COMP 472

Artificial Intelligence Project Assignment

Part 1

Team name: **AK_18**

Data Specialist: **Nadim Khalife** (Student ID: 40188245)

Training Specialist: **Jungsoo Lee** (Student ID: 40174025)

Evaluation Specialist**: Victor-Thyreth Ouy** (Student ID: 40208821)

Project Repository: https://github.com/jungsoolee1/COMP-472-Project

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

Signatures and ID Numbers:

__________________________ Date: 2024-05-30

Jungsoo Lee 40174025

__________________________ Date: 2024-05-30

Victor-Thyreth Ouy

__________________________ Date: 2024-05-30

Nadim Khalife

I. Dataset

A. Overview

The dataset used for the artificial intelligence project is extracted from the FER 2013 (Facial Expression Recognition 2013) dataset. Here are the key details:

- **Total Number of Images**: The FER 2013 dataset contains 35,887 grayscale images of faces.
- **Selected Images**:
  - Approximately 400 images per class for training the model.
  - 100-200 images per class for testing the model.
- **Classes Used**: Neutral, angry, engaged, and happy.
- **Image Size**: All selected images are 48 x 48 pixels.
- **Diversity**:
  - Includes people of different ages, backgrounds, and ethnicities.
  - Ensures the dataset is robust and unbiased, working well across various backgrounds.
- **Image Characteristics**:
  - Most images are frontal face shots, suitable for training models focused on frontal face analysis.
  - Some images have unorthodox framing and positions, beneficial for testing to determine which angles the program evaluates best.

B. Justification for Dataset Choice

- <u>Abundant Facial Expressions:</u> The dataset contains a substantial number of images of various facial expressions, allowing us to train the model effectively.
- <u>Consistent Image Size:</u> The images are uniformly sized at 48x48 pixels, simplifying the data preparation process.
- <u>Aligned Emotions:</u> The emotions depicted in the images (neutral, angry, engaged, and happy) align well with our project's goal of detecting students' emotions.
- <u>Diverse Subjects:</u> The dataset includes people of different ages, backgrounds, and ethnicities, making the model more robust and unbiased.

C. Dataset Challenges

- <u>Resolution Limitations</u>: The 48x48 pixel resolution might miss finer details in facial expressions. This limitation can affect the model's ability to capture subtle changes in facial features that are critical for distinguishing between similar emotions.
- <u>Similar Expressions:</u> Differentiating between neutral and engaged expressions is challenging due to their similarities. Both expressions often appear similar, with

minimal differences in facial features, making it difficult to categorize them accurately.

- Detail Loss: The fixed resolution may limit the detail in facial expressions, potentially impacting the model's training effectiveness. Higher resolution images could capture more intricate details that are essential for precise emotion recognition.

D. Provenance information

Image Batch	Source	Licensing Type
All Images	FER2013 Dataset Kaggle	Public Domain / Open Source

II. Data Cleaning

A. Techniques and Methods Applied for Standardizing the Dataset

The primary goals of our data cleaning is to standardize the resolution of images, adjust their brightness, contrast, and sharpness, and ensure consistent naming conventions. The cleaned dataset provides a uniform foundation for effective machine learning model training.

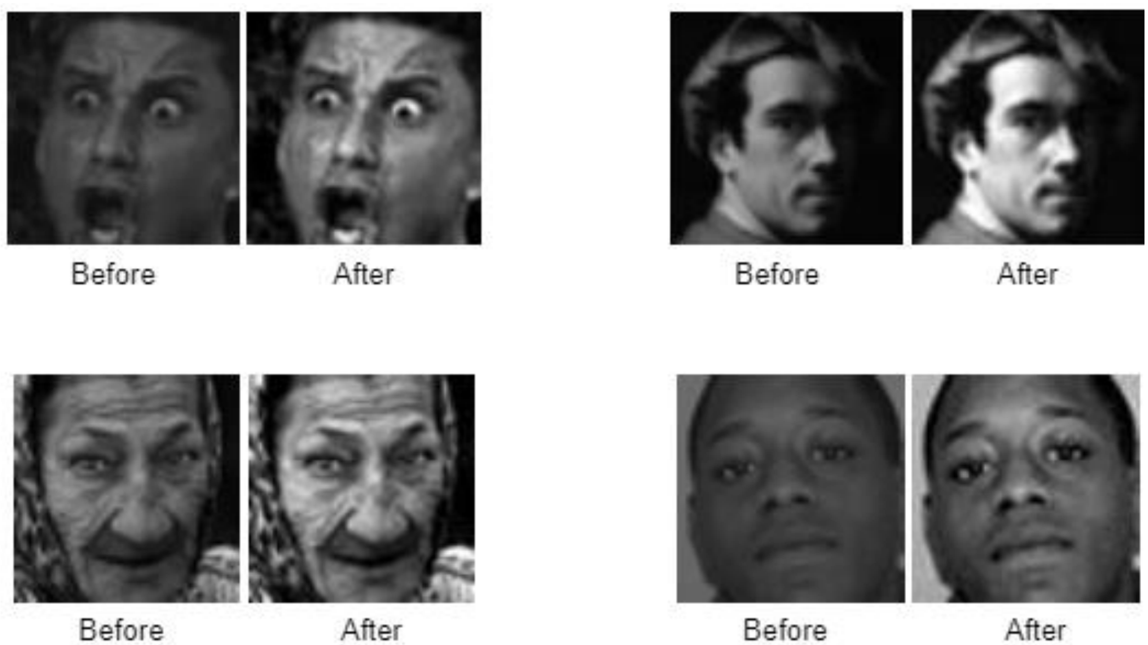

Figure 1: Comparison before and after cleaning dataset images

1. **Resizing and Conversion to Grayscale**:
    - <u>Resizing:</u> All images are resized to a standard resolution (64x64 pixels). This standardization ensures that each image has the same dimensions, facilitating efficient processing and model training.
    - <u>Conversion to B&W:</u> Images are converted to black and white (grayscale). This simplifies the data by reducing it to a single channel, which can reduce computational complexity and focus the model on essential features without the variability introduced by color. Additionally, the dataset that we found was already set in grayscale.
2. **Brightness Adjustment**:
    - <u>High Brightness:</u> Images that are too bright are adjusted to reduce their brightness. This is achieved by calculating the brightness level and reducing it proportionally if it exceeds a certain threshold of 150.
    - <u>Low Brightness</u>: Conversely, images that are too dark are adjusted to increase their brightness. This ensures that all images have a similar brightness level, enhancing the model's ability to learn from them.]

3. **Contrast Adjustment**:
    - High Contrast: Images with excessively high contrast are adjusted to reduce their contrast. This helps in preventing certain features from being overly pronounced, which can mislead the model.
    - Low Contrast: Images with very low contrast are adjusted to increase their contrast. This helps in highlighting features that might otherwise be too subtle for the model to detect.
4. **Sharpness Adjustment**:
    - High Sharpness: Images with high sharpness are adjusted to reduce their sharpness. This helps in smoothing out overly crisp edges that might create noise in the data.
    - Low Sharpness: Images with low sharpness are adjusted to enhance their sharpness. This ensures that the important features and edges within the images are clear and distinguishable.
5. **Naming**:
    - Each image is renamed using a consistent naming convention that includes the directory names and a unique identifier. This helps in maintaining a clear and organized dataset structure, which is crucial for effective data management and retrieval.

B. Challenges Encountered and Solutions

1. **Balancing Brightness and Contrast Adjustments**:
    - Challenge: Determining the right thresholds and adjustment levels for brightness, contrast, and sharpness to enhance image quality. We did not want to add a set fixed increase/decrease size since this would not be proportional to some images on either extremes.
    - Solution: The script uses proportional adjustments based on how much the image's properties deviate from the thresholds. It takes into account how much brighter/darker an image is compared to the threshold and increases/decreases based on that amount.
2. **Handling Existing Cleaned Dataset Directory**:
    - Challenge: If a cleaned dataset directory already exists, running the cleaning process again would result in conflicts.
    - Solution: The script first checks if the cleaned dataset directory exists and removes it if it does. This ensures that the cleaning process starts fresh each time, avoiding conflicts and ensuring the directory contains only the latest cleaned data.
3. **Ensuring Unique File Names**:
    - Challenge: When renaming and saving cleaned images, ensuring that each image has a unique name to prevent overwriting.
    - Solution: The script generates unique file names based on the parent directory, subdirectory, and a count. This systematic approach ensures that each image name is unique and follows a consistent pattern.

III. Labeling

The FER 2013 dataset is already pre-labeled, which simplified the process of labeling the images with their corresponding emotions. However, we had to meticulously reviewed the dataset to retain only the most accurate images that matched their designated emotions and discarded those that did not. This was crucial for achieving better results when training the model, especially since we are using a dataset of 400 images. Each team member was responsible for one emotion category, and we labeled the images accordingly until all four categories were completed. The dataset was then divided into two folders: one for training and one for testing, with each folder containing the four emotions: Happy, Engaged, Neutral, and Angry.

One significant challenge we encountered was labeling the neutral and engaged emotions, as these facial expressions appear very similar. The FER 2013 dataset includes a neutral folder with 1223 images, containing both purely neutral expressions and those with more focused, engaged looks. We worked to separate this folder into two distinct categories: neutral and engaged. For the happy and angry folders, we selected images that most accurately represented these emotions. Some facial expressions were difficult to categorize; for instance, a subtle smile might be mistaken for a neutral expression.

To ensure accuracy, we focused on categorizing each facial expression as appropriately as possible. For example, faces with wide smiles or laughs were categorized as happy, while faces with eyes open wider than usual were placed in the engaged category, indicating a more focused demeanor. Finally, to ensure that the faces were represented accurately, we cross-checked each other's work. This process helped identify and correct any inconsistencies in our dataset, ensuring that the facial expressions were correctly labeled.

IV. Dataset Visualization

A. Class Distribution:

The following is a bar graph showing the number of images in each class. The following data visualization was done using Matplotlib and Pillow [1], [2].

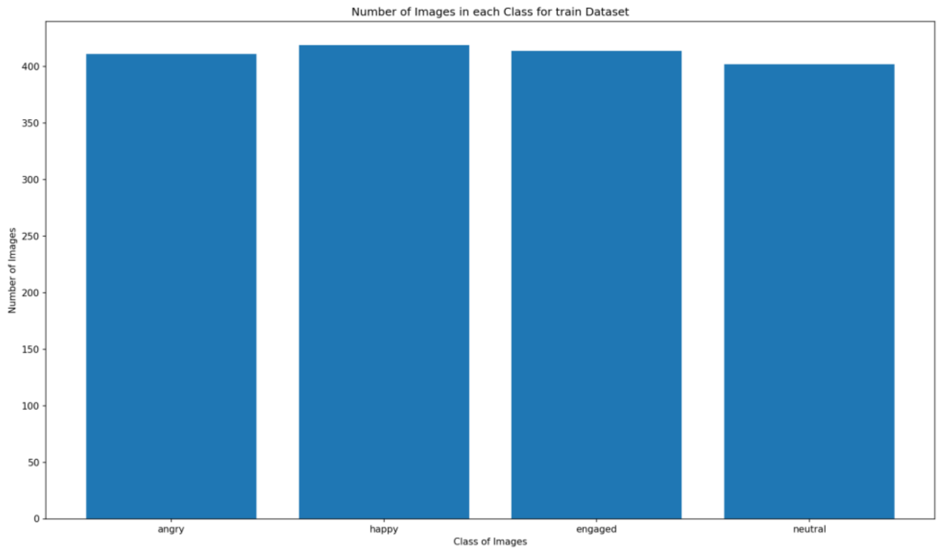

Figure 2: Number of images in each class for the training dataset

The bar graph above shows a near equal number of images for each class. We can thus conclude that none of the classes are overrepresented or underrepresented.

B. Pixel Intensity Distribution:

The following are histograms of the aggregated pixel intensity for each class. Since the images are in grayscale, the pixel intensity will measure the pixel's brightness ranging from 0 to 255, where 0 represents black and 255 is white. The frequency label on the y-axis indicates the number of pixels with the associated pixel intensity.

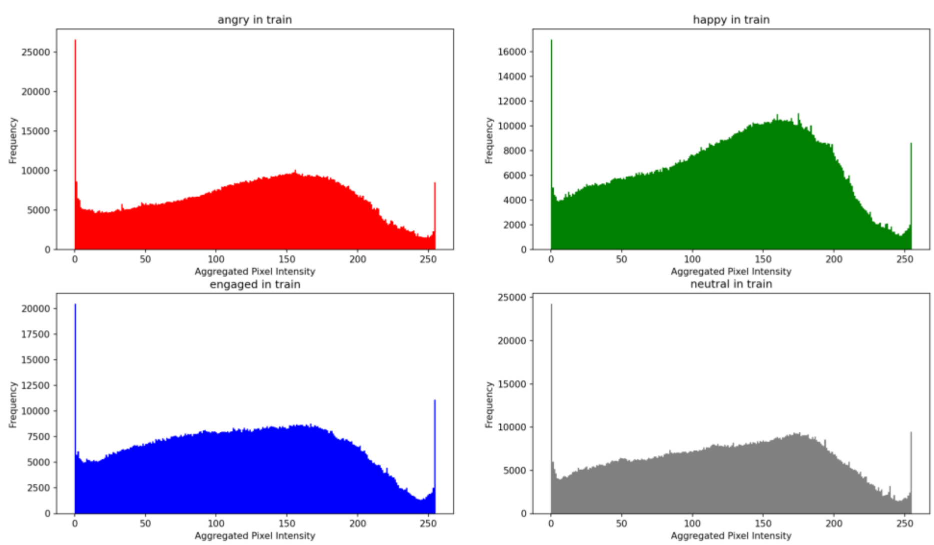

Figure 3: Histograms of aggregated pixel intensity for each class

C. Sample Images:

The following are pixel intensity histograms of 15 randomly sampled images for each class.

<u>Neutral Class:</u>

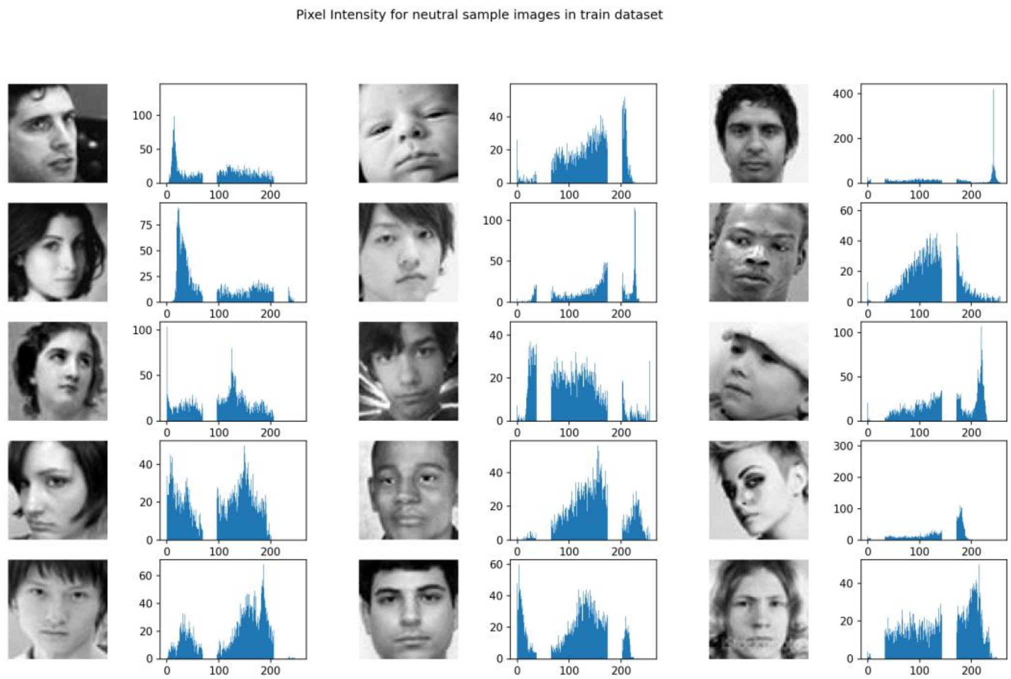

Figure 4: Pixel Intensity Histograms of Images in Neutral Class

<u>Engaged Class:</u>

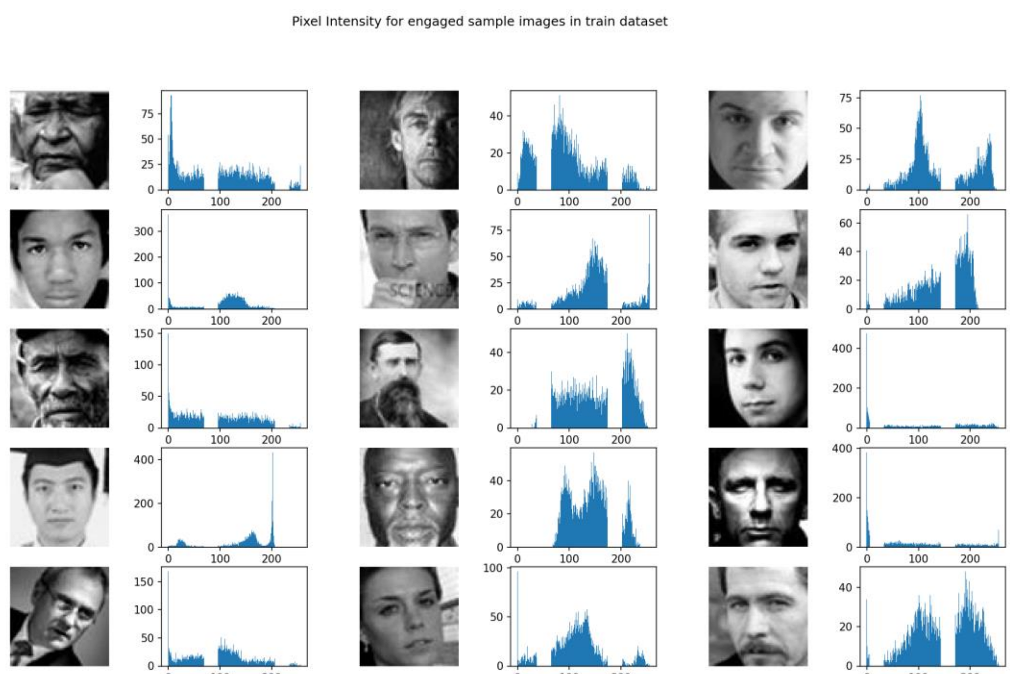

Figure 5: Pixel Intensity Histograms of Images in Engaged Class

Happy Class:

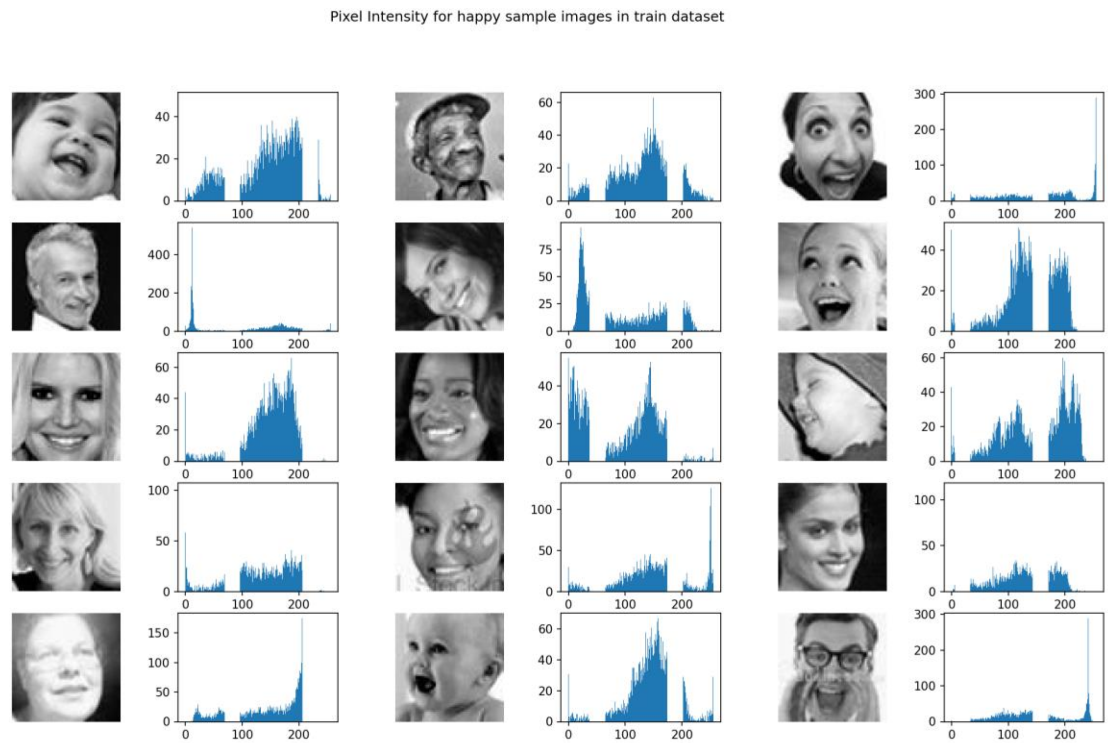

Figure 6: Pixel Intensity Histograms of Images in Happy Class

Angry Class:

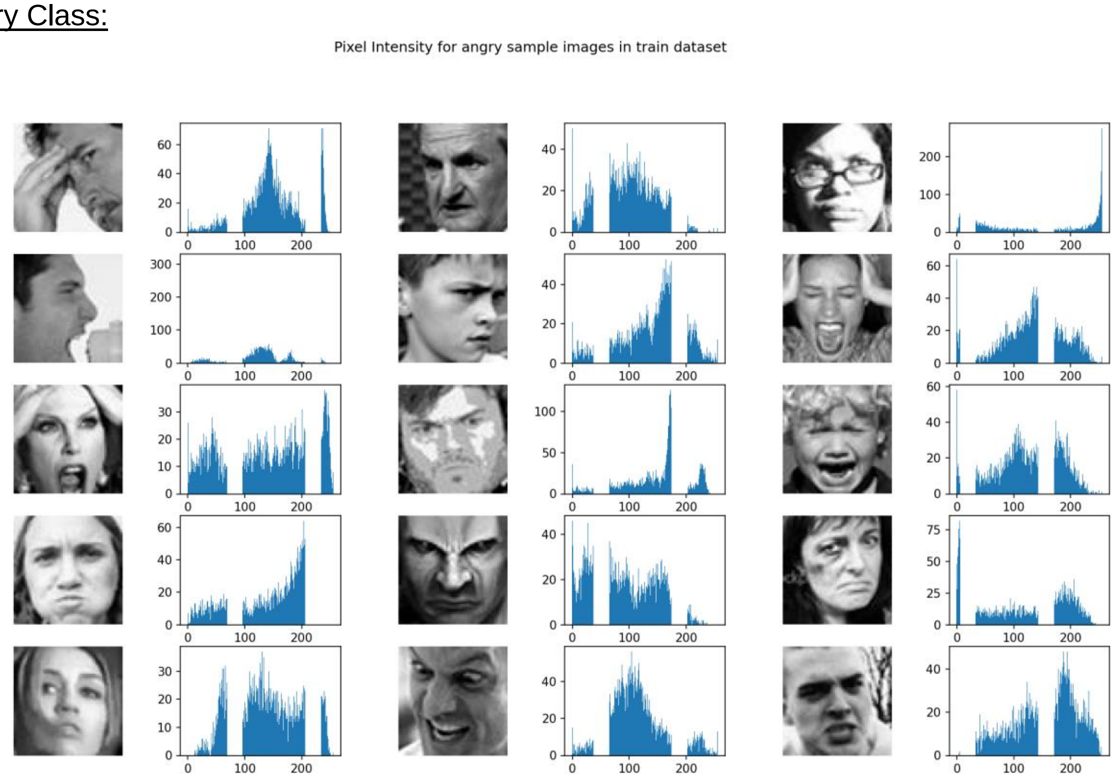

Figure 7: Pixel Intensity Histograms of Images in Angry Class

References

[1] Matplotlib, “Matplotlib: Python plotting — Matplotlib 3.1.1 documentation,” Matplotlib.org, 2012. https://matplotlib.org/

[2] A. C. (PIL F. Author), “Pillow: Python Imaging Library (Fork),” PyPI. https://pypi.org/project/Pillow/

[3] P.-L. Carrier and A. Courville, "Facial Expression Recognition 2013 (FER2013)," University of Montreal. Available: https://www.kaggle.com/datasets/msambare/fer2013. [Accessed: May 30, 2024].