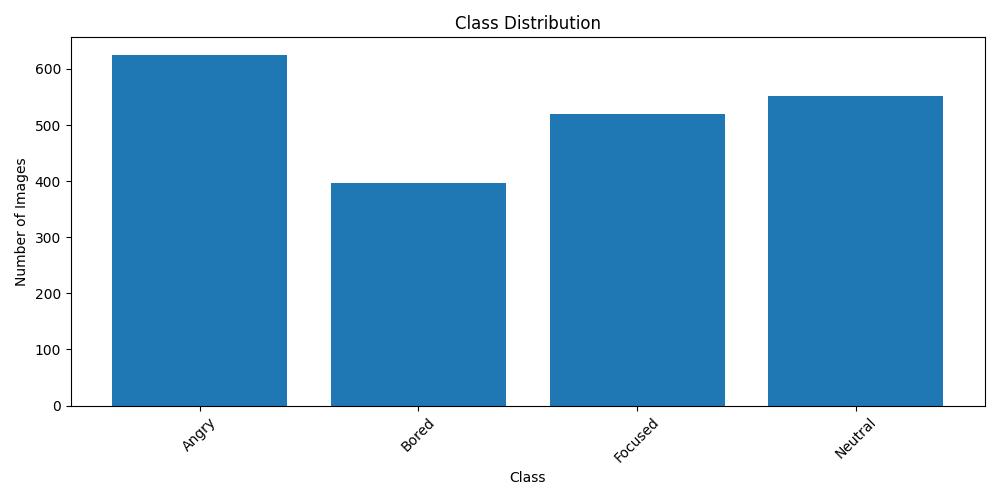
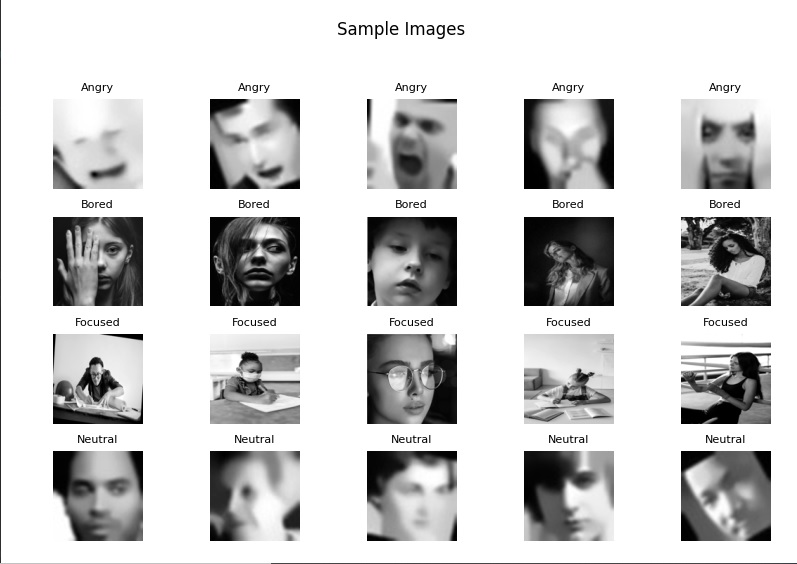
**1. Class Distribution Analysis**



Figure

* **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 1)**:
  + 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.

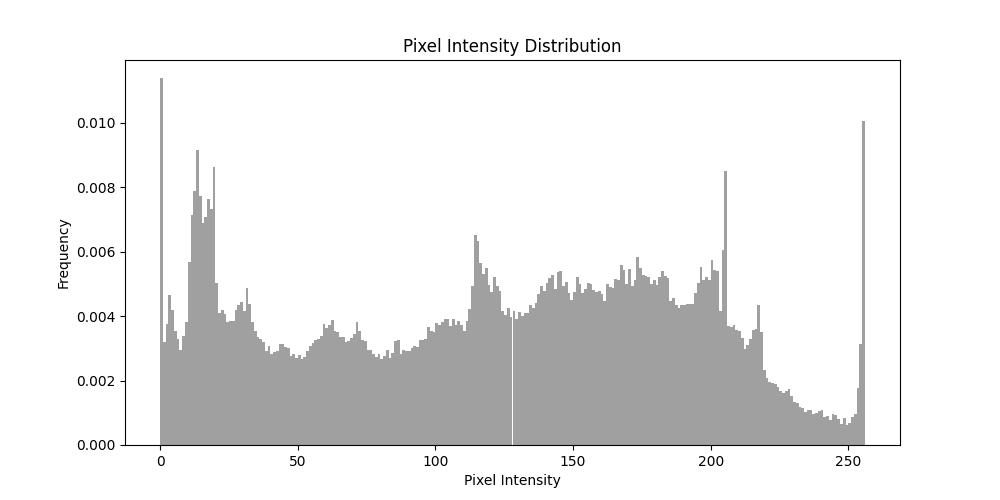
**2. Sample Image Visualization**



Figure

* **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 2)**:
  + **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.

**3. Pixel Intensity Distribution Analysis**

****

Figure

* **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 3)**:
  + **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.