**Data cleaning phase**

The selected classes for emotional recognition consist of angry, bored, focused, and neutral. The data has been gathered from two different datasets, including FER-2013 and an online stock of photos. The reason behind choosing two different datasets is that the classes 'bored' and 'focused' do not exist in the FER-2013 dataset.

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

1. **Deleting background ():** To enhance the model's precision, background detection and elimination are 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.

**A person with her hand on her face

Description automatically generated**

Figure 1 'Bored' class

1. **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:

**A blurry image of a person's face

Description automatically generated**

Figure 2 'Angry' class

1. **Saturation adjustment ():** Saturation adjustment is useful for enhancing color vibrancy and regulating color intensity. 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.

**A person standing in front of a bridge

Description automatically generated**

Figure 3 ‘Bored’ class

1. **Gray scale ():** 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:
   1. 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.
   2. 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.

A person with a bun in her hair

Description automatically generatedA person with a bun in her hair

Description automatically generated

Figure 4 ‘Focused' class

1. **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

1. **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 smoothes boundaries, while erosion is effective at removing small noise or isolated pixels in an image.

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.

A close-up of a person's face

Description automatically generatedA blurry image of a person's face

Description automatically generated

Figure 6 ‘Neutral’ class

1. **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.