

IMAGE BASED EMOTION RECOGNITION SYSTEM

Presented by group 5
Ashi Jain
Ridit Jain
Runishka Rao
Lakshit Wasan

INTRODUCTION

- In our project, we've developed an advanced image-based emotion recognition system that leverages a range of digital image processing techniques—including equalized histogram, Sobel and frequency filters, Gaussian blur, as well as grayscaling, normalization, and resizing—to refine images for accurate emotion detection.
- This system not only recognizes emotions but also simultaneously predicts age and gender using a sophisticated machine learning model.
- A unique feature of our system is its ability to enhance user experience by providing personalized music recommendations based on the detected emotions, seamlessly integrating emotion recognition with an interactive, user-centric application.

Literature Review

Paper No.	Main Approach/Methodology	Unique Features/Nuances	Limitations/Challenges
[1]	Facial emotion recognition with CNN	Integration of preprocessing and CNN	Dependence on EMOTIC dataset limits generalizability
[2]	Graph Embedded CNN (GECNN) for EEG-based emotion recognition	GECNN with graph convolutional layers and attention mechanisms	Biases in EEG data; limitations in dataset applicability
[3]	CNN with image edge detection	Edge detection to enhance accuracy	Dataset diversity limitations
[4]	Image edge detection and CNN	Edge detection for improved training efficiency	Dataset diversity limitations
[5]	EEG signals analyzed with various CNNs	Use of AlexNet, ResNet50, VGG16, and configurable CNNs	Dependence on specific signal processing parameters; limited dataset for testing

Literature Review

[6]	Facial expression classification with ResNet-50V2	Testing different epochs for performance evaluation	High accuracy vs. loss trade-off
[7]	Facial emotion recognition with transfer learning	Utilization of transfer learning techniques	Dependence on predefined datasets and classifier models
[8]	Mini_Xception architecture for emotion recognition	Custom Mini_Xception design	High computational demands for training
[9]	Open-source tools for emotion recognition	Development of a test system for simulation	Dependence on simulation; lack of real-world testing
[10]	CNN with transfer learning	Fine-tuning pre-trained VGG model; reduced training time	Need for large datasets; high-performance hardware requirements

DataSet

01

Dataset for Emotion Recognition

This file contains a 48x48 pixel grayscale image of a face

There are 28,709 samples in the training set and 3,589 samples in the test population.

There are seven categories based on the emotion on the face (0 = angry, 1 = disgusted, 2 = afraid, 3 = happy, 4 = high caring, 5 = surprised, 6 = natural)

02

Dataset for Age and Gender Prediction

The UTKFace dataset, accessible on Kaggle, includes more than 20,000 different facial images commented on with age, gender, and identity information. This comprehensive dataset covers a wide range of ages, working with both study of age estimation and regression.

Methodology for Emotion Recognition

Exploratory Data Analysis

The provided process involves extracting a dataset from an archive with "Train" and "Test" folders containing emotion-classified images. It analyzes each image's RGB properties, generates histograms, and then assesses emotion distribution in both sets. It creates pie charts to visually represent the emotion balance, ensuring dataset representation and quality assessment.

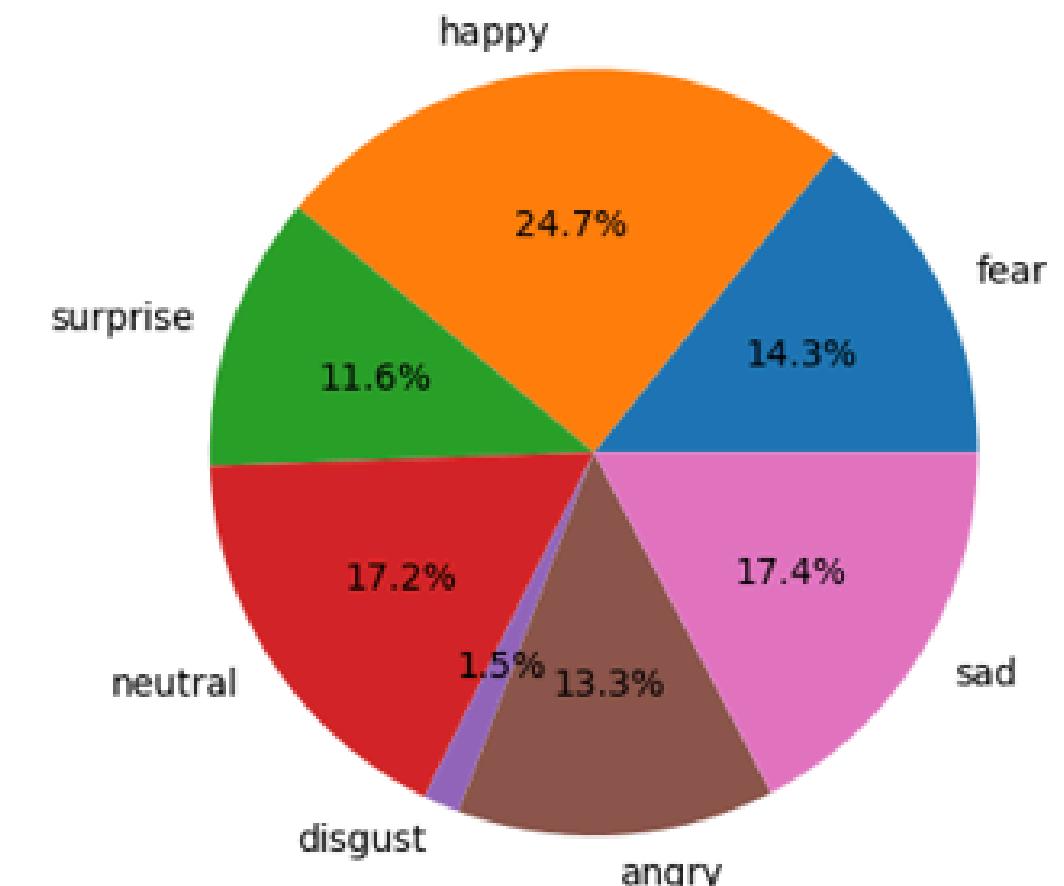
category count

0	fear	1024
1	happy	1774
2	surprise	831
3	neutral	1233
4	disgust	111
5	angry	958
6	sad	1247

category count

0	fear	4097
1	happy	7215
2	surprise	3171
3	neutral	4965
4	disgust	436
5	angry	3995
6	sad	4830

Train set



Test set

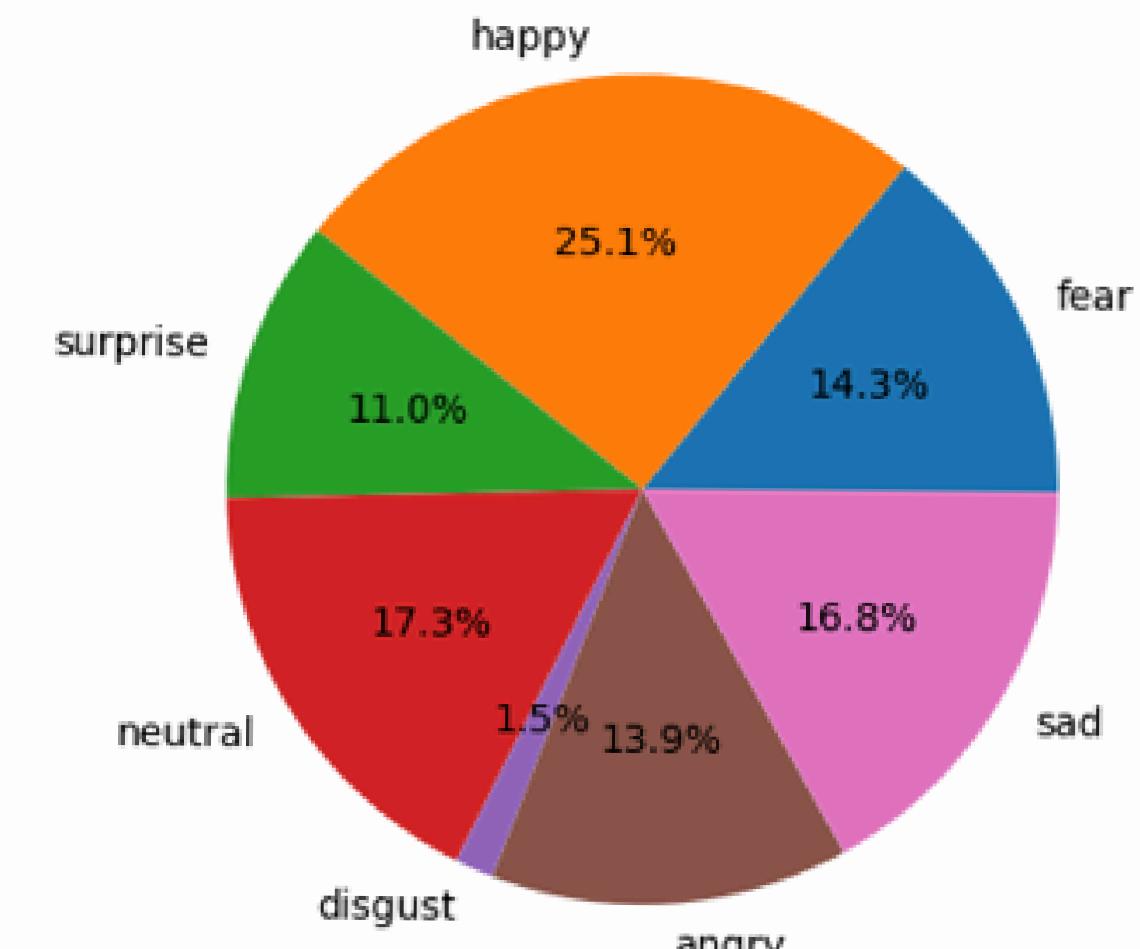


Image Preprocessing

- **Histogram equalization** is an efficient and simple image processing technique that adjusts dynamic range and contrast using histograms.
- **Sobel filter** in image processing highlights edges by convolving with 3×3 kernels for gradient approximation in horizontal and vertical directions.
- **Frequency filter** selectively modifies or extracts specific frequency components from signals or images, aiding tasks like noise reduction and feature extraction.
- **Data augmentation** enhances computer vision datasets by introducing variations, improving model generalizability. **Resizing images** to a consistent (256×256) size ensures uniform input dimensions for deep learning. **Image normalization** scales pixel values for faster convergence and better weight initialization. **Grayscale transformation** preserves luminance data but isn't applied as images are already in grayscale format.

Pre-Processing

Edge Detection

Segmentation

Smoothening

Sharpening

Input Image

Feature Extraction

Classification

Decision

Methodology for Age and Gender Prediction

Exploratory Data Analysis

The UTKFace dataset was downloaded from Kaggle using the Kaggle API and extracted locally. A home directory was created for efficient data access. Images were labeled with age and gender information, and a structured database was constructed for organized data management. Visual inspection involved displaying sample images and creating visualizations to analyze age and gender distributions, providing valuable insights into the dataset's characteristics.

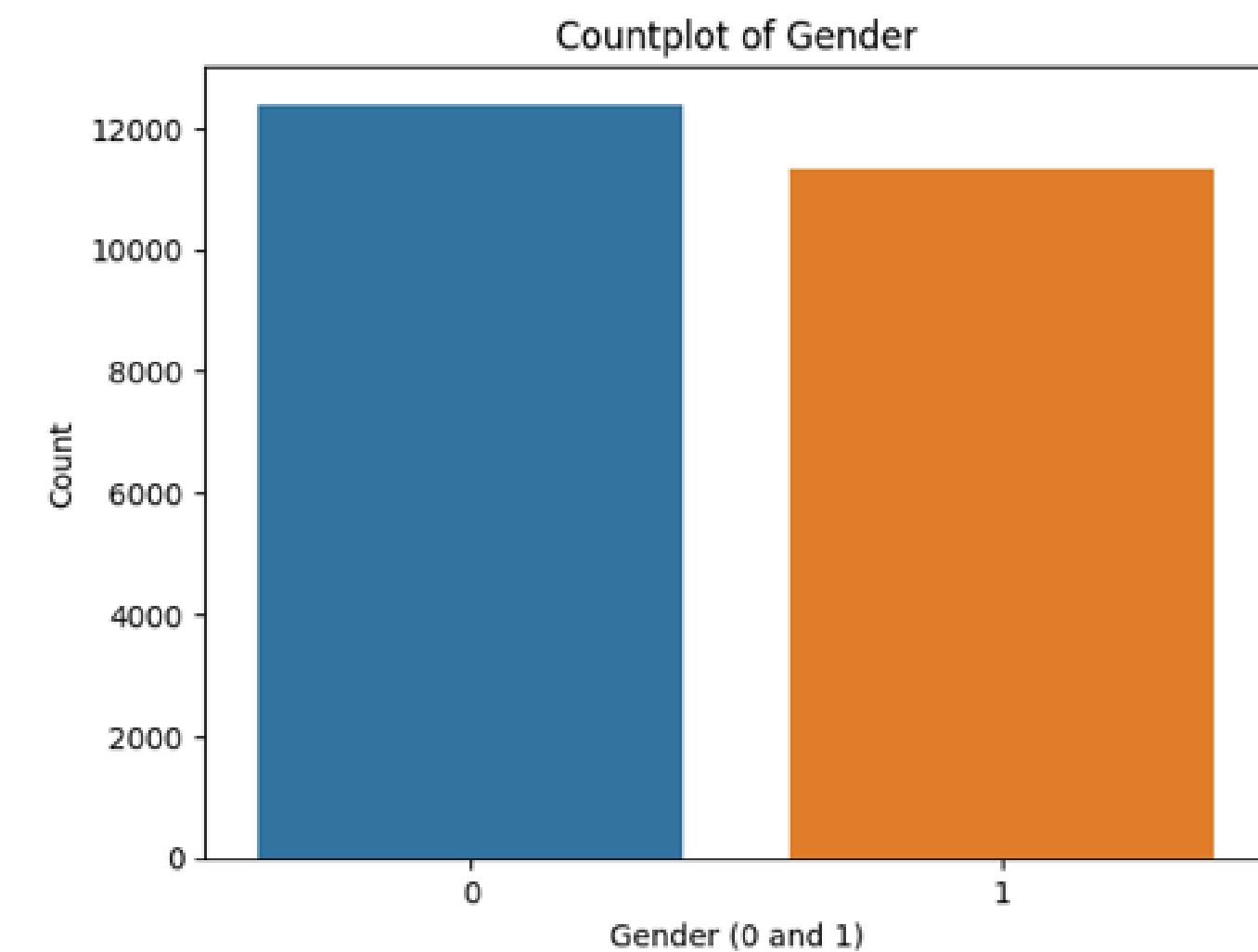
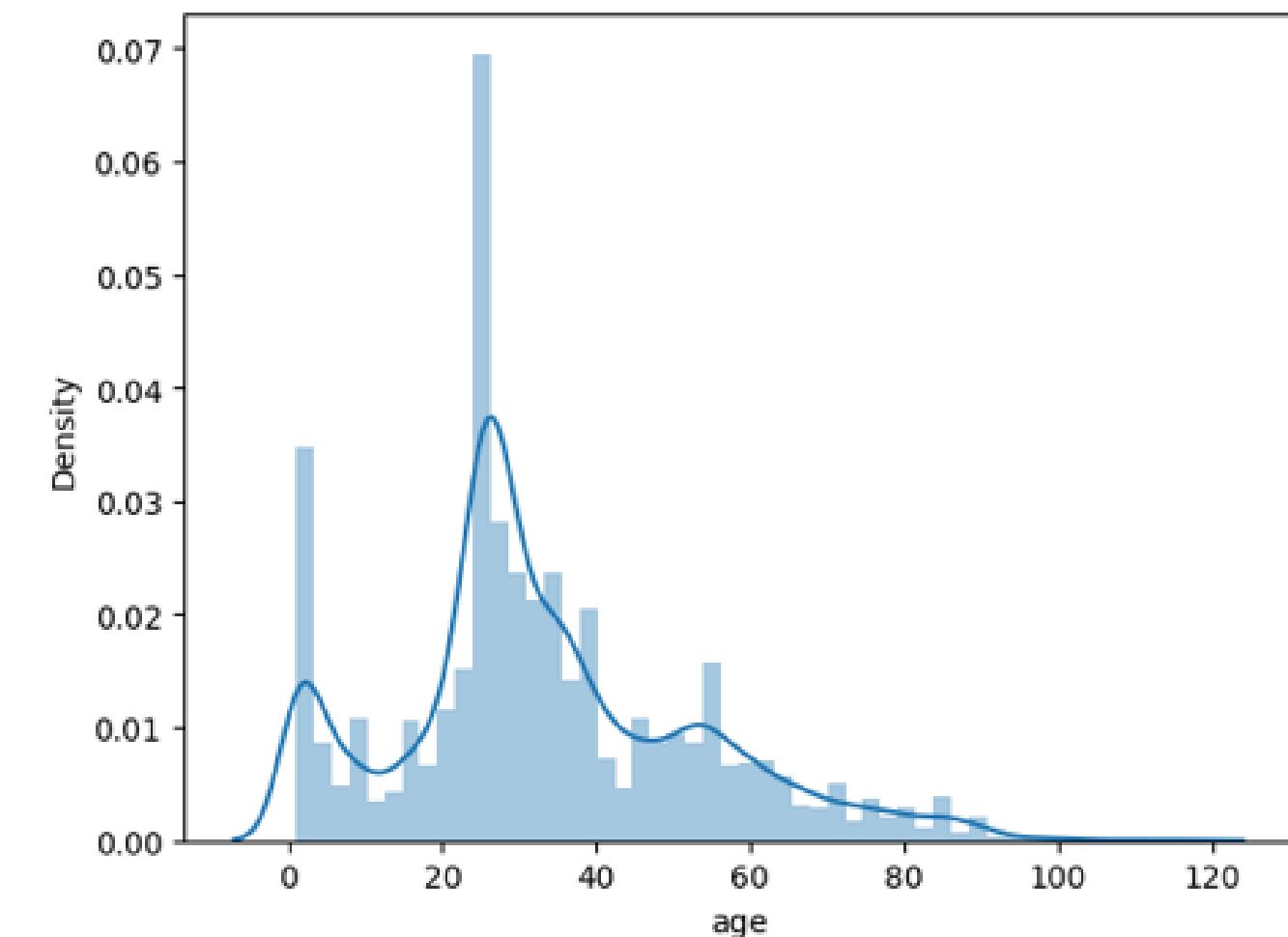
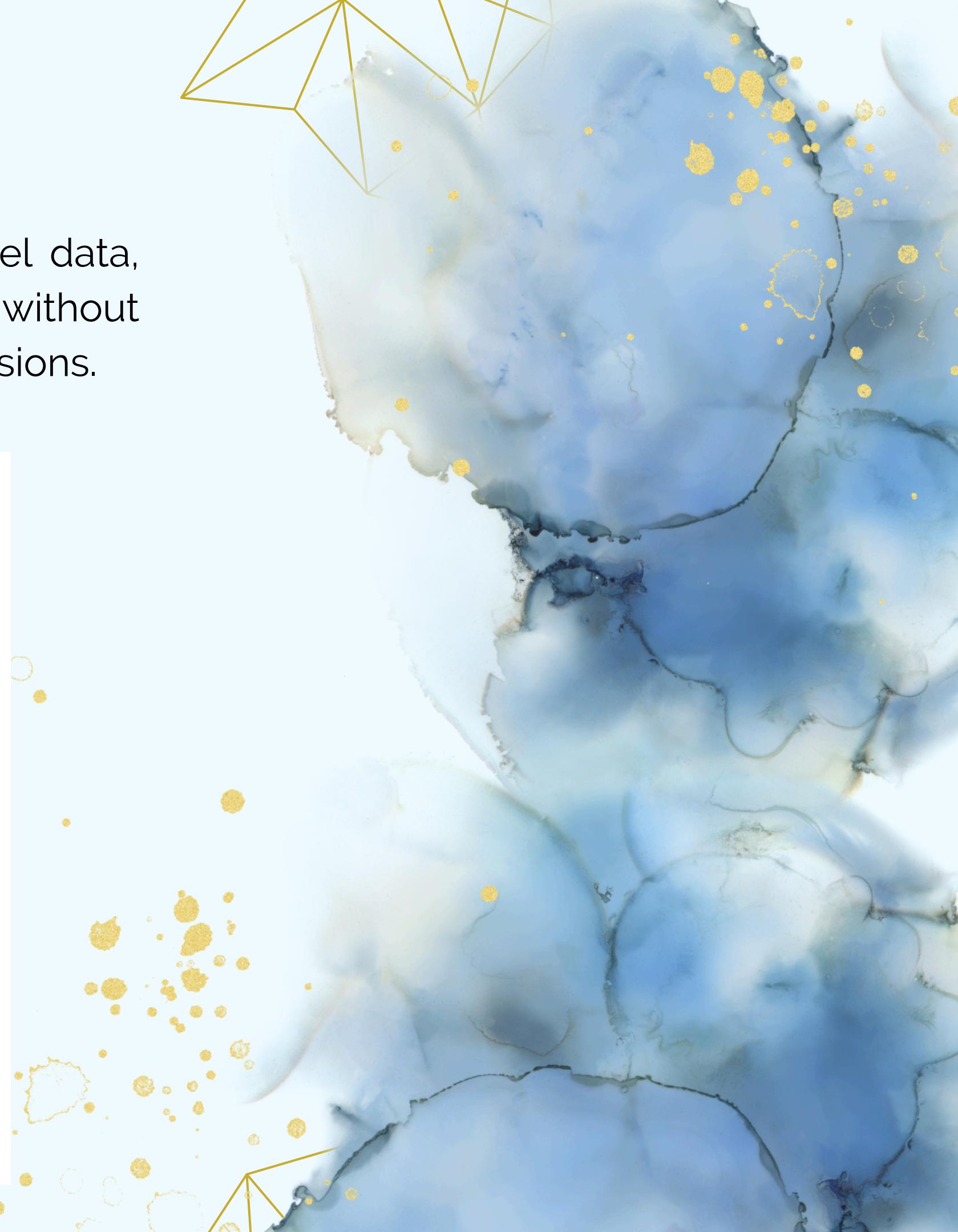
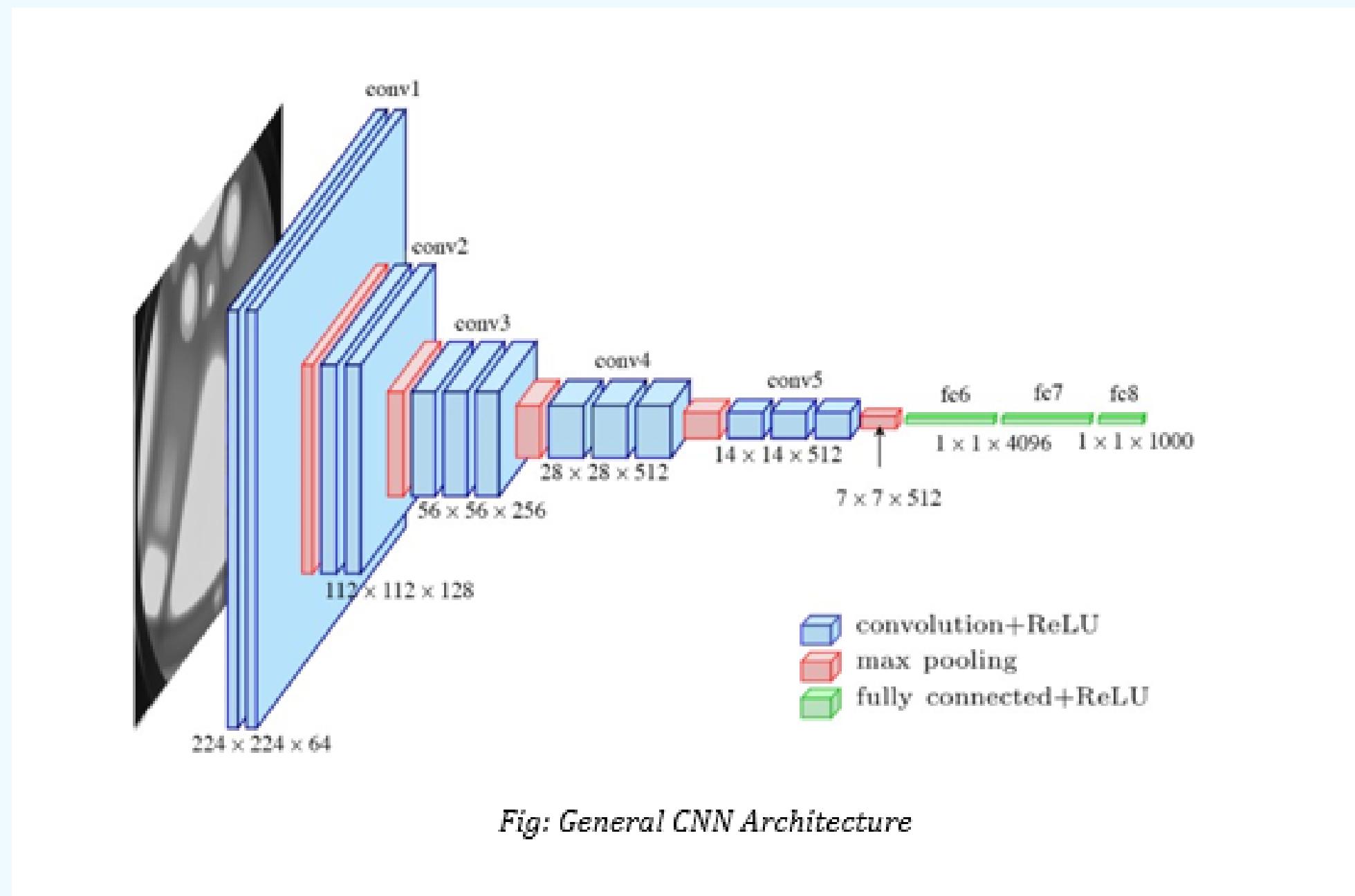


Image Preprocessing

- **Resizing images** to 128x128 pixels ensures uniformity, standardized input size, and enhances feature learning during feature extraction.
- **Normalization scales** pixel values to [0, 1] range, ensuring consistency, aiding model convergence, and improving generalization during training.
- **Grayscale** conversion simplifies data, preserves facial features, and enhances model performance by focusing on structural aspects in age and gender prediction from images.

Model Building

CNNs excel in emotion recognition by learning from raw pixel data, capturing spatial hierarchies, and handling various expressions without manual feature engineering. Ideal for diverse emotional expressions.



Experimental Result

Comparison Analysis

	Accuracy
Prediction of emotion using pre-processing	66.4%
Prediction of emotion using gaussian blur and equalized histogram	74.4%
Prediction of emotion using filter and CNN model	80.6%
Prediction of age and gender	84% and 94%
Prediction of song suggestion based on emotion	94.6%

Table: Prediction of different techniques with their accuracy

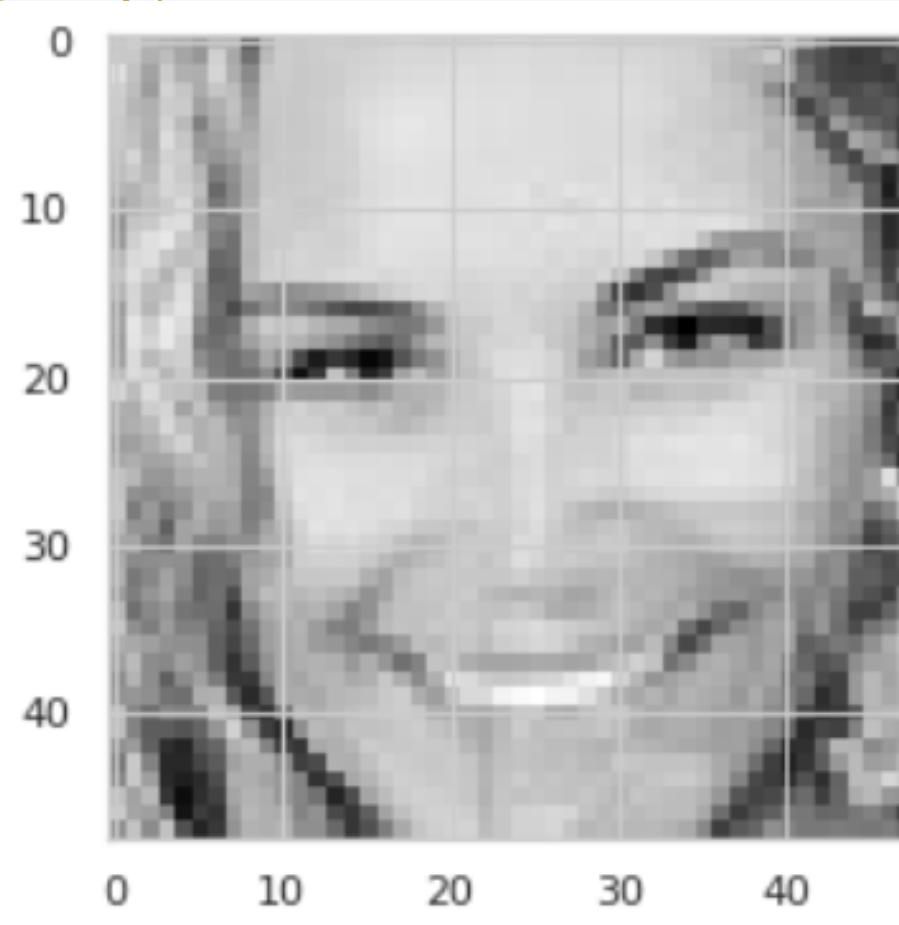
Predicted emotion: Fear



Predicted emotion: Sad

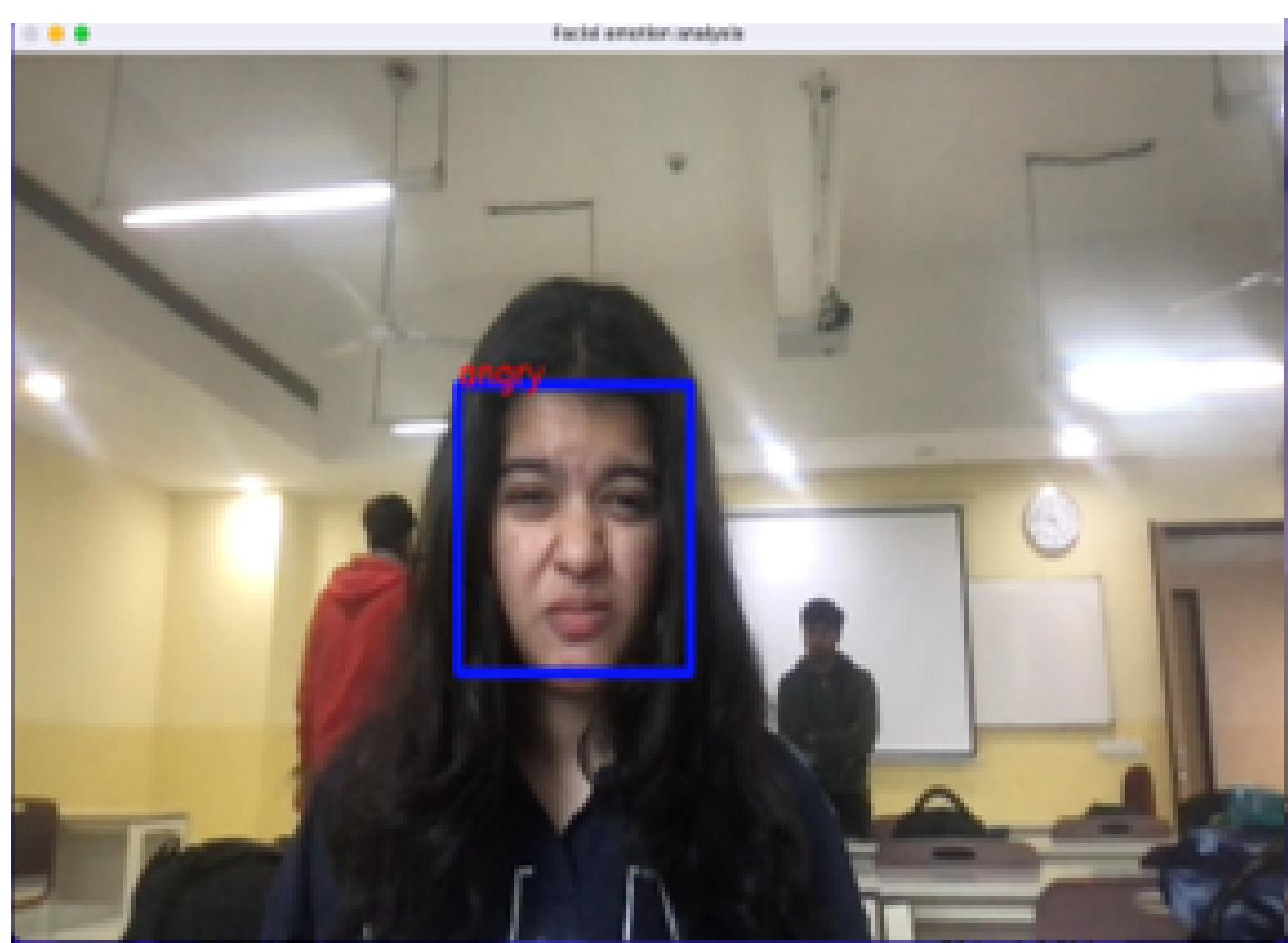


Original Gender: Male Original Age: 41
1/1 [=====] - 0s 229ms/step
Predicted Gender: Male Predicted Age: 42

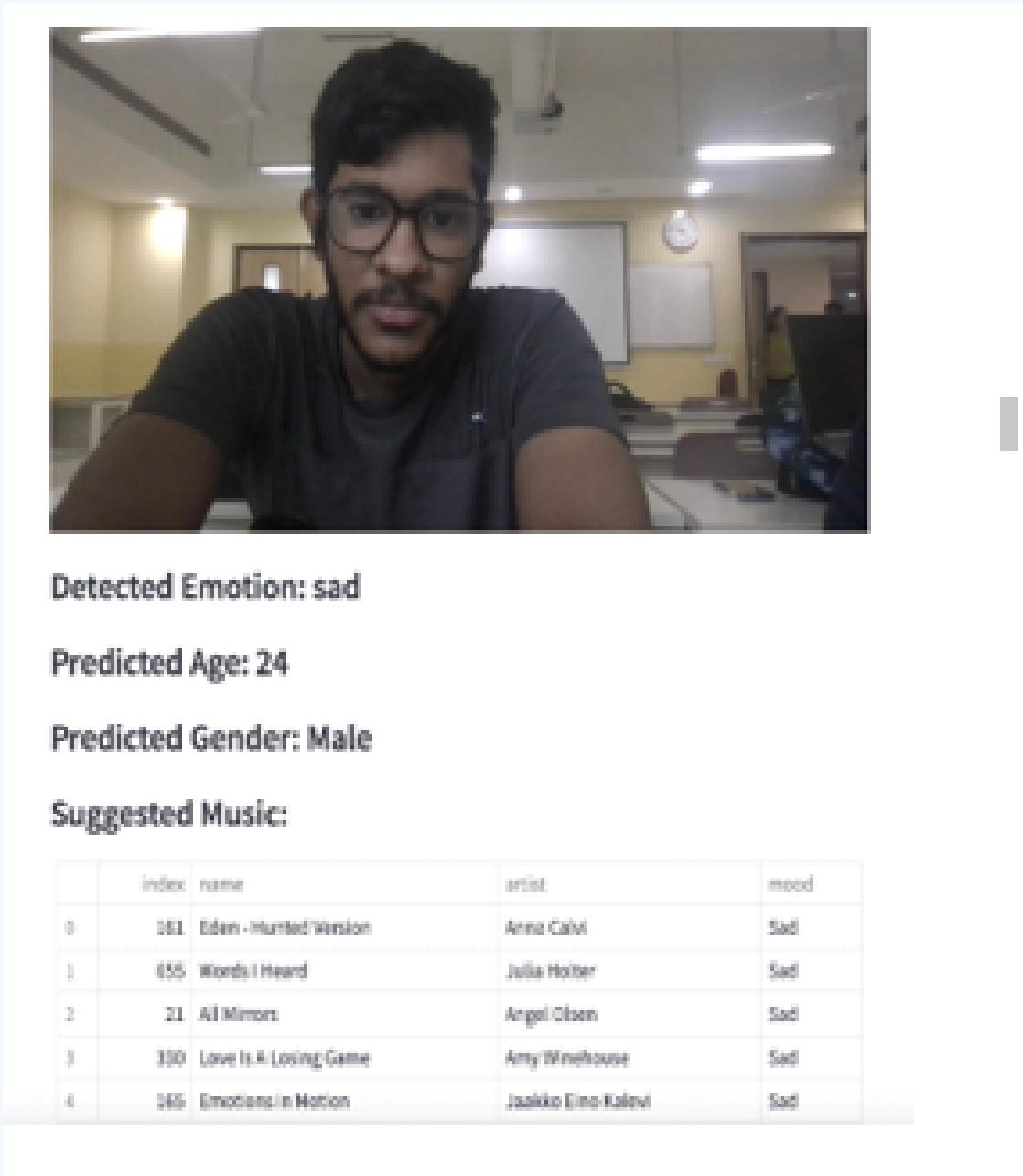


index		name	artist	mood
0	594	Underwater - Tinlicker Remix	Gabriel & Dresden	Happy
1	246	Humans (Let Me Love You)	Sick Individuals	Happy
2	136	Deep Dark Jungle	Nicky Romero	Happy
3	514	Streets	Kensington	Happy
4	587	Tubthumping	Chumbawamba	Happy

Deployment



Here we have used camera to take video of the user and predicted what is the emotion of the user.



Detected Emotion: sad
Predicted Age: 24
Predicted Gender: Male
Suggested Music:

	index	name	artist	mood
0	001	Eden - Hunted Version	Anna Calvi	Sad
1	055	Words I Heard	Julia Holter	Sad
2	21	All Mirrors	Angel Olsen	Sad
3	300	Love Is A Losing Game	Amy Winehouse	Sad
4	365	Emotions in Motion	Jasika Neo-Kalevi	Sad

Here we have used camera to take picture of the user and predicted what is the emotion of the user, age ,gender and the song reccomendation based on his emotion so that it can help the user to calm.



A decorative background featuring a large, horizontal, abstract blue wash at the top, resembling ink or paint. Below it, there are several gold-colored geometric shapes: two large, irregular hexagons on the left and right sides, and two smaller, more defined hexagons in the center. The background is also scattered with numerous small, gold-colored circles of varying sizes.

THANK YOU