



ML APPROACH FOR MUSICAL THERAPY USING FACIAL EXPRESSIONS

Mrs. Kavyashree G. M^[1], Bhargavi^[2], Chinmayi^[3], K. Kanchana^[4], M. S. Sudhir Srinivas^[5]

^[1] Asst Professor, Dept of CS Engg Sir M. Visvesvaraya Institute of Technology, Bangalore, Karnataka, India

^[2]^[3]^[4]^[5] UG Student, Dept. of CS Engg Sir M. Visvesvaraya Institute of Technology, Bangalore, Karnataka, India

Abstract: Music significantly contributes to development and improves one's mood because it is a major source of amusement and motivation to continue. It can help us and give us confidence when we're down. We frequently experience a fall in mood when we listen to depressing music. We experience happiness when we listen to upbeat music. Our paper proposes a web interface, which suggests music to the user depending on his present emotions and thus uplift the user's mood. This system uses the device camera to take a picture of the user and recognizes his face & facial expressions. We have used ResNet 50 Convolutional Neural Network to build the facial expression recognizing model. We have used a dataset containing 7000+ real-time images, created by ourselves.

IndexTerms - Facial expression, Emotion Recognition, Musical therapy, Music Recommendation, CNN, ResNet50

1. INTRODUCTION

Music is one of the most powerful forms of media, which can evoke strong emotions and overwhelm listeners with subliminal messages. It effectively manipulates our feelings, which in turn affects how we feel. Other forms of media include books, movies, and television dramas; however, music conveys its message far more quickly than these other forms of media. It can help us and give us confidence when we're down. We frequently experience a fall in mood when we listen to depressing music. We experience happiness when we listen to upbeat music.

The enormous advancements in machine learning and artificial intelligence have accelerated the automation of numerous tiresome operations very challenging to accomplish manually. Another area that has recently gained importance is emotion recognition. In many situations, machine interpretation is superior to human interpretation. A person's emotional state can be read and identified by automated decision-making systems for a variety of purposes, including suggestions.

Humans typically indicate their current mood through nonverbal means such as hand movements, facial expressions, and voice tones. For this project, the face is our main point of focus because it is so important. Mood and music are frequently related. Depending on the song, we may feel joyful depressed, energized, or relaxed. Due to the significant impacts that music may have on a person's outlook and wellbeing, for managing a range of medical conditions, mood enhancement, and anxiety/stress reduction, music therapy has been explored. This program will use a camera to take real-time facial images and conduct face feature detection, mood interpretation, and emotion prediction before connecting to the music playing module which will recommend the user a playlist that will enhance their mood. Here we use the 7 basic emotions Anger, Disgust, Fear, Happy, Sad, Neutral and Surprise. This system will be applicable to the users wearing spectacles, where in none of the existing similar systems provides that feature.

2. RESEARCH OBJECTIVE

The objective of our research is to provide personalized music recommendations to users based on their current mood or emotional state. The system aims to enhance the user experience by suggesting songs that match their mood, which can have a significant impact on their overall satisfaction with the music streaming service. This involves identifying the most effective machine learning algorithms for analyzing the user's mood based on the facial image of the image. The research focuses on developing a user interface that effectively communicates

the mood-based music recommendations to the user.

The research focuses on the potential impact of the system on user engagement and satisfaction. This involves conducting user studies to evaluate the effectiveness of the mood-based recommendations compared to other recommendation approaches, as well as exploring the impact of the recommendations on the user's emotional state and overall listening experience. The research also investigates the potential for the system to increase user retention and reduce churn, as well as the feasibility of integrating the system into existing music streaming platforms. Ultimately, the research objective should aim to develop a robust and effective mood-based music recommendation system that delivers value to both users and music streaming providers.

3. LITERATURE REVIEW

The strategies suggested by many studies are discussed here.

Borui Zhang, et al[19] have suggested that the Gabor Wavelet Transform filter can be used to extract features from the facial image across five scales and eight rotations in order to capture the key visual characteristics.

Shahli, et al[14] proposed a music recommendation system implemented in MATLAB with an average accuracy of 94%. The used Bandlet Transformation on the detected face area by Viola Jones algorithm to extract geometric structures of the face.

Ankita Mahadik, et al[3] intended to create an application to recommend music based on the user's mood. They used viola-Jones algorithm for face detection, using keras, deep learning is applied to detect facial expressions. For recognizing the facial expressions and a hybrid approach of music recommendation.

Prof. Jaychand Upadhyay, et al[2] proposed a CNN model whose accuracy on the training set was 97.42% thanks to batch normalization.

Krupa K S et al[7] proposes a system for detecting the user's emotions during interactions with an emotionally intelligent chatbot using facial landmarks and semantic analysis.

Renuka S. Deshmukh, et al[16] conducted research that suggests a system to describe human emotions. Here, automatic facial emotion recognition is used as a general approach for the facial emotion recognition system. In this paper, motion of the facial parts for each of the emotions is mentioned.

Aurobind V. Iyer, et al[17] have used the Adaboost Learning Algorithm to construct strong classifiers by weighing the multiple weak classifiers and they are in turn used while cascading classifiers for face detection.

A. Phaneendra, et al[1] proposed that songs will be played once the emotion has been identified in this case. Here, the client side uses an API call to request songs from the Spotify app.

M.Keerthana, et al[4] proposed that the system has three modules namely Emotion extraction, Audio feature extraction and emotion audio recognition. Here they used a technique called Audio emotion recognition (AER) that classifies audio signals into moods.

Shlok Gilda, et al[18] have used the stochastic gradient descent ML technique to construct an adaptive music player. The user has the option to modify the class of a certain song, and SGD is applied for those users only while taking the new label into account.

4. PROPOSED SYSTEM

Our proposed system includes the below mentioned steps:

Create a Real Face Emotion Dataset:

The first step in building the website is to create a real face emotion dataset. This involves collecting images of people's faces in different emotional states, such as happy, sad, angry, etc. It is important to ensure that the dataset is diverse and balanced, with an equal number of images for each emotion.

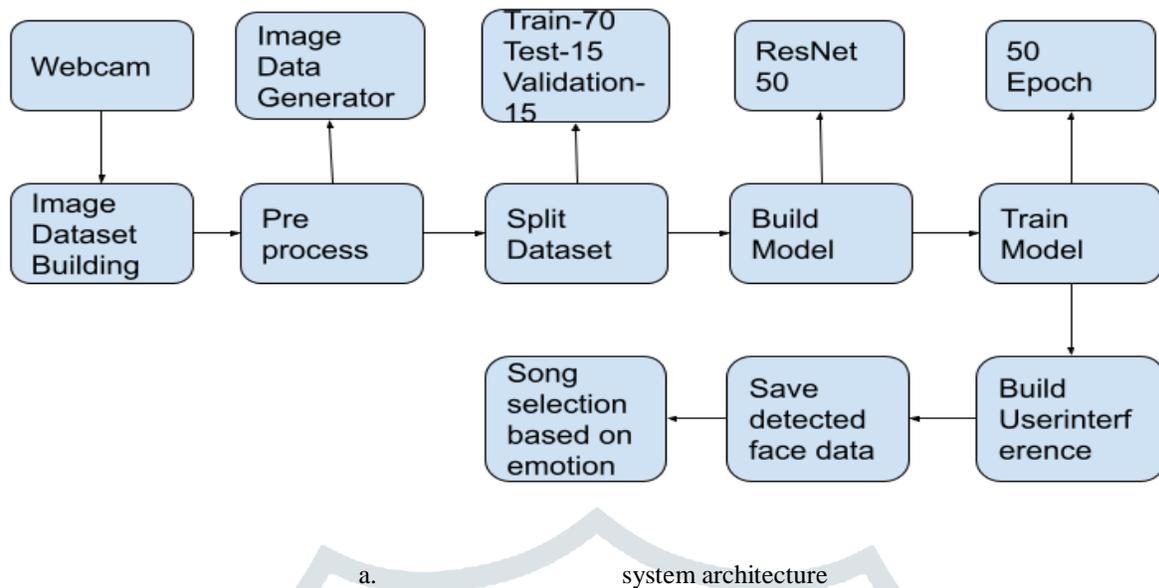
Preprocess the Dataset:

Once the dataset has been collected, it needs to be preprocessed. This involves resizing the images to a uniform size and converting them to a format that can be used by the deep learning model.

Augment the Dataset:

To improve the robustness and generalization of the model, the dataset can be augmented by applying transformations to the images. This can include rotating, flipping, and cropping the images, as well as adjusting the brightness and contrast.

The figure given below shows the workflow of our proposed system:



Build Deep Learning Model:

The next step is to build a deep learning model that can classify the emotions in the images. A popular architecture for this task is the ResNet50 model, which has been pre-trained on a large image classification dataset and can be fine-tuned for the emotion recognition task.

Train the Model:

Once the model has been built, it needs to be trained on the preprocessed and augmented dataset. This involves splitting the dataset into training and validation sets, and then feeding the images into the model in batches.

Test the Model:

After training the model, it needs to be tested on a separate test set to evaluate its performance. This involves feeding the test images into the model and comparing the predicted emotions to the ground truth labels.

Evaluate the Model:

Once the model has been tested, it needs to be evaluated to determine if it is suitable for the face emotion-based music recommendation task. This involves analyzing the model's strengths and weaknesses, as well as its performance on different emotions.

Build HTML for Frontend:

With the model in place, the next step is to build the HTML for the frontend of the website. This involves designing a user interface that allows users to interact with the website and select the emotion that they want to listen to music for.

Use Flask for Backend:

The backend of the website can be built using Flask, a lightweight web application framework for Python. Flask can be used to handle user requests, call the deep learning model to predict emotions, and serve up the appropriate music recommendations.

Deploy Web Application:

Finally, the web application needs to be deployed so that it can be accessed by users. Once the website is live, users can select their desired emotion and listen to music playlist that is tailored to their mood.

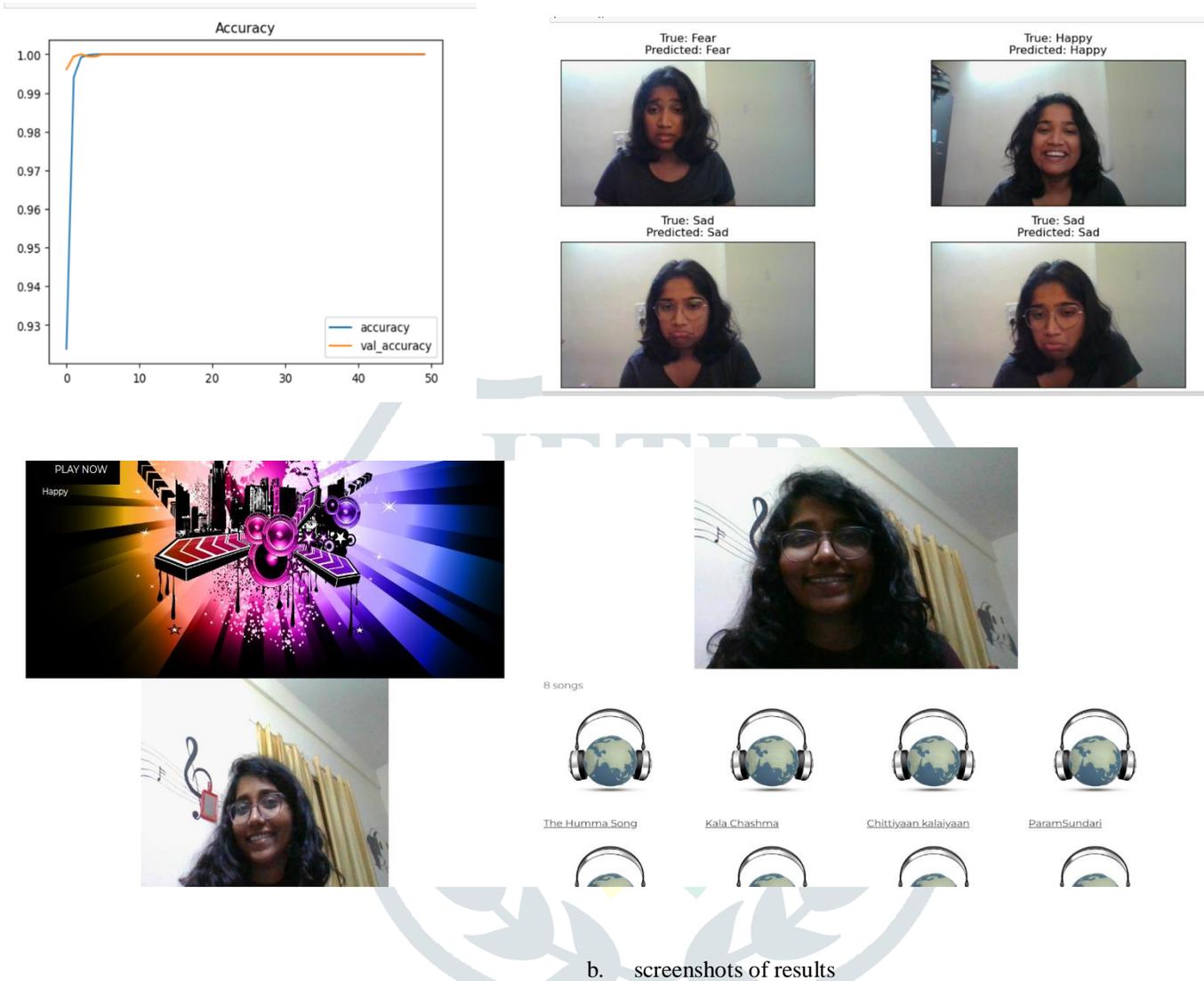
ResNet50 is a deep convolutional neural network architecture consisting of 50 layers that was introduced in 2015 as a solution to the problem of vanishing gradients in very deep neural networks. It uses skip connections to allow for the direct propagation of information from early layers to later layers, which helps to mitigate the degradation problem and enable the training of very deep networks. ResNet50 has been pre-trained on large image datasets such as ImageNet and can be fine-tuned for a variety of computer vision tasks, including image classification, object detection, and semantic segmentation.

None of the existing systems focus on the users wearing spectacles, but our proposed system will be trained to serve the users who wear spectacles as well.

5. RESULTS AND CONCLUSIONS

We have built a ML Model which suggests a playlist of songs to the user that will uplift the user's mood. This model can predict the user's emotion through captured facial image. This model takes user's facial image as input and suggests a playlist of songs as output. Based on the emotion detected from the image captured, a playlist of 8 songs is suggested to the user. Our system is applicable for users wearing spectacles as well.

The model was trained using 7000+ real-time captured images. The model was trained using ResNet-50 Convolutional Neural Network. This model gives training & testing accuracy of 100% which is far better than the existing models.



b. screenshots of results

6. REFERENCES

- [1] Madhusmitha Muduli, A. Phaneendra,, Siri Lakshmi Reddy, R. Veenasree "EMUSE – An Emotion Based Music Recommendation System" *International Research Journal of Modernization in Engineering Technology and Science*.
- [2] Vaibhav Murari , Prof. Jaychand Upadhyay , Sharan Shetty, Jarvis Trinidad "Mood Based Music Playlist Generator Using Convolutional Neural Network" *International Journal for Research in Applied Science & Engineering Technology* 2022.
- [3] Shambhavi Milgir, Ankita Mahadik, Vaishali Kavathekar and Prof. Vijaya Bharathi Jagan. "Mood based Music Recommendation System" *International Journal of Engineering Research & Technology*, 2021
- [4]M. Shruthi,M. Keerthana, S. Aravind Kumar " Emotion based music recommendation system." *IJCRT Issue 6 June 2021*.
- [5] Bhargav Satish Kumar Y, Vinay p, Raj Prabhu T, Jayanth P, A. SUNEETHA, "Facial Expression Based Music Recommendation System" *International Journal of Advanced Research in Computer and Communication Engineering* 2021.
- [6] ,Megha Gupta,Madhuri Athavle, Deepali Mudale, Upsana Shrivastav. "Music recommendation based on face emotion recognition." *Journal of informatics, Electrical and Electronics Engineering*, 2021.

[7] Sahil Choudhury, Krupa K S, Ambara G, Kartikey Rai. "Emotion aware Smart Music Recommender System using Two Level CNN" IEEE Xplore 2020

[8] Prof. Pankaj Kunekar,Pranjul Agrahari, Ayush Singh Tanwar, Biju Das, "Musical Therapy using Facial Expressions"International Research Journal of Engineering and Technology. 01 Jan 2020

[9] Yilin wu, Peizhuo Liu,Ziyang yu, Mengda Zhao, Hexu chen. "Research on automatic music recommendation algorithms based on facial micro expression recognition" 39th Chinese Control Conference, 2020.

[10] Mr.Akshay Choudhary,Prof.Sumeet Pate, Miss.Poonam Harane, Miss.Shreya Zunjarrao "Mood Based Music Player Using Real Time Facial Expression Extraction" International Journal for Research in Engineering Application & Management (IJREAM) 2019.

