Emotion Recognition for Alexithymia: CNN Approach

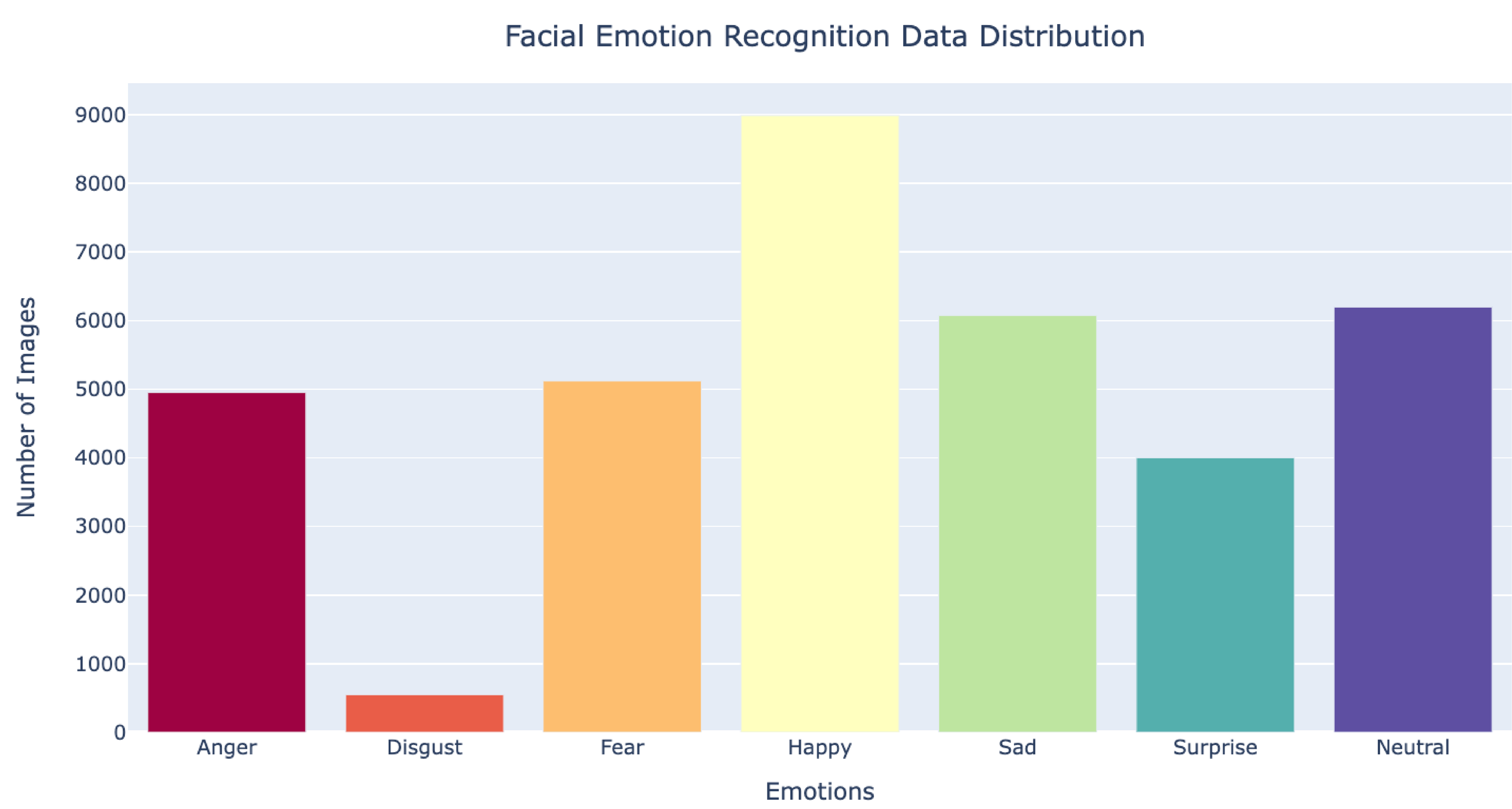
Authors
Clementine Surya, 22200226 (clementine.surya@ucdconnect.ie)
Liu Ye, 22200868 (ye.liu1@ucdconnect.ie)
ACM40960 Projects in Maths Modelling, MSc Data and Computational Science, University College Dublin

MOTIVATION

- To provide assistance to individuals with Alexithymia, helping them overcome challenges in recognizing and expressing emotions by creating real-time Facial Emotion Recognition (FER) through webcam video streams.
- To explore the diverse applications of FER such as personalized services, customer behavior analysis, healthcare, employment, education, public safety, and crime detection.

DATASET

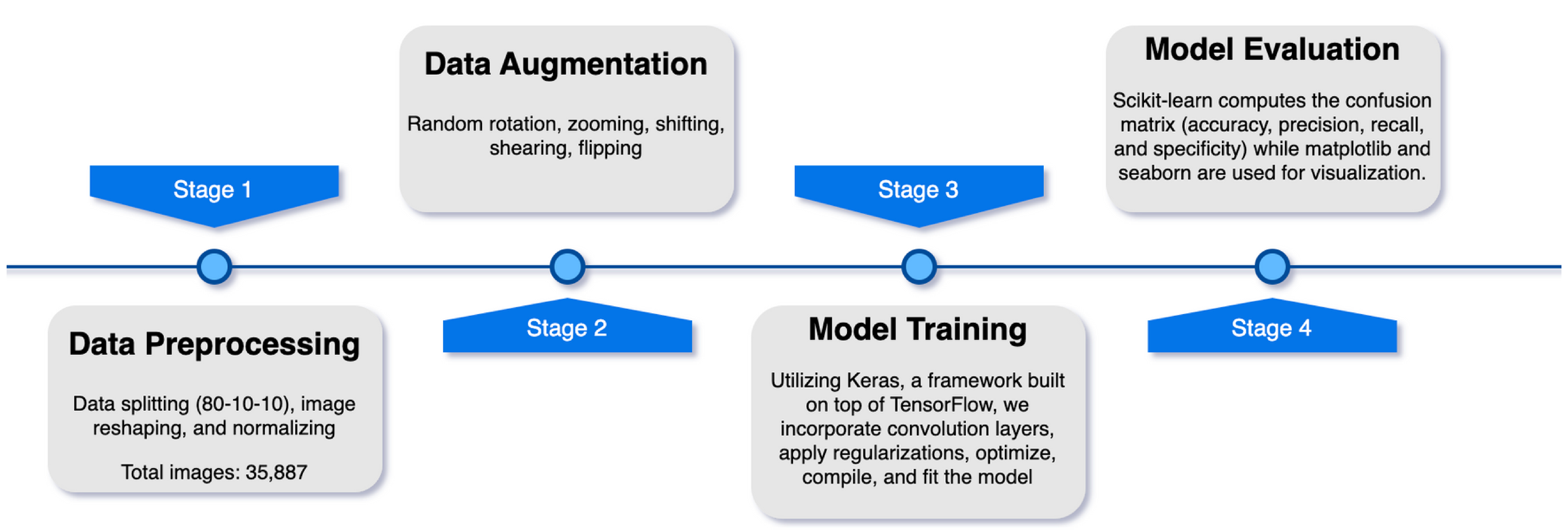
FER2013: comprises 35,887 grayscale images with 48x48 pixels, distributed among 7 different classes: "Anger," "Disgust," "Fear," "Happy," "Sad," "Surprise," and "Neutral."



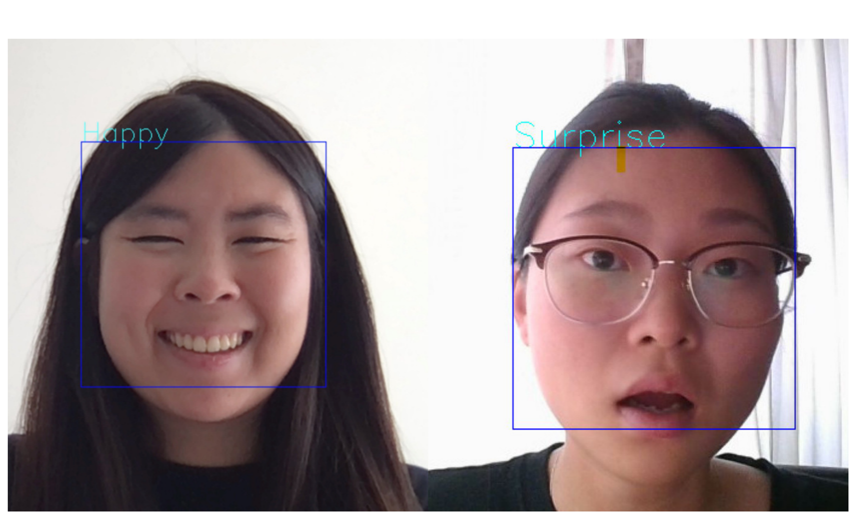
The classes show imbalance, 'Happy' being the dominant and 'Disgust' being the minority

METHODOLOGY

Using **Python** as the programming language, the model is implemented. The entire model is simulated in the **Jupyter Notebook**.



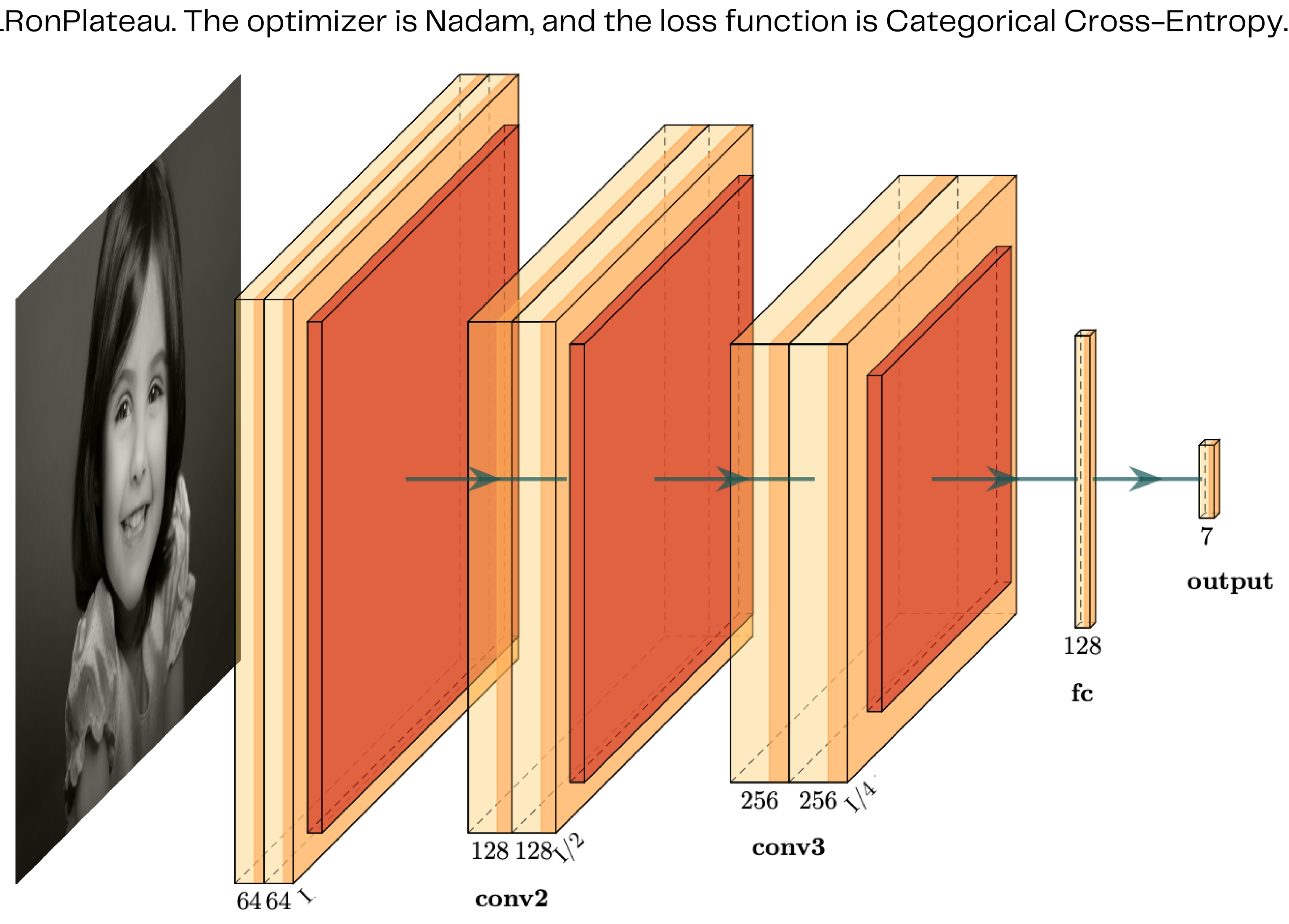
For real-time facial expression recognition, we employed the **Haar Cascade classifier**, a feature-based object detection algorithm, implemented through **OpenCV**. This approach allowed us to detect faces in live video streams from the laptop's webcam. Subsequently, the CNN model was applied to recognize facial expressions in real-time.



Screenshot of the outcome from our webcam's facial emotion recognition model

MODEL

Three blocks with 2 convolutional layers, BatchNormalization, MaxPooling, and Dropout (0.4-0.6), followed by a 128-unit FC layer and a softmax layer. Convolutional layers in each block use 64, 128, and 256 filters of size 3x3. MaxPooling layers have 2x2 kernels. Training involves a batch size of 32 for 100 epochs. ReLU activation and HeNormal kernel initializer are used. Callbacks include Early Stopping & ReduceLROnPlateau. The optimizer is Nadam, and the loss function is Categorical Cross-Entropy.



[1] People with Alexithymia have difficulties recognizing and communicating their own emotions, and they also struggle to recognize and respond to emotions in others. (Source: MedicalNewsToday)

ABSTRACT

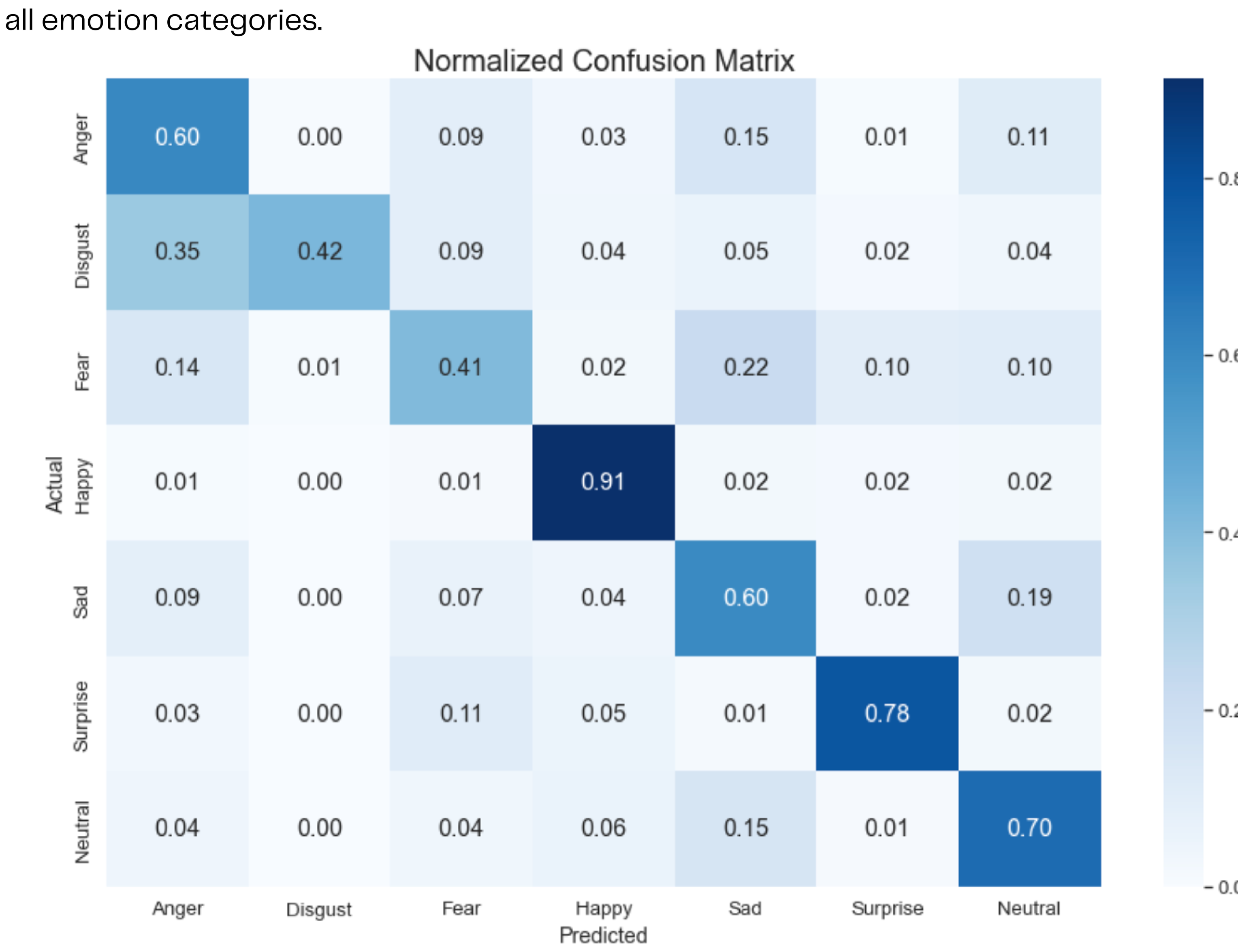
The advancement of Artificial Intelligence has significantly contributed to the field of technology. Among its successes, machine learning and deep learning algorithms have been widely applied in various areas, including classification systems, recommendation systems, and pattern recognition. Human emotions play a crucial role in shaping thoughts, behaviors, and feelings. This project aims to benefit individuals experiencing Alexithymia.[1] With this thought, we have developed a Convolutional Neural Network (CNN) model capable of classifying seven distinct human facial emotions. Through the training and testing, the model achieves an accuracy level of approximately **68.24%.**

ANALYSIS & DISCUSSION

Training & Validation Loss and Accuracy - Gradually improving, and it stops at epoch 52 because of the Early Stopping callback. By that point, the training and validation accuracy are around 70%, and the training and loss accuracy are about 0.9.

Confusion Matrix

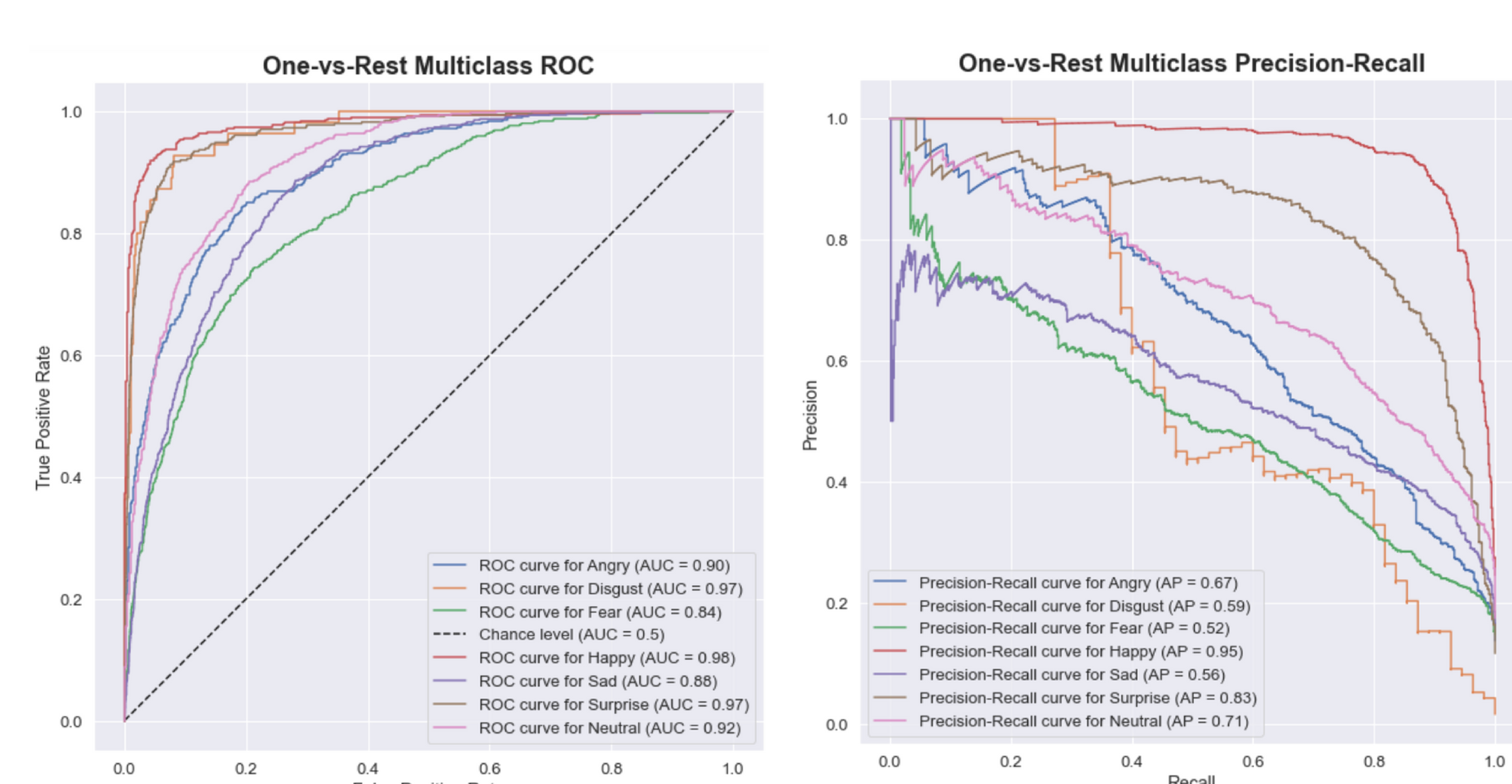
Disgust images frequently predicted as Anger, with 19 instances of misclassification across all images. This might be because we have a smaller set of data for this category. Notably, Happy demonstrated exceptional classification performance, with 801 accurate predictions across all images, the highest among all emotion categories.



Classification Performance Metrics

The model's performance on the test set achieves an overall accuracy of approximately 68.24%. Given the highly imbalanced class, evaluating the model through metrics such as F1-score and Precision-Recall (PR) curve become more appropriate as they focus on the positive class. Notably, we can see that negative emotions such as "Angry", "Disgust", "Fear", and "Sad" have lower F1-score and PR curve value than positive or neutral emotions with "Fear" having the lowest score. Looking at the images in the dataset again, it's tough for even people to tell the difference between "Fear" and other emotions like "Anger" or being "Sad" or to distinguish between "Disgust" and "Anger". This is also true in real life - **detecting the negative emotion is challenging.**

Classes	Precision	Sensitivity (Recall)	Specificity	F1-Score	Accuracy
0 - Angry	0.604	0.603	0.937	0.603	0.892
1 - Disgust	0.719	0.418	0.997	0.529	0.989
2 - Fear	0.563	0.405	0.946	0.471	0.866
3 - Happy	0.877	0.911	0.959	0.894	0.947
4 - Sad	0.529	0.598	0.894	0.561	0.845
5 - Surprise	0.778	0.776	0.971	0.777	0.948
6 - Neutral	0.635	0.698	0.915	0.665	0.877



ROC curve misleads due to highly imbalance dataset; PR curve is realistic, focusing on positive class.

FUTURE WORK

Exploring transfer learning methods with pre-trained models, facial landmark alignment, additional data augmentation, addressing class imbalance and expanding the dataset to include more varied examples could improve the model's classification capabilities. Additionally, to investigate human performance in emotion detection and compare it with the outcomes of the CNN model.

RELATED LITERATURE

- A survey on facial emotion recognition techinques: A state-of-art literature review. Author: Felipe Zago Canal, Tobias Rossi Mu¨ller, Jhennifer Cristine Matias, Gustavo Gino Scotton, 2022, Information Sciences
- Extended deep neural network for facial emotion recognition. Author: Deepak Kumar Jain, Pourya Shamsolmoali, Paramjit Sehdev, 2019, Pattern Recognition Letters.
- CERN: Compact facial expression recognition net. Author: Darshan Gera, S.Balasubramanian, Anwesh Jami,2022, Pattern Recognition Letters.