COMP 6721 Applied Artificial Intelligence (Fall 2023)

Project: AI-Ducation Analytics (Part 1)

Team Name: AK_7

Team members details:

Name	Student ID	Specialization
Gowtham Nalluri	40262135	Data Specialist
Protim Ghosh	40185075	Training Specialist
Md Sakib Ullah Sourav	40264066	Evaluation Specialist

Github link of the project:

https://github.com/Sakibsourav019/AI-Ducation-Analytics-COMP-6721---AK_7-

Contents

1. Dataset

We used two datasets for this project. The first primary dataset [1] we adopted for the purpose of accomplishing this project is taken from Kaggle. The original dataset is composed of images including two sections: training and validation. The entire dataset has seven classes, namely, angry, disgust, fear, happy, neutral, sad and surprise.

Another dataset [2] we used to create our custom dataset which is also taken from Kaggle.This dataset also has seven classes but different than in [1], namely, anger, contempt, disgust, fear, happy, sadness and surprise. But this dataset does not have any subparts like the first one. It only corresponds to 981 images in total for all the seven classes.

As both of the datasets have limitations for certain image classes of our interest and to make our model more efficient, we plan to feed and train it with more diverse data. Which is why we used a python script to scrape images of our desired classes from the internet.

1.1. Overview of the datasets

The first dataset [1] is quite imbalanced when it comes to the number of images per class. From Figure 1 below, we can see that the class "disgust" has a very low number of images in comparison to other classes while "happy" has the highest number of images.

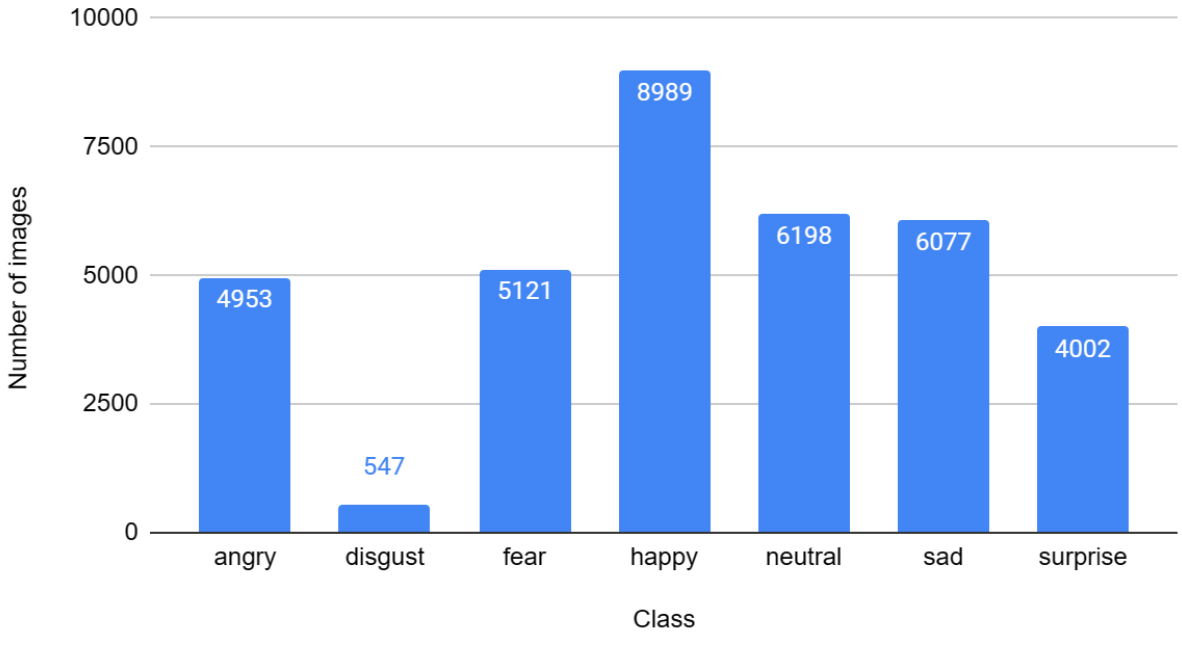

Figure 1: Number of Images vs Class in Dataset 1 [1].

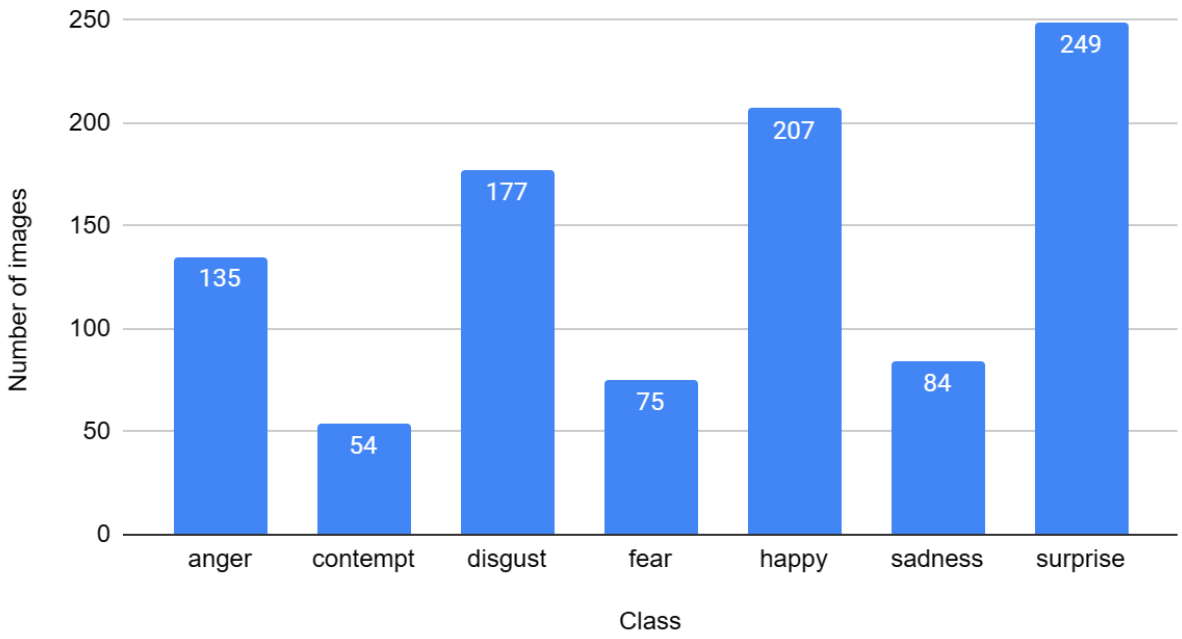

Figure 2: Number of Images vs Class in Dataset 2.

From the dataset 2, it can also be seen that the same trend as the dataset 1 is consistent here. Overall, we can conclude that both the datasets are quite imbalanced.

As mentioned above, we scraped images of all four classes of our interest from the internet using a script that uses Bing search engine to collect images. Thus we collected 39 “bored” images, 33 “neutral” images, 36 “engaged” images and 48 “angry” images as shown in Figure 3.

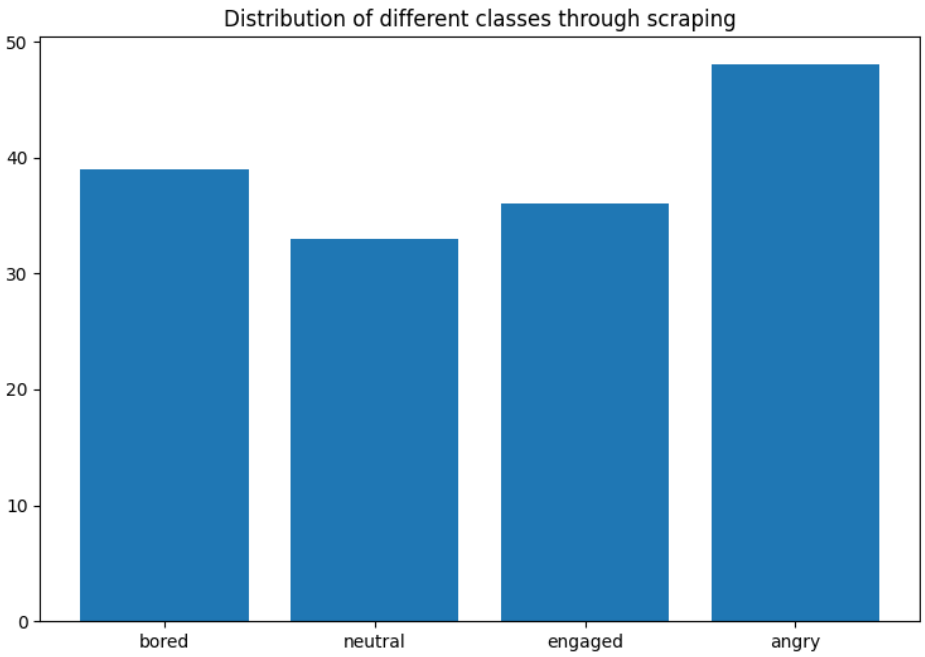

Figure 3: Number of Images vs Class through web scraping.

1.2 Justification for the datasets choice

While searching for the datasets, we came across various sources and options. Yet, the type of image data suitable for our task was not adequate. Moreover, some of the data classes we needed for this project have scarcity in the existing available datasets. Hence, we decided to scrape some images and merge it with the datasets so that we can construct our desired classes by combining those. All the data we adopted have the frontal face shots with no visible scenic background that makes the datasets well equipped for our desired goal.

1.3 Provenance information

Below in table 1, we list the informations about the datasets we used in this project-

Table 1: Overview of the two primary datasets that we adopted in this work

	Dataset 1	**Dataset 2**	**Web Scraping**
Name	CKPLUS	Face expression recognition dataset	Bing Search Engine Script
Number of images	35887 (Training - 28821, Validation - 7066)	981	156
Data classes	angry, disgust, fear, happy, neutral, sad, surprise.	anger, contempt, disgust, fear, happy, sadness and surprise	Neutral, engaged, bored, angry
Nature of images	Frontal face shots	Frontal face shots	Frontal face shots
Authors	Unknown	Unknown	Unknown
License	CC0: Public Domain	Unknown	NA
Links	https://www.kaggle.com/datasets/shawon10/ckplus/data	https://www.kaggle.com/datasets/jonathanoheix/face-expression-recognition-dataset	NA

In table 2, we will get the overview of our customly developed dataset for this project out of the two primary datasets mentioned above.

Table 2: Overview of the dataset we developed for this project

Number of images	**Image sources**	**Image classes**
Training - 1600 Testing - 600	1. Training - dataset 1 [1] 2. Validation - dataset 2 [2] 3. Web Images through scripts	Neutral, engaged/focused, bored/tired, angry/irritated

2. Data Cleaning

All the images in the datasets [1,2] are in grayscale color and in 224 x 224 size.To comply, we have resized the scraped images into 224 x 224 too.

Below, we can see a demonstration of an “angry” image that we scraped and resized the original rgb image into a 224 x 224 grayscale image.

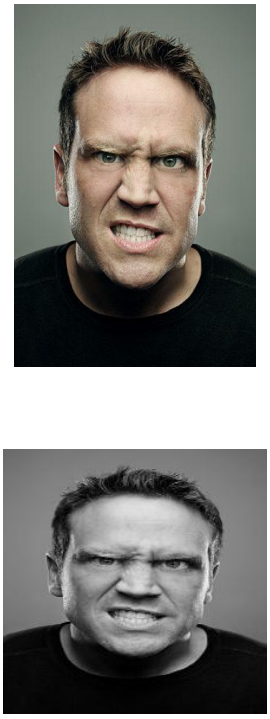

Figure 4. Before and after image resize operation.

Apart from that, the datasets we used as primary sources have more images in total than we need. We operated an in-depth manual eye-skimmed data cleaning process so that we can include most deserving images to the corresponding classes. Hence, we believe the model can get good training data and the model can perform better in the next parts of this project.

3. Labeling

As we mentioned earlier, two datasets [1, 2] have been utilized to define the data classes we need. However, among the four data classes of our interests, two of them named “neutral” and “angry” are common in every dataset. So, we picked the best matched images for our model. But the other two classes “bored/tired” and engaged/focused” were unusual.

Hence, to label these two classes we skimmed all the images in the dataset and looked for other sources in the internet beyond to connect images that fall in these classes. Finally, we came up with a solution. As the solution, we firstly took help from the python script to download internet images using Bing Image Search through a series of written prompts. And secondly, we skim through the images from the dataset [1] to see which images are best fit to “bored/tired” and "engaged/focused” classes according to the criterias and features of expression mentioned in the project guideline. Thus we labeled the images to these two classes.

4. Dataset Visualization

We used the python function matplotlib and seaborn to show the image data that we constructed for our task.

4.1 Class Distribution

We plotted the class distribution of our training and testing image data below in Figure 5 and 6. Number of training "bored" images are 439, number of training "neutral" images are 433, number of training "engaged" images are 436, and number of training "angry" images are 448.

In the same way, the number of test "bored", "neutral", "engaged" and "angry" images each are of 150.

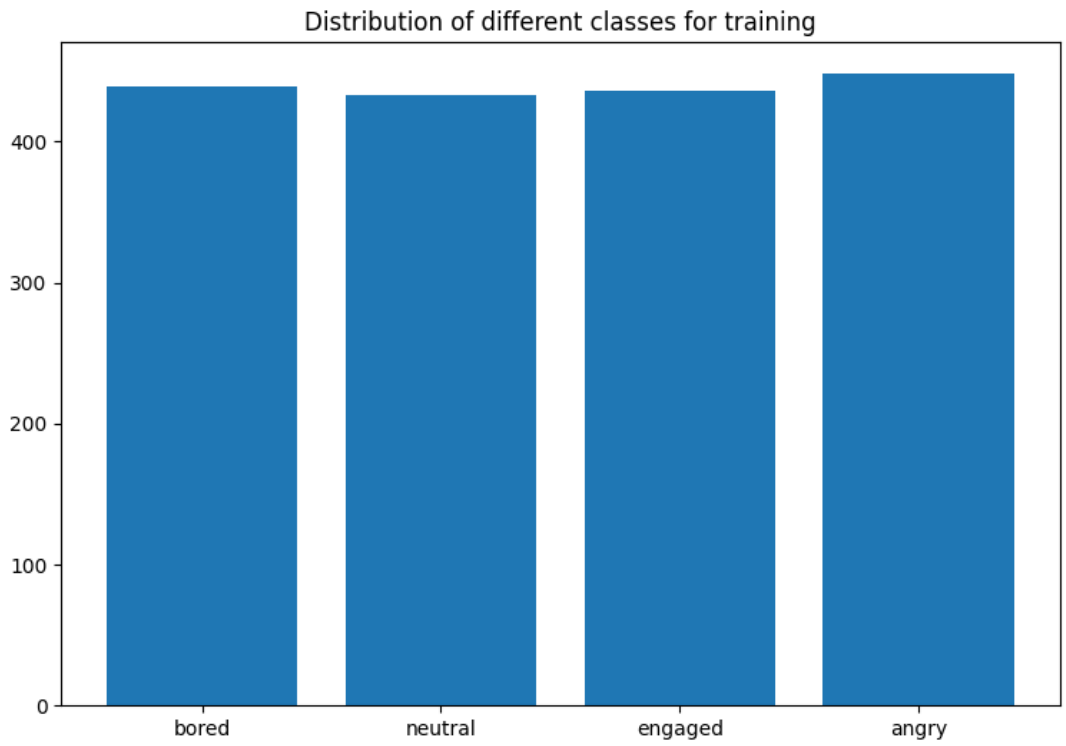

Figure 5. The training data images of four classes.

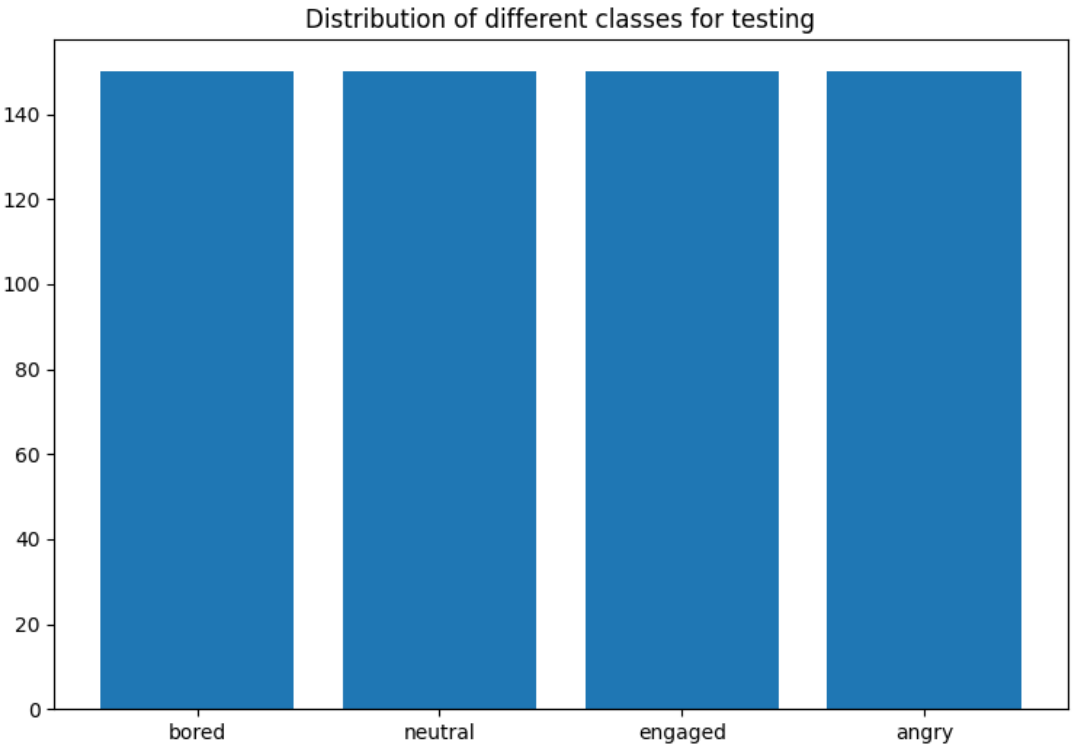

Figure 6. The testing data images of four classes.

4.2 Sample Images

As described and asked in the project guideline, below we listed the outputs of 5x5 grids, each image contains 25 images of four classes one after another.

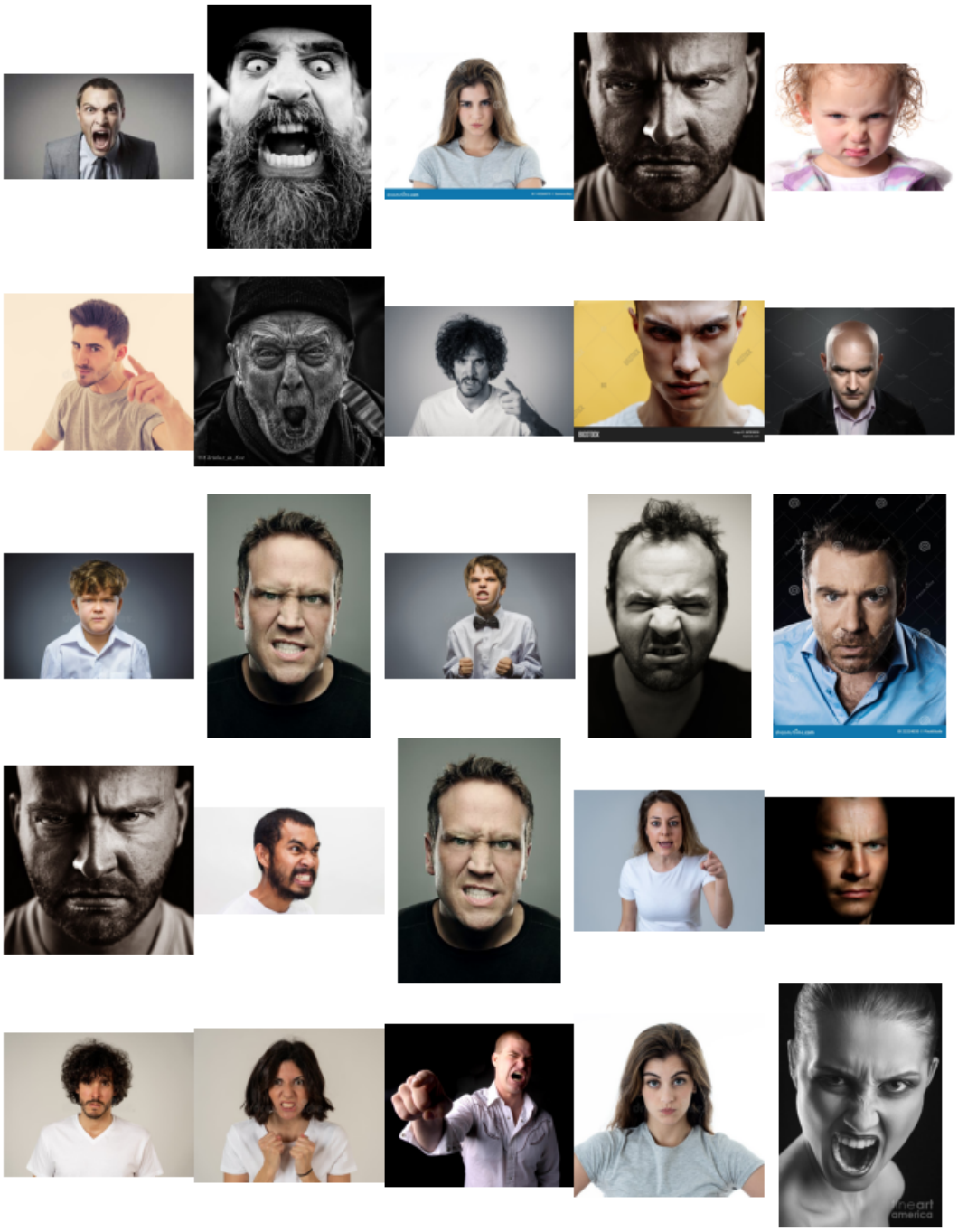

Figure 7. 25 “Angry” images from our dataset in a 5X5 grid.

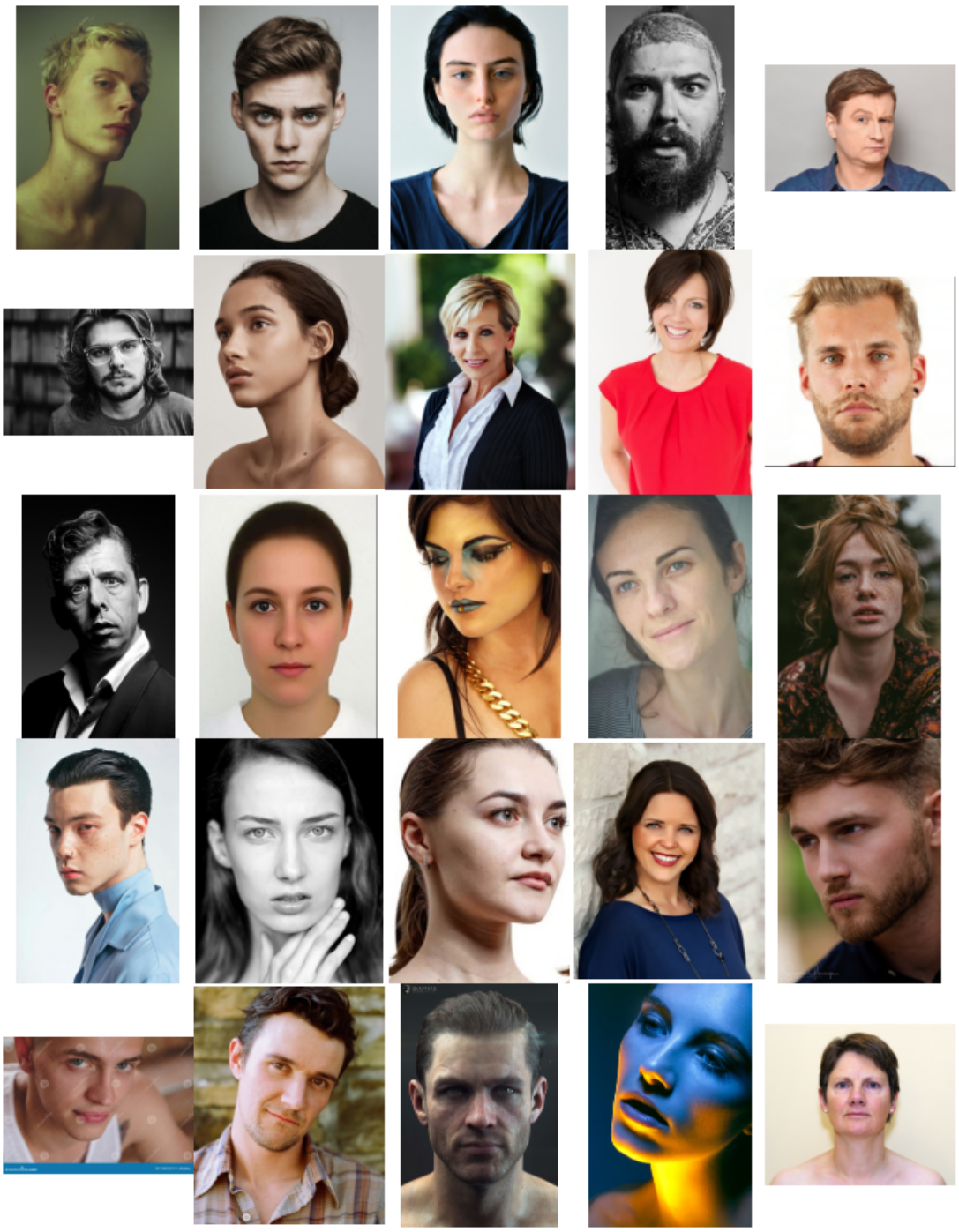

Figure 8. 25 “Engaged” images from our dataset in a 5X5 grid.

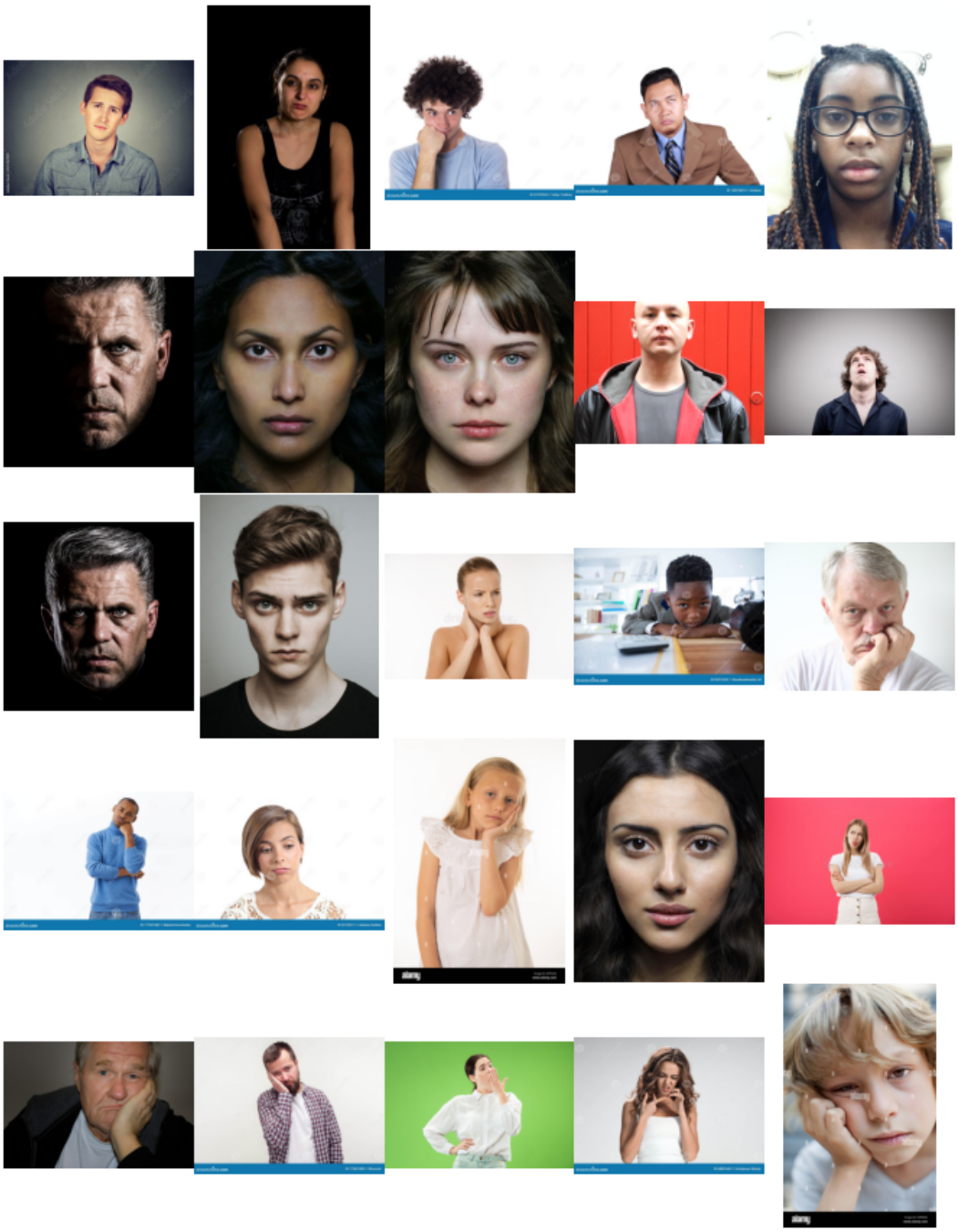

Figure 9. 25 “Bored” images from our dataset in a 5X5 grid.

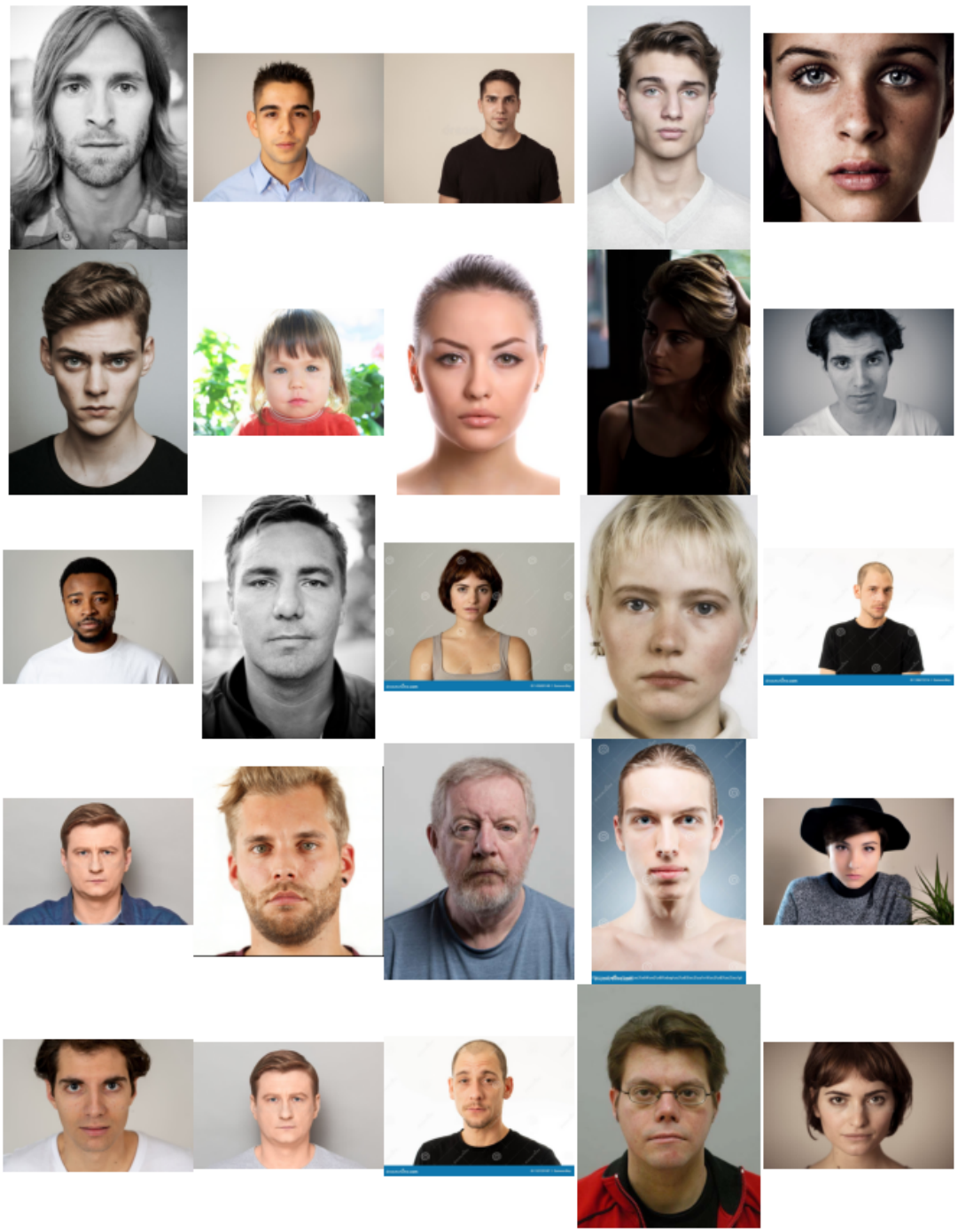

Figure 10. 25 “Neutral” images from our dataset in a 5X5 grid.

4.3 Pixel Intensity Distribution

Below, we have shown the plot that we get from the histogram that shows the pixel intensity distribution of the random images we have shown in section 4.2.

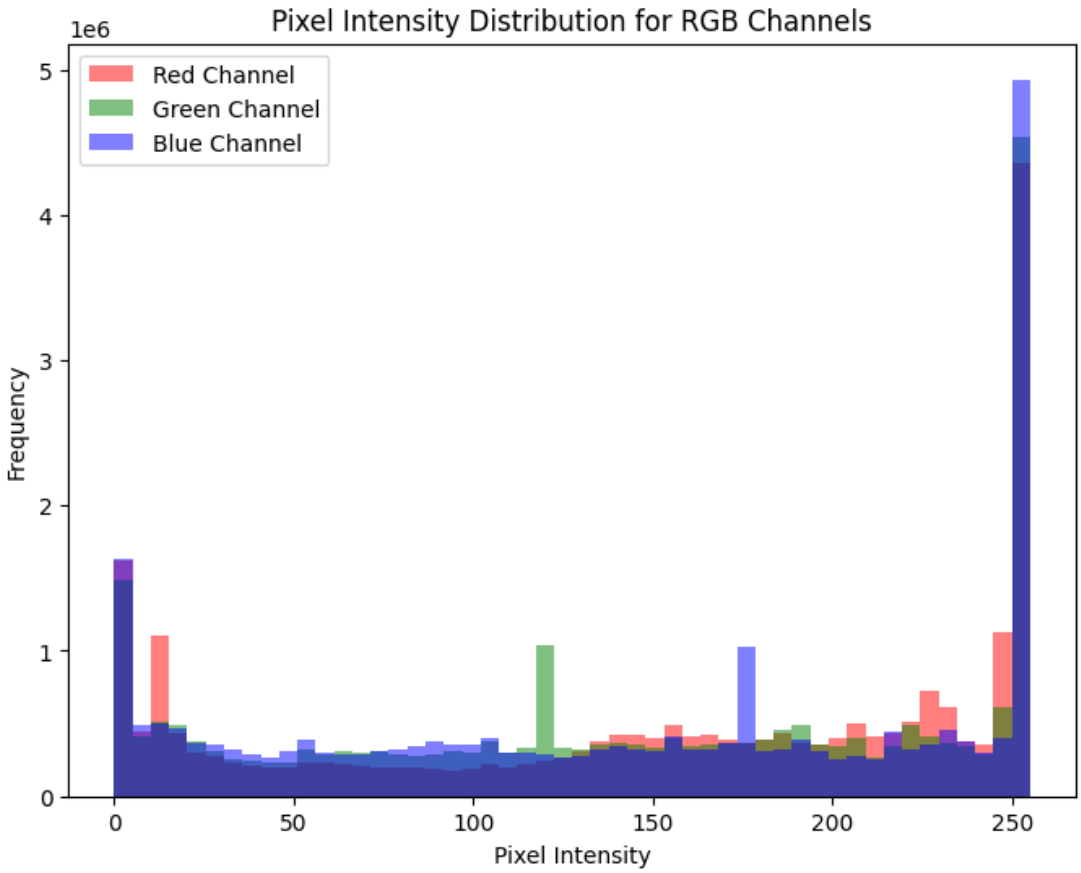

Figure 11. Histogram of random images in RGB color

Reference

[1] “Face expression recognition dataset,” *Kaggle*, Jan. 03, 2019. https://www.kaggle.com/datasets/jonathanoheix/face-expression-recognition-dataset

[2] “CKPLUS,” *Kaggle*, Oct. 16, 2018. https://www.kaggle.com/datasets/shawon10/ckplus/data