Team Name

AK-18

Team Member

Mina Jaberi

Student ID: 40176988

Specialization: Data Specialist

Mahshad Mahdavi Moghadam

Student ID: 40295634

Specialization: Training Specialist

Mina Mahmoud Roshanzamir

Student ID: 40253186

Specialization: Evaluation Specialist

Link of Project:

https://github.com/Minajab/AppliedAI_EmotionDetection/tree/main

1. Dataset

i. Overview

In this section, our aim is to provide a comprehensive overview of the dataset that we used for the emotion recognition project.

The dataset was compiled from multiple sources due to the unavailability of all required classes within a single publicly accessible dataset. Specifically, we obtained data for the 'neutral' and 'angry' emotion classes from a publicly available dataset. However, the remaining two classes were sourced from a specific online stock photo.

In our dataset, we have a total of 2,115 images, which are categorized into four distinct classes. Table 1 shows the distribution of instances of each class.

Class	Number of Instances
Neutral	555
Focused	532
Tired	402
Angry	626
Total	2115

Table 1 Number of Instances in Each Class

The data for our project has been collected from two distinct sources: FER-2013 and an online stock of photos[1]. To manage the scale of the dataset, we opted to select images randomly from both sources. FER-2013 alone contains an extensive collection of 35,887 grayscale images, which exceeds the scope of our project.

This dataset exhibits distinctive characteristics that significantly contribute to its suitability for our research:

- Facial Alignment:

  Most of the images in the dataset have been automatically captured to ensure that the face is centered and occupies a consistent amount of space within the image. This feature enhances the dataset's suitability for facial analysis and related research objectives.

- Diverse Backgrounds:

  In addition to consistent facial alignment, the dataset features images collected from online stock photos, offering a variety of backgrounds. This diversity in backgrounds enriches our dataset, providing a broad range of environmental contexts for our analysis.

[1] https://www.pexels.com/

ii. Dataset Selection Justification

In this section, we provide an explanation for our choice of utilizing the FER2013 and Pexels photo stock datasets in our research project. These datasets offer unique advantages and address specific needs in our study.

FER2013 Dataset: The FER2013 dataset plays a pivotal role in our research for the following reasons:

- Dataset Size and Suitability: FER2013 is a relatively large dataset, making it well-suited for the training of neural network algorithms. Its size allows for robust model development.
- Public Availability: This dataset is publicly available, enhancing accessibility for both researchers and students. This openness aligns with our commitment to transparency and collaboration.
- Research Validity: FER2013 has been utilized in numerous research papers, with demonstrated success in facial expression recognition. Its usage in prior studies affirms its reliability and suitability for our research objectives.

However, it's important to acknowledge certain limitations in the FER2013 dataset:

- Resolution Challenge: The images in FER2013 are of relatively low resolution (48x48 pixels). This can pose a challenge as it may not adequately capture fine-grained facial features, potentially limiting our model's ability to recognize subtle emotional expressions.
- Class Limitations: FER2013 has limitations in terms of the classes it includes. Notably, it lacks complex facial expressions such as 'tired' and 'focused.' These constraints prompted us to seek additional resources for these intricate emotional expressions.

Pexels Photo Stock: The choice of the Pexels Photo Stock dataset is driven by the following considerations:

- Resource Versatility: Pexels Photo Stock serves as a valuable resource for projects requiring diverse, high-quality visuals. Its extensive collection of images provides a rich source of visual content, essential for our research.
- Accessibility: Pexels Photo Stock is readily accessible, contributing to the efficiency of our project.

However, it is important to recognize certain challenges associated with Pexels Photo Stock:

- Content Overuse: The popularity of Pexels may result in some images being overused across various projects. This overuse can impact the uniqueness of the content we utilize.
- Differing Image Quality: Images in Pexels Photo Stock vary in quality, which may necessitate careful selection to ensure consistency in the dataset.

In conclusion, our choice of the FER2013 and Pexels Photo Stock datasets is driven by their respective strengths and the specific needs they fulfill in our research. Acknowledging both their advantages and limitations allows us to make informed decisions and utilize the datasets effectively in our project.

iii. Provenance Information

In Table 2, we provide complementary information about our collected dataset.

Batch Image	**Source**	**Dataset Reference**	**License**	**Open Access**
Neutral	FER2013[2]	https://www.kaggle.com/datasets/msambare/fer2013	https://opendatacommons.org/licenses/dbcl/1-0/	Public Domain
Angry	FER2013	https://www.kaggle.com/datasets/msambare/fer2013	ttps://opendatacommons.org/licenses/dbcl/1-0/	Public Domain
Tired	Pexels	https://www.pexels.com/	https://www.pexels.com/license/	Attribution not required
Focused	Pexels	https://www.pexels.com/	https://www.pexels.com/license/	Attribution not required

Table 2 Information of Each Image Batch

[2] Carrier, Pierre-Luc, Aaron Courville, Ian J. Goodfellow, Medhi Mirza, and Yoshua Bengio. "FER-2013 face database." Universite de Montral 3 (2013).

2. Data Cleaning Phase

A set of clearance functions has been applied to the dataset, which is listed in order below.

1. **Deleting background ():** To enhance the model's precision, background detection, and elimination is essential. In the FER-2013 dataset, the focus is solely on the faces, whereas the pictures obtained from another source exhibit the opposite scenario. As a result, to improve the model's accuracy for two distinct emotions, we have implemented this function for the 'focused' and 'bored' classes. An example output for this function can be found below.

Figure 1 Bored class. The leftmost image is the original, and the rightmost one is the processed version.

2. **Rotation ():** During testing, the input image can exhibit various degrees of rotation. Therefore, to make the model more robust, it is valuable to introduce slight rotations within a subset of the data. In the rotation function, for each training image, a random number between zero and one should be generated. If the random value is less than 0.5, the image should be rotated. The degree of rotation should be a random number between -30 and 30 degrees. Here is an example output:

Figure 2 Angry class. The leftmost image is the original, and the rightmost one is the processed version.

3. **Saturation adjustment ():** The sentence is clear and well-structured. It effectively communicates the purpose of saturation adjustment. There are no apparent issues with the sentence. This function includes a parameter known as the 'saturation factor.' When the value of this factor is greater than one, it increases the purity of colors in an image. This function has been applied to the 'bored' and 'focused' classes, which contain colorful images with a saturation factor of 1.5 (resulting in a slight alteration). An example output is as below.

Figure 3 Bored class. The leftmost image is the original, and the rightmost one is the processed version.

4. **Grayscale ():** The 'angry' and 'neutral' classes consist of grayscale images, while the remaining classes have colorful images. Therefore, the grayscale transformation has been applied to the colorful classes. The benefits of transforming the images to grayscale are as follows:
    a. Grayscale images contain only one channel, as opposed to RGB images, which have three channels. Consequently, converting to grayscale results in reduced memory usage and a faster training procedure.
    b. Grayscale images retain essential information while eliminating potentially distracting color information. The training model's focus is solely on the intensity and luminance of the pixels, simplifying the feature extraction and training stages.

Figure 4 Focused class. The leftmost image is the original, and the rightmost one is the processed version.

5. **Contrast Regulation ():** After applying the grayscale function to all emotional classes, contrast regulation should be performed. The rationale behind this is that emotional expressions are often conveyed through facial features, and ensuring consistent and controlled contrast can help the model better focus on these features. The 'contrast-factor' parameter determines the intensity of contrast and has been set to 1.1, resulting in a slight increase in contrast.

Figure 5 Neutral class. The leftmost image is the original, and the rightmost one is the processed version.

6. **Morphology (), Denoise (), and Deblock ():** The images collected from the FER-2013 dataset are notably noisy. When zoomed in, artifacts are easily visible. Therefore, three functions—Morphology (), Denoise (), and Deblock ()—have been applied in sequence.

   Morphological operations involve a chain of dilation and erosion, which helps reduce or eliminate artifacts and blocks. Dilation regularizes and smooths boundaries, while erosion is effective at removing small noise or isolated pixels in an image [3].

   For denoising, the fast mean denoising method has been used. Unlike some noise reduction techniques that blur or smooth images, non-local means denoising aims to preserve important details and structures in the image. As a result, it is a good choice for denoising the FER-2013 images, which may not be of high quality.

   In the deblocking process, we have two steps: the application of the bilateral filter and Gaussian blur. The bilateral filter reduces compression artifacts, often visible as blocky patterns while preserving edges and fine details in the image. Since relying solely on the bilateral filter was not sufficient for mitigating the blocks, we subsequently applied Gaussian blur, making the artifacts less pronounced and visually disruptive. Example outputs after applying these three deblocking functions are as follows.

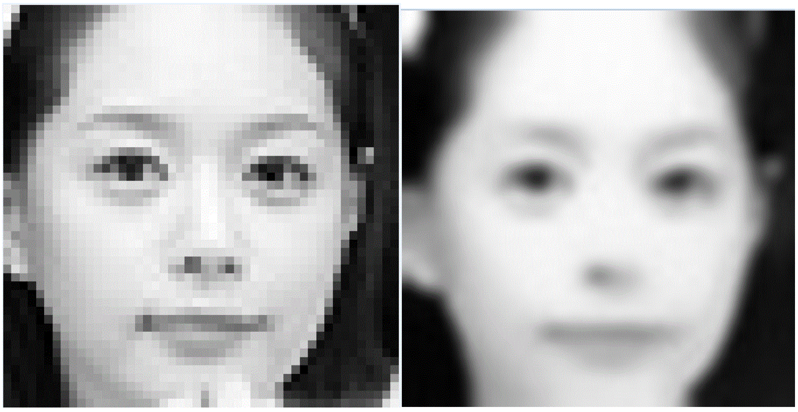

Figure 6 Neutral class. The leftmost image is the original, and the rightmost one is the processed version.

7. **Resizing the image ():** We calculated the maximum width and height among the images from different classes and then adjusted the size of the images to match the maximum width and height.

3. Labeling:

While we collected our dataset from different resources, all our data was pre-labeled.

- FER2013 Dataset: For the classes related to emotions (Angry, Neutral), we utilized the dataset, which was publicly available and had pre-labeled emotional expressions.
- Pexels Photo Stock: For classes such as Tired and Focused, we sourced images from the Pexels Photo Stock website. These images were categorized based on the presence of tired or focused expressions.

The utilization of the FER2013 dataset and images from Pexels Photo Stock was instrumental in labeling our dataset accurately and efficiently.

4. Visualization phase

i. Class Distribution Analysis

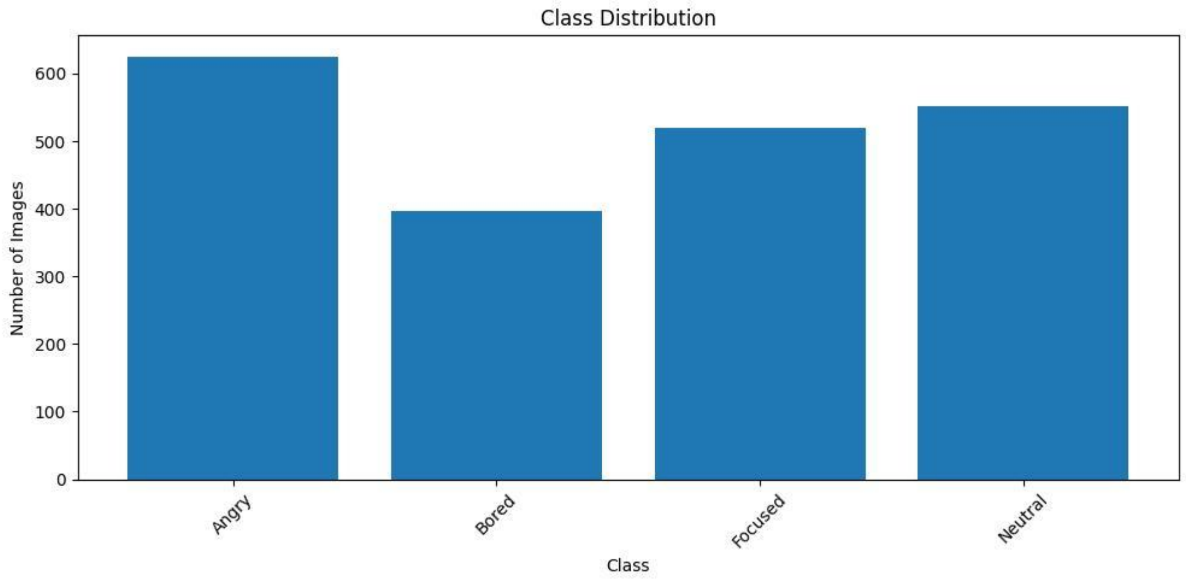

figure 7 (class distribution)

- **Methodology**: We utilized the class_dist function, which leveraged Matplotlib to generate a bar graph showcasing the number of images in each class.

- **Findings (Refer to Figure 7)**:
  - We have four distinct emotion classes: Angry, Bored, Focused, and Neutral.
  - The Angry class contains the highest number of images, highlighting ample data for this emotion.

- ○ The Bored class has fewer images, hinting at the possible need for data augmentation to balance this class.
- ○ Focused and Neutral classes are almost equal in terms of the number of images they contain.
- ○ An imbalance in dataset distribution can lead to a model bias, making this visualization crucial for identifying such concerns beforehand.

ii. Sample Image Visualization

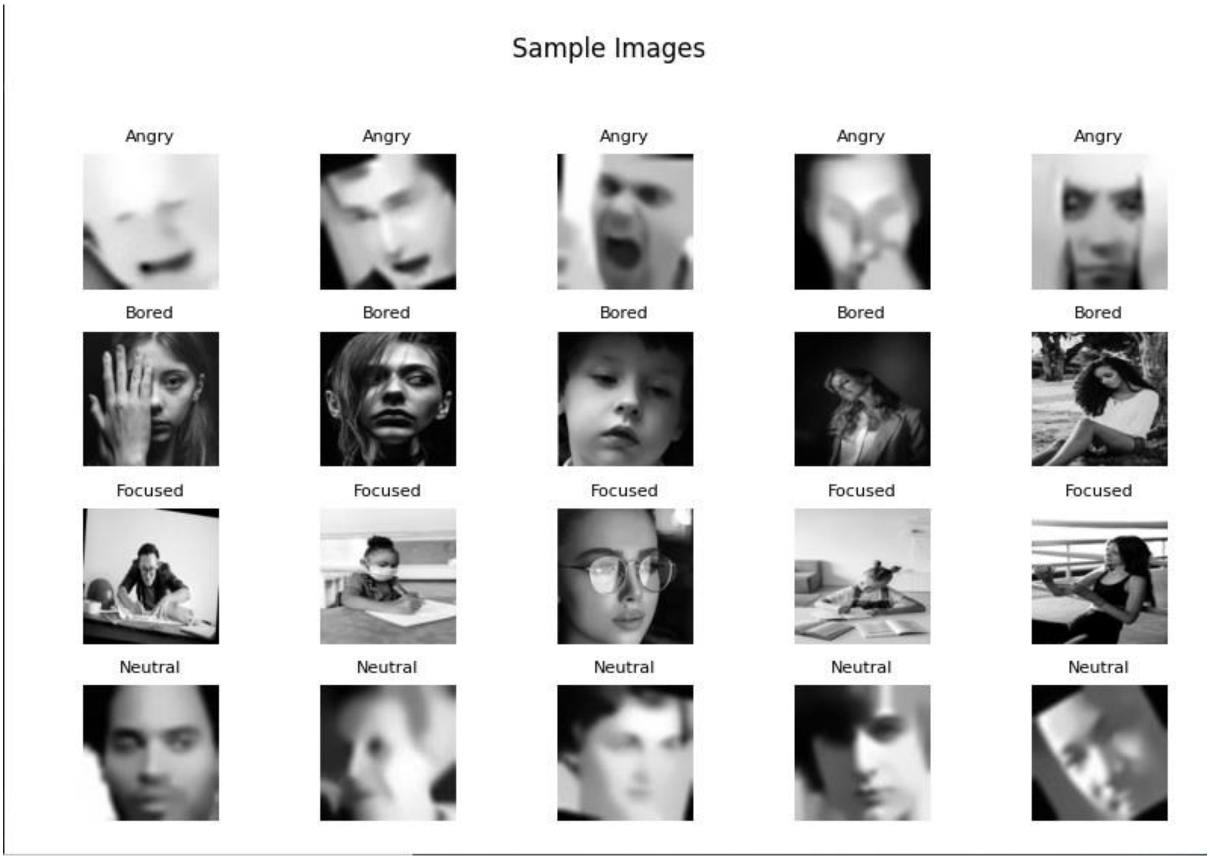

Figure 8 (sample images)

- **Methodology**: We employed the sample_imgs function, which is designed to showcase a set of images from various classes in a 5x5 grid format.

- **Findings (Refer to Figure 8)**:
  - ○ **Quality and Clarity**: There is a variation in the quality of images; some are sharp, while others are blurred.
  - ○ **Variability**: The dataset showcases a range of subjects in terms of people and their poses, emphasizing the dataset's diversity.
  - ○ **Potential Anomalies**: It is essential to periodically check these images to identify any mislabeling or inconsistencies, as these can affect model performance.

iii. Pixel Intensity Distribution Analysis

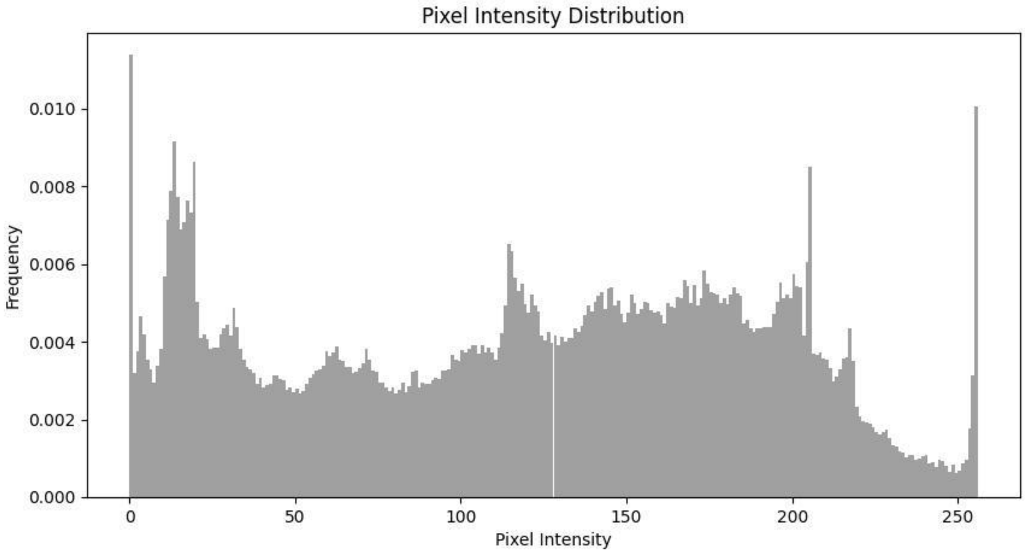

Figure 9 (pixel intensity distribution)

- **Methodology**: We introduced the pixel_dist function, aimed at displaying a histogram to depict the distribution of pixel intensities across a list of images.

- **Findings (Refer to Figure 9)**:
  - **Dark Pixels**: A notable peak around the 0-50 intensity range indicates many dark pixels in the images, possibly due to shadows, dark backgrounds, or objects.
  - **Mid-Tone Pixels**: The 100-150 range displays a balanced distribution of pixels, likely representing objects in typical lighting conditions.
  - **Bright Pixels**: A significant peak near the 250-intensity range suggests a considerable number of very bright or white pixels, which could be due to overexposed areas or flash reflections.

- **Insights**:
  - **Variations in Lighting**: The wide range of intensities indicates that the dataset contains images with varied lighting conditions, which is beneficial for training models.
  - **Potential Anomalies**: The distinct peaks and troughs in the histogram can help in spotting anomalies like overexposed or underexposed images. In such cases, data augmentation might be necessary.

Reference

[1] "Pexels," [Online]. Available: https://www.pexels.com/. [Accessed: October 15, 2023].

[2] P.-L. Carrier, A. Courville, I. J. Goodfellow, M. Mirza, and Y. Bengio, "FER-2013 face database," Université de Montréal, 2013.

[3] Maulion, M. (2021, January 31). Morphological Operations: A Cleaning Technique in Image Processing. Medium. https://mattmaulion.medium.com/morphological-operations-a-cleaning-technique-in-image-processing-155baed8fcd1