batch Normalization - It is the technique of adding extra layers to a deep neural network to make it faster and more stable. In this, first, the input will be normalized, and later rescaling and offsetting will be performed.

4.4.2. Getting the best model parameters - We will try different models to check which model works better for us and then we'll move ahead with that as per the accuracy results.

4.4.3. Visualization - In this phase, we will try to visualize some of the outcomes.

4.4.4. Documentation - This is an ongoing process that is related to other tasks in this section.

References

1. S Metilda Florence and M Uma 2020 IOP Conf. Ser.: Mater. Sci. Eng. 912 062007
2. Ramya Ramanathan, Radha Kumaran, Ram Rohan R, Rajat Gupta, and Vishalakshi Prabhu, an intelligent music player based on emotion recognition, 2nd IEEE International Conference on Computational Systems and Information Technology for Sustainable Solutions 2017. https://doi.org/10.1109/CSITSS.2017.8447743
3. Roberto De Prisco , Alfonso Guarino , Delfifina Malandrino and Rocco Zaccagnino. Induced emotion-based music recommendation through reinforcement learning. *Appl. Sci.* 2022, *12*, 11209
4. Zhuwei Qin, Fuxun Yu, Chenchen Liu, Xiang Chen. How convolutional neural networks see the world - A survey of convolutional neural network visualization methods. Mathematical Foundations of Computing, May 2018.
5. Frans Norden and Filip von Reis Marlevi, A Comparative Analysis of Machine Learning Algorithms in Binary Facial Expression Recognition, TRITA-EECS-EX-2019:143
6. Research Prediction Competition, Challenges in representation learning: facial expression recognition challenges, Learn facial expression from an image, (KAGGLE)