COMP 472 Artificial Intelligence

Team Name: OB_03

Team Members:

Gevorg Alaverdyan (#40202177) - Data Specialist

Jay Patel (#40203705) - Training Specialist

Joud Babik (#40031039) - Evaluation Specialist

Github: https://github.com/JRB958/COMP472-PROJECT

We certify that this submission is the original work of members of the group and meets the Faculty's Expectations of Originality

Signatures:

GA (#40202177) 5/28/2024

JB (#40203705) 5/28/2024

JP (#40031039) 5/28/2024

Thursday, 30 May 2024

Table of Contents

Dataset

Overview

The AMFD [1] contains 219 faces (smiling and neutral expression poses) with mixed-race heritage. 109 of the images are neutral while 110 are smiling. Characteristics: All frontal face shots, diverse backgrounds, naive observers.

The Natural Human Face Images for Emotion Recognition [2] has 5500 + images with 8 emotions categories – anger (890 images), contempt (208 images), disgust (439 images), fear (570 images), happiness (1406 images), neutrality (524 images), sadness (746 images) and surprise (775 images). All images contain grayscale human faces (or sketch). Each image is 224 x 224 pixel grayscale in PNG format.

Adobe Stock [3] was also used for gathering images. Since the images were found using a search bar there are no concrete classes. 65 Neutral faces were found using this.
Characteristics: All frontal face shots, diverse backgrounds.

Human Faces dataset [4] was used as a dataset. It contains 7200 images of diverse people (age, race, gender). The images were not labeled.

happy_or_sad_Binary_Image_Classification dataset [5] has 165 images.
Class Distribution: Sad Faces: 73, Happy Faces: 92.
Characteristics: Mostly frontal face shots, diverse backgrounds.

Facial Recognition dataset (Human) dataset [6] comprises images featuring diverse human faces annotated with various emotions like happiness, sadness, anger, neutral, and surprised. It's divided into three directories: test, train, val. Each contains happiness, sadness, anger, neutral, and surprised classes. There are 1823 images in total and 42 happy in the test folder, 327 happy in train and 41 happy in val.

Justification

1. American Multiracial Faces Database (AMFD)

- **Relevance:** The AMFD is particularly useful for projects focusing on the recognition and analysis of mixed-race faces. This dataset addresses the need for diverse facial representations in machine learning models, ensuring inclusivity and reducing biases that might arise from using homogeneous datasets.
- **Unique Features:** Contains 219 faces with mixed-race heritage. Includes both smiling and neutral expressions, providing a range of emotional data. Accompanied by ratings from naive observers, offering additional data for subjective analysis.

- **Challenges:** Limited number of images (219 faces), which might be insufficient for training deep learning models.

2. Natural Human Face Images for Emotion Recognition

- **Relevance:** This dataset is highly suitable for projects aiming to develop or test emotion recognition algorithms. It offers a substantial number of images across eight emotional categories, facilitating comprehensive training and evaluation of models.
- **Unique Features:** Over 5,500 images categorized into eight different emotions (anger, contempt, disgust, fear, happiness, neutrality, sadness, and surprise).
- **Challenges:** Sourced from various online platforms, which might lead to inconsistencies in image quality and style. The dataset's manual annotations might introduce subjectivity, affecting the uniformity of the data.

3. Adobe Stock

- **Relevance:** This tool was used because it was convenient to fill up any dataset that was missing pictures. The search engine is effective and provides corresponding face images
- **Challenges:**Needed manual work to download and search for images one by one.

3. Human Faces Dataset

- **Relevance:** This dataset is ideal for general-purpose facial recognition and classification tasks. It offers a broad mix of faces from different demographics, which is beneficial for creating unbiased models.
- **Unique Features:** Contains over 7,200 images covering diverse age groups, races, and profiles.
- **Challenges:** The presence of GAN-generated images might complicate the training process if not properly managed.

4. happy_or_sad_Binary_Image_Classification

- **Relevance:** This dataset is specifically designed for binary emotion classification, making it an excellent choice for projects focused on distinguishing between happy and sad expressions.
- **Unique Features:** Contains 165 images divided into two classes: sad and happy faces.
- **Challenges:** Limited size (only 165 images), which might not be sufficient for training robust models. Binary classification restricts the scope to only two emotions, limiting broader application.

5. Facial Recognition Dataset (Human)

- **Relevance:** This dataset is suitable for projects that require a variety of emotional annotations, offering a balanced mix of expressions.
- **Unique Features:** Includes images annotated with multiple emotions such as happiness, sadness, anger, neutral, and surprise.

- **Challenges:** Standard resolution might not be high enough for certain detailed facial analysis tasks.

Provenance Information

Dataset	Source (see References)	License
American Multiracial Faces Database (AMFD)	[1]	Creative Commons Attribution 4.0 International Public License
Natural Human Face Images for Emotion Recognition	[2]	Public Domain/Various online sources
Adobe Stock	[3]	Adobe Stock License
Human Faces Dataset	[4]	Public Domain/Various online sources
Happy_or_sad_Binary_Image _Classification	[5]	Public Domain/Various online sources
Facial Recognition Dataset (Human)	[6]	Public Domain/Various online sources

* For complete reference of all images we created an excel:

https://docs.google.com/spreadsheets/d/1hsD1Ln5UQay_Z60LH0DsVyfRwsQEbv89/edit?usp=sharing&ouid=104831330757010305938&rtpof=true&sd=true

Data Cleaning

Amongst all the datasets, imperfections with most images were noticed. Techniques such as image resizing and the utilization of scripts were used to facilitate the procedure for data cleaning. Throughout the process, we noticed imperfections with some of the images we went through. In particular, we dealt with 3 challenges that needed to be addressed for data cleaning as shown below.

- **Images less than 10kb**: Many images were under 10kb, which is generally not ideal due to poor resolution and increased noise. Ideally, we would want our CNN model to be trained on qualitative data that well portrays each of the facial emotions we would be dealing with. This issue was easily resolved by sorting the storage size of the images from low to high, allowing us to focus only on images above 10kb.

- **Inconsistent sizing**: Another deficiency between all images was that they did not conform to a standardized size; images of different size formats do not contribute to a consistent input to our training model. As such, we would want our images to be of the same size to prevent preprocessing overhead. Although manually resizing each image was an option, this is a very naive approach as it would take a significant amount of time to complete. Instead, we used the following script which iterates through all images and resizes them with 224x224 dimensions.

- **Watermarks**: Although this was not an issue with all images, some were watermarked with a logo, which is not something we want our CNN model to pick up on. Given that very few images were watermarked, we decided to avoid using such images within our final dataset.

Below is an example demonstrating the importance of cleaning images so that they can be used for our training model. Pictures not of the right size put less emphasis on the person's face, which can potentially distort the training model. By cropping the images accordingly, we ensure that the model only focuses on the facial expressions within the image and nothing else.

Figure 1 *Before and after cleaning image process*

Labeling

Dataset labeling consisted of identifying all of the images and labeling them based on one of the four emotions. Given that we will likely be working with supervised learning, all training data must be labeled as either happy, angry, neutral or focus.

As we were not able to find a single dataset that contained 500 images per each of the 4 facial expressions chosen, we had to go through several datasets, merge them together and compile all of the images. This process served to reach the goal of obtaining a net of 2000 images that would be used for our training model. To simplify the process of combining all images across the multiple datasets, we used the following script, which iterates through all images and inserts them within an excel spreadsheet.

Although we were fortunately able to do so successfully, finding datasets of qualitative images posed to be a challenge as there were very few that provided us with what we needed. Extensive search had to be done

Amongst all the datasets, we were fortunate enough for some of them to be already labeled, based on the folders images resided in. For instance, the “Natural Human Face Images for Emotion Recognition” [2] dataset labeled all images by the appropriate emotion and grouped them within separate folders. The data being already labeled in this case made it easier for us to gather all necessary images that would be used as training and testing data.

However, it was not the case for all datasets to be labeled in such a way as described above. Images from such datasets were often unlabeled in any way, resulting in the set of all images being grouped together. This made it challenging for us as it required more time from us to label each of the images we had to go through. In addition to this, it was often hard to distinguish images as either neutral or focused as depending on the facial structure, they can look very similar to one another. To overcome this issue, we looked for key facial expressions that would determine whether they were focused or neutral. Such expressions include placement of eyebrows, degree of squinting eyes, mouth positioning and so on. This allowed us to iterate through all of the unlabeled images and label them ourselves without the use of external platforms or tools.

Dataset Visualization

Class Distribution

The count of all images in a class were counted and Figure 2 presents the bar chart representation of the count

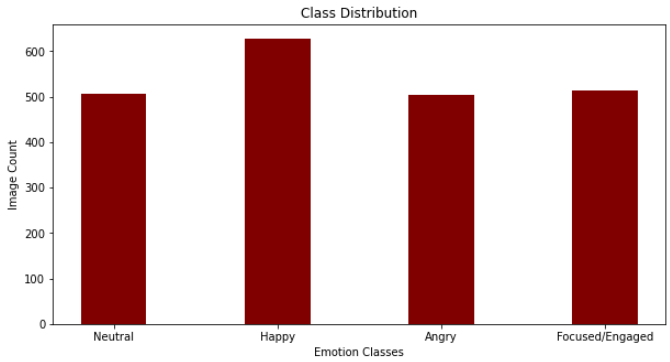

Figure 2 *Class Distribution for all Emotion Class*

Pixel Intensity Distribution

For Pixel Intensity Distribution we use a script to process all the images in a class as follows:

- Mask the images to remove background colors that will skew the histograms to a pixel value that corresponds to it.
- Pre-process 3-channel grayscale to a one channel image.
- Normalize images
- Plot the histogram in 3 channels for RGB or single channel for Grayscale.

Figures (3-10) present the aggregated pixel density distribution for every class in RGB and Grayscale based on the images selected for each class.

Neutral Class

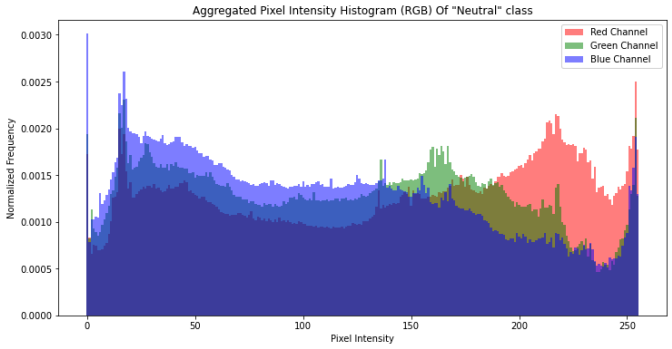

Figure 3 *Aggregated RGB Pixel Intensity Histogram for Neutral Class Images*

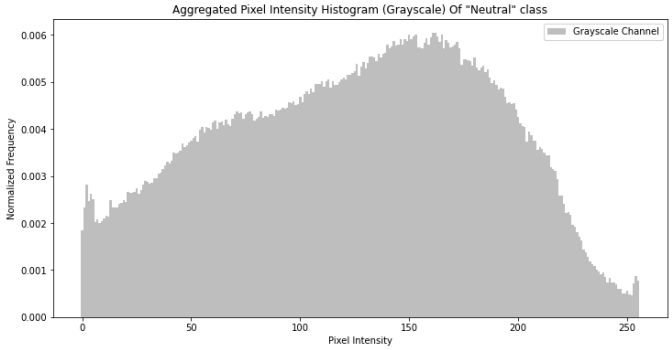

Figure 4 *Aggregated Grayscale Pixel Intensity Histogram for Neutral Class Images*

Angry Class

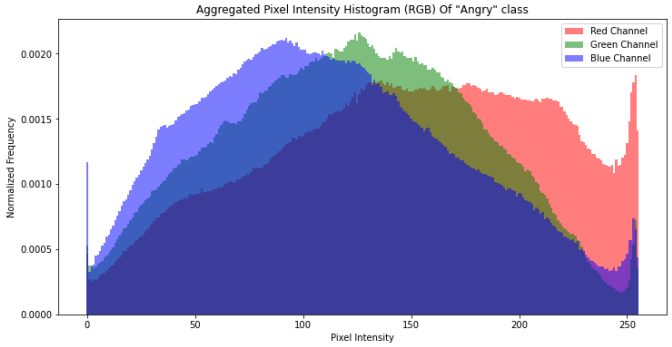

Figure 5 *Aggregated RGB Pixel Intensity Histogram for Angry Class Images*

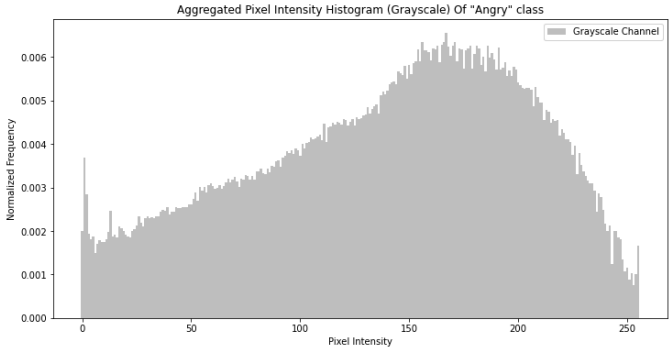

Figure 6 *Aggregated Grayscale Pixel Intensity Histogram for Angry Class Images*

Happy Class

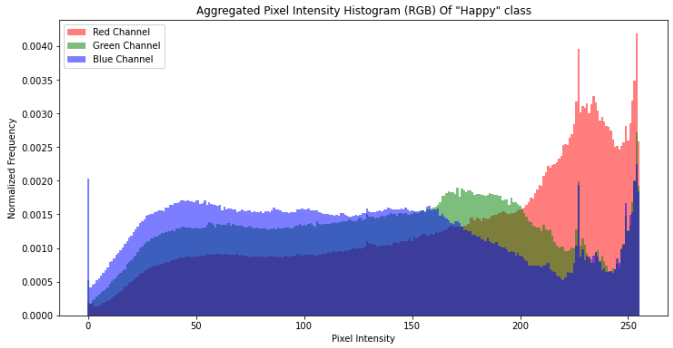

Figure 7 *Aggregated RGB Pixel Intensity Histogram for Happy Class Images*

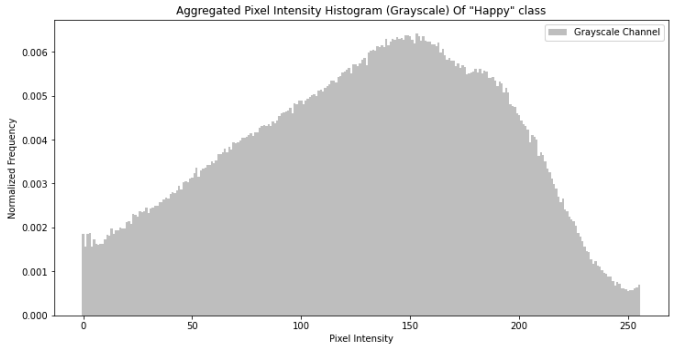

Figure 8 *Aggregated Grayscale Pixel Intensity Histogram for Happy Class Images*

Focused/Engaged Class

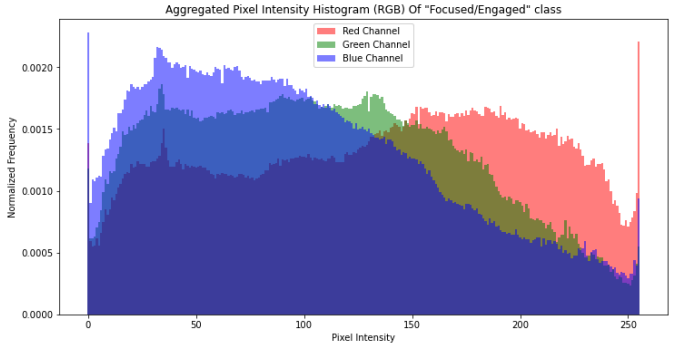

Figure 9 *Aggregated RGB Pixel Intensity Histogram for Focused/Engaged Class Images*

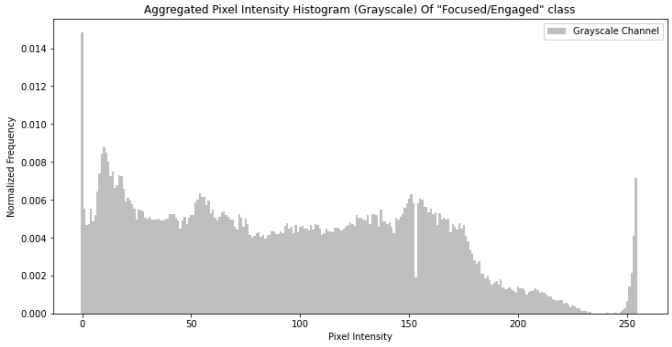

Figure 10 *Aggregated Grayscale Pixel Intensity Histogram for Focused/Engaged Class Images*

Sample Images

We use a script to randomly pick 15 images from the class folder and generate the histogram to be included in the report. Tables (1-4) show 15 sample images for every class with its corresponding masked histogram.

Neutral Class

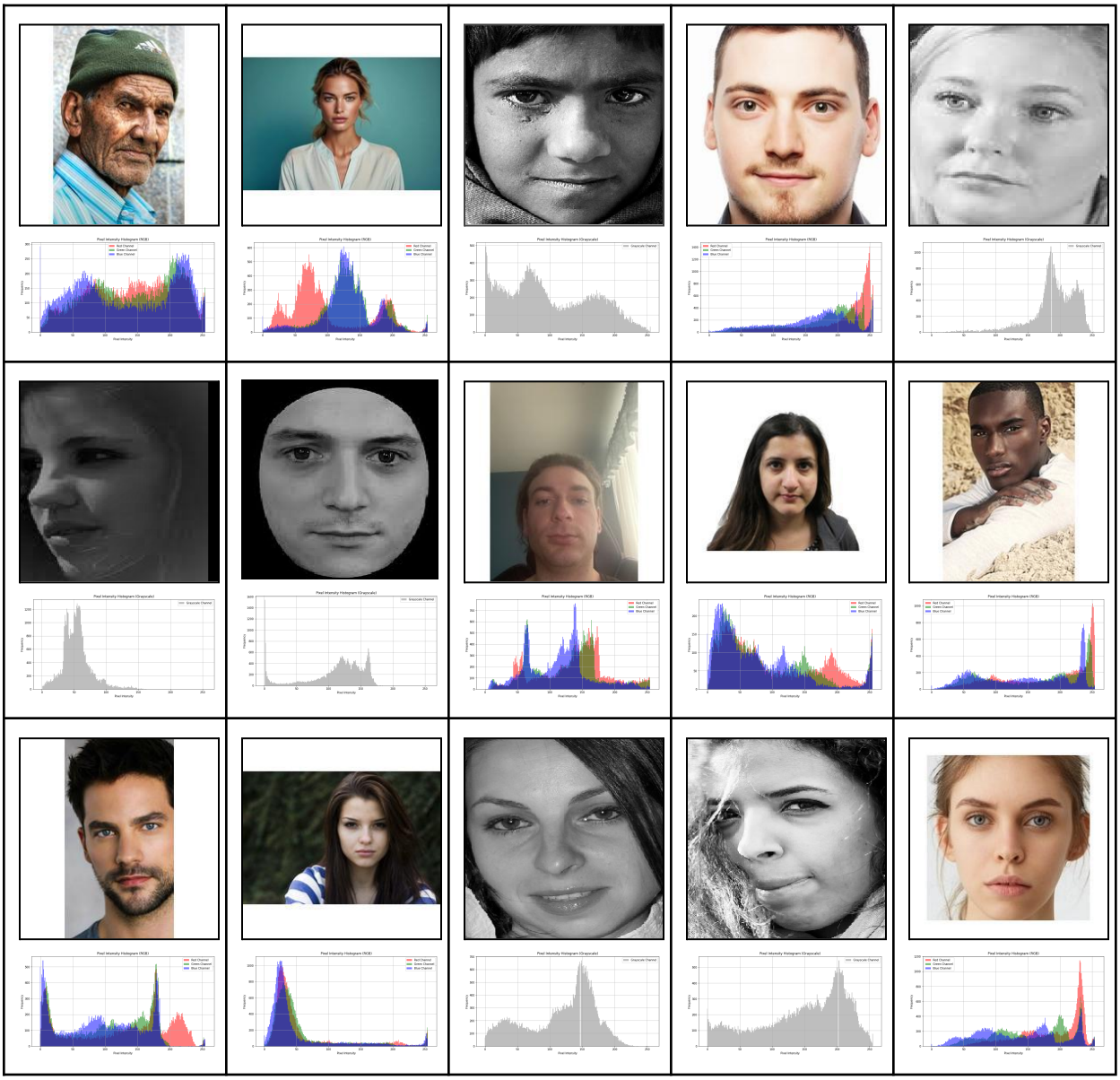

Table 1 *Sample Neutral Images With Histograms*

Angry Class

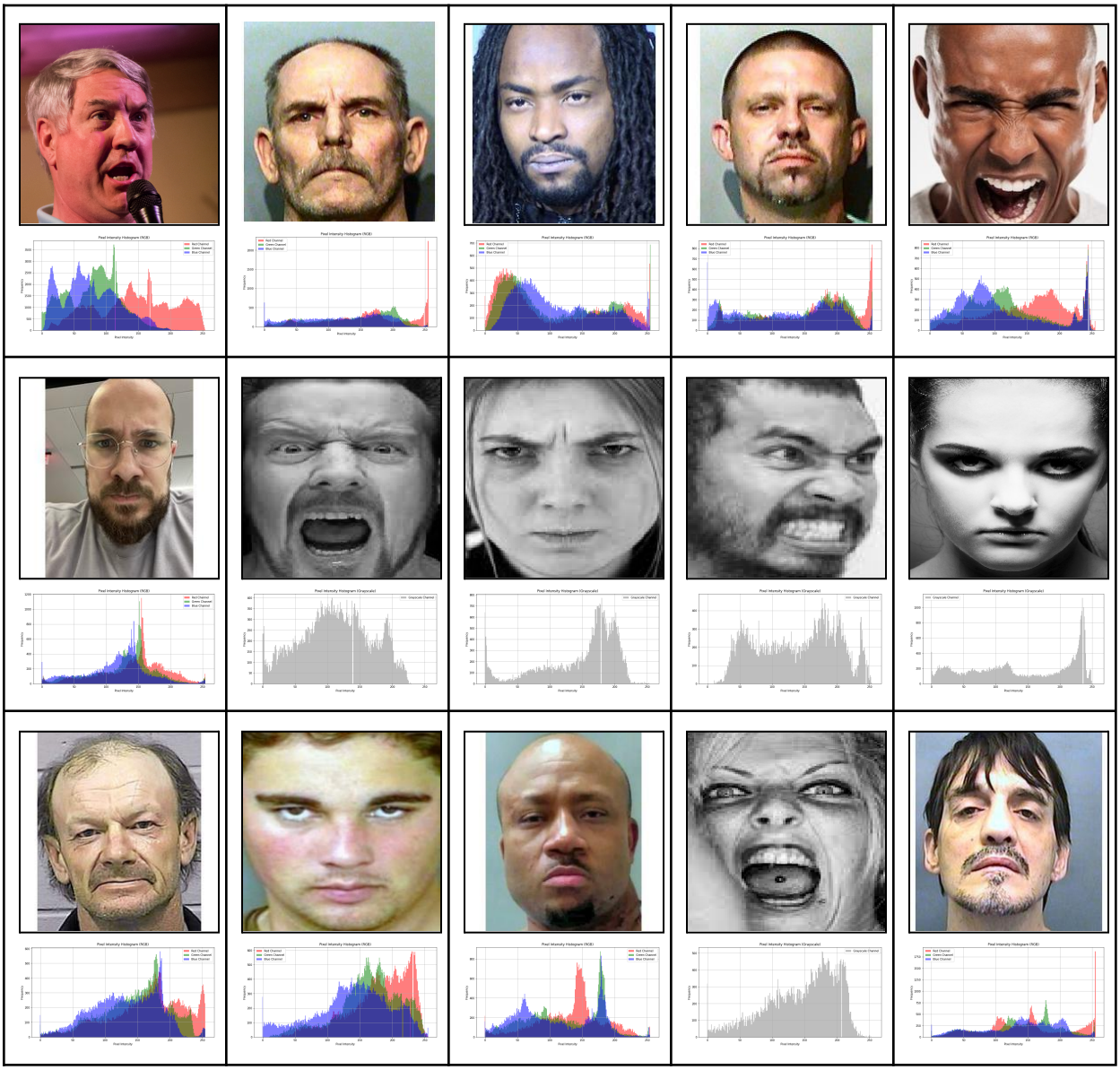

Table 2 *Sample Angry Images With Histograms*

Happy Class

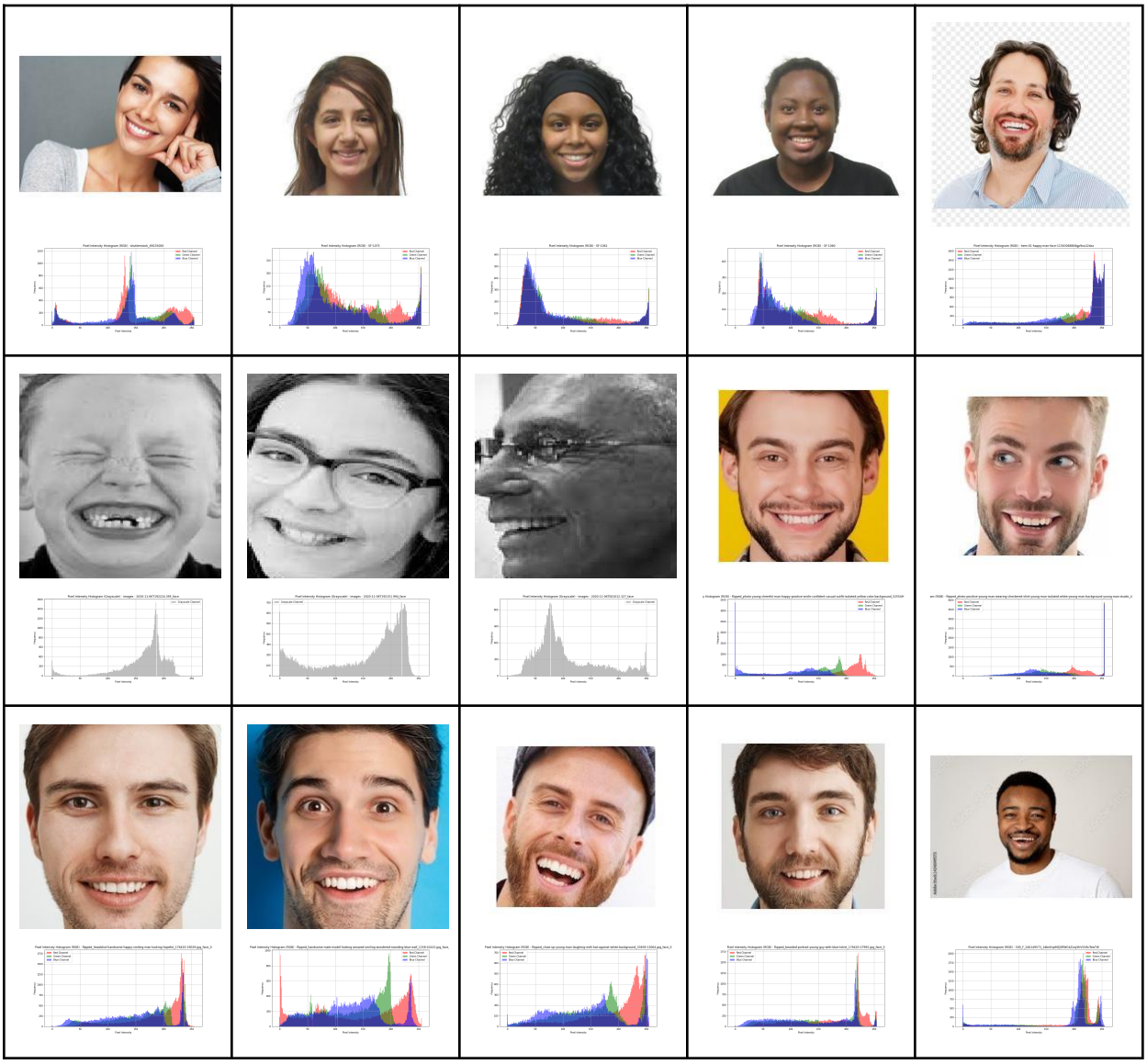

Table 3 *Sample Happy Images With Histograms*

Focused/Engaged Class

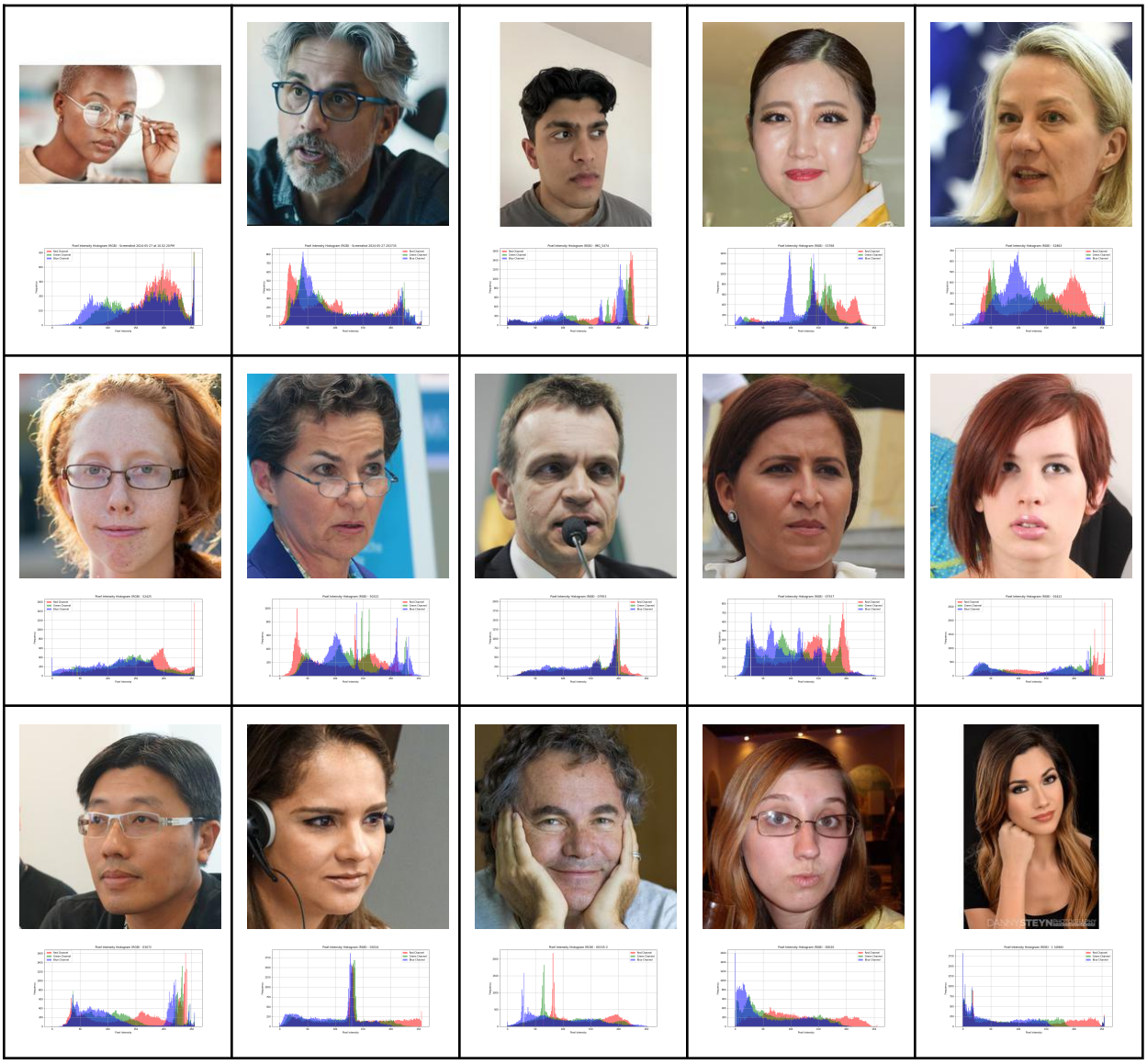

Table 4 *Sample Focused/Engaged Images With Histograms*

References

[1] J. M. Chen, J. B. Norman, and Y. Nam, “Broadening the stimulus set: Introducing the American Multiracial Faces database - behavior research methods,” SpringerLink, https://link.springer.com/article/10.3758/s13428-020-01447-8#Bib1 (accessed May 28, 2024). (Neutral, Happy)

[2] S. Vaidya, “Natural human face images for emotion recognition,” Kaggle, https://www.kaggle.com/datasets/sudarshanvaidya/random-images-for-face-emotion-recognition (accessed May 28, 2024). (Neutral, Happy, Angry)

[3] “Neutral face images – browse 48,577 stock photos, vectors, and video,” Adobe Stock, https://stock.adobe.com/ca/search?k=neutral+face (accessed May 28, 2024). (Neutral, Focused)

[4] A. Gupta, “Human faces,” Kaggle, https://www.kaggle.com/datasets/ashwingupta3012/human-faces (accessed May 28, 2024). (Neutral, Focused, Angry)

[5] E. Koar, “Happy_or_sad_binary_image_classification,” Kaggle, https://www.kaggle.com/datasets/eneskosar19/happy-or-sad-binary-image-classification (accessed May 28, 2024). (Happy)

[6] Z. Khan, “Facial recognition dataset (human),” Kaggle, https://www.kaggle.com/datasets/zawarkhan69/human-facial-expression-dataset (accessed May 28, 2024). (Happy, Neutral)