COMP 6721

Applied Artificial Intelligence

Fall 2023

Project Assignment Part 1 – AK_8

Name	Specialization	Student ID
Krutik Gevariya	Data Specialist	40232386
Shivam Patel	Training Specialist	40226428
Nayankumar Rajeshbhai Sorathiya	Evaluation Specialist	40227432

GitHub Link : https://github.com/nayansorarhiya/AAI_Project.git

Dataset Information

Overview

In our emotion recognition project, we employed two datasets: "Dataset A" and "Dataset B" to classify emotions into four distinct categories, namely **Neutral**, **Engaged**, **Bored**, and **Angry**.

- **Dataset A:**
  - Total Number of Images: 1,118
  - Number of Images per **Angry**, **Neutral** Class: 559

- **Dataset B:**
  - Total Number of Images: 1,110
  - Number of Images per **Boredom**, **Engagement** Class: 555

Total : 2,228

Figure 1. Facial Expression Categories

Attributes : Dataset A predominantly comprises frontal facial images featuring varied lighting conditions, a wide range of ages, gender, and diverse backgrounds.

Justification for Dataset selection

- We opted for **Dataset A** because of the high quantity of images and the concentration of **frontal face shots**, which are beneficial for training a reliable face recognition model.
- **Dataset B** was chosen to add variability to the training data, since it includes photos taken in various contexts. This combination of datasets allows our model to have a successful performance in **real-world scenarios** where lighting and backgrounds may differ. Nonetheless, this also presents a challenge when it comes to managing the augmented diversity of the data.

Provenance Information

Dataset	Image Source	Licensing Type	Relevant Information
Dataset A (1)	Link for Dataset A (1) [1]	Creative Commons	559 images of class Neutral
Dataset A (2)	Link for Dataset A (2) [2]	Creative Commons	559 images of class Anger
Dataset B	Link for Dataset B [3][4][5]	Custom License "Terms of Use."	555 images of class Boredom and Engagement

Data Cleaning

We used a number of strategies and procedures to preprocess the dataset during the "Data Cleaning" stage of our emotion recognition project to make sure it was consistent and of high quality for further analysis. The procedures followed, difficulties encountered, and illustrations of the cleaning effects before and after are included in this section.

Techniques and Methods

1. Standardization of the Dataset

1.1. Resizing Images

Resizing the photos to a consistent resolution was one of the first steps in data cleaning. This guaranteed data homogeneity while simultaneously lowering the computational cost of our model. Since **224 × 224** pixels is a standard input size for many deep learning models, we scaled every image to that size.

1.2. Single format files

We implemented file format conversion techniques to help make the model more resilient to changes in image format **png** to **jpg**. By doing so, we aimed to reduce the impact of file format change variations on the model's performance.

2. Challenges and Solutions

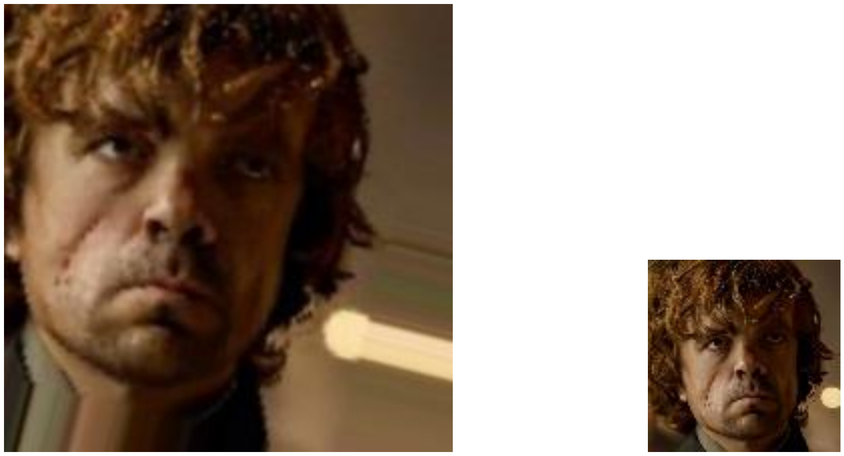
2.1. Data Imbalance

A problem with data imbalance arose throughout the data cleansing process. Our dataset's emotion categories were not evenly distributed, with considerably less samples for **boredom** and **engagement** emotions than for others. We created a more balanced dataset by under sampling the overrepresented categories and oversampling the underrepresented ones in order to remedy this problem.

2.2. Noise and Artifacts

Our datasets contained images with noise, artifacts, and outliers that could adversely affect model training. To address this, we applied data augmentation techniques such as resize the image to reduce the impact of noise.

3. Example



Before After

Figure 2. Image resize

Labeling

1. Merging Datasets and Class Mapping

1.1. Dataset Merging

Our project involved the combination of multiple datasets, each with its own set of emotion labels. Merging these datasets required mapping the emotion labels to a common set of emotional categories (Neutral, Engaged, Bored, Angry) to maintain consistency across the entire dataset. Dataset is stored in **CSV** file with fields file **name**, type of **class**, **size** of images, and file **format**.

1.2. Challenges Faced

Inconsistent Labeling Schemes: Different source datasets had varying emotional labels. For example, one dataset used "Happiness" instead of "Engaged." This inconsistency posed a challenge in mapping labels accurately.

Data Skew: Some datasets had an uneven distribution of emotions, which required oversampling and under-sampling to achieve class balance.

1.3. Solution

All dataset is mapped to one single CSV file with related categories of class.

	A	B	C	D	E	F
1	Image Name	Emotion	Image For	Width	Height	Size (bytes)
2	image0028549.jpg	Angry	jpg	224	224	7749
3	image0028550.jpg	Angry	jpg	224	224	7138
4	image0028561.jpg	Angry	jpg	224	224	8507
5	image0028562.jpg	Angry	jpg	224	224	6344
6	image0028563.jpg	Angry	jpg	224	224	7369
7	image0028566.jpg	Angry	jpg	224	224	8134
8	image0028572.jpg	Angry	jpg	224	224	7053

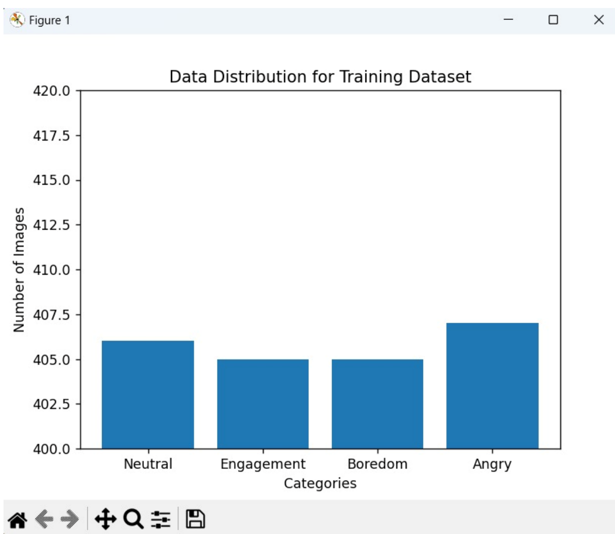
Figure 3. CSV file for image labeling

Dataset Visualization

In this section, we present visualizations that provide insights into our dataset. Effective visualization is crucial for understanding the data distribution, content, and potential challenges.

1. Class Distribution

To gain an understanding of the dataset's class distribution, we plotted a bar graph depicting the number of images in each emotion category. Class distribution is a critical factor to identify if any class is overrepresented or underrepresented. The following bar graph shows the class distribution in our dataset:



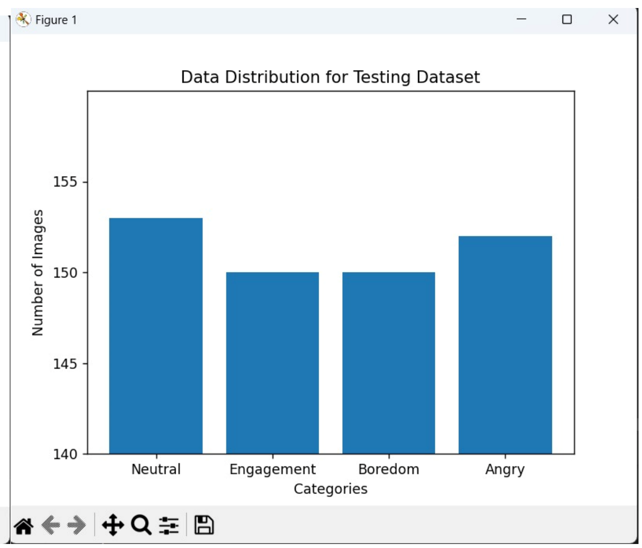


Figure 4, 5. Bar graph for Training and Testing images

2. Sample Images

We randomly selected 25 images, random from each class, and arranged them in a 5x5 grid. We can use this visualization to gain insight into the range of facial expressions and check for any unusual patterns or incorrectly labeled data.

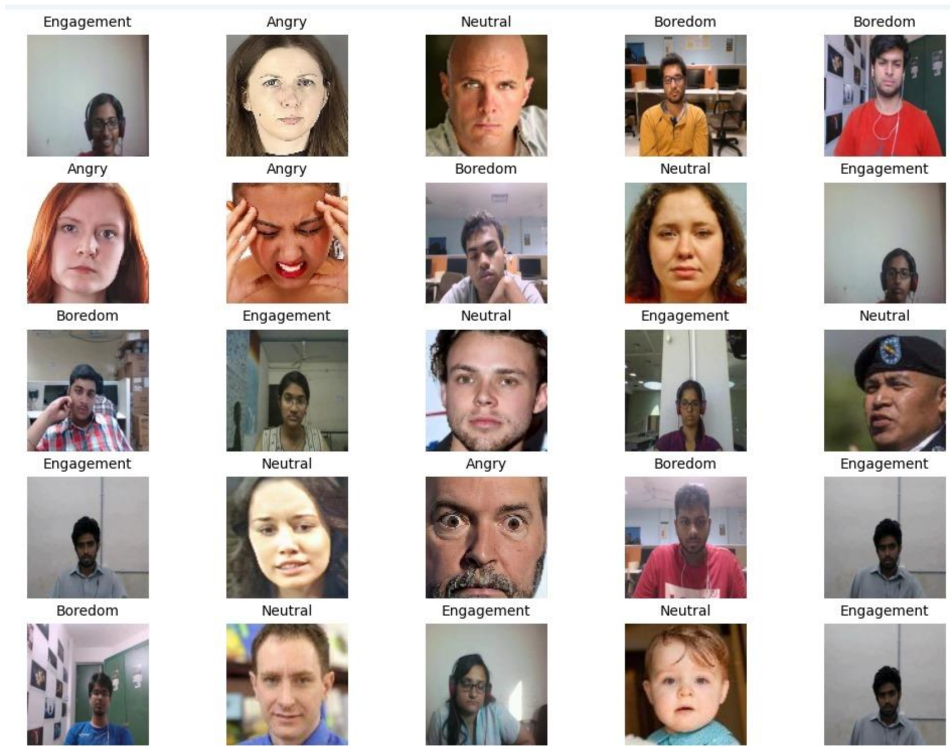


Figure 6. Labels for 25 random images

Analyzing these photographs, we can see different facial expressions and illumination, which are important components for identifying emotions. This also helps us validate the dataset's quality and maintain the accuracy of labeling.

3. Pixel Intensity Distribution

For color (RGB) images, we overlaid the intensity distributions of the Red, Green, and Blue channels on a single histogram for the same random 25 images.

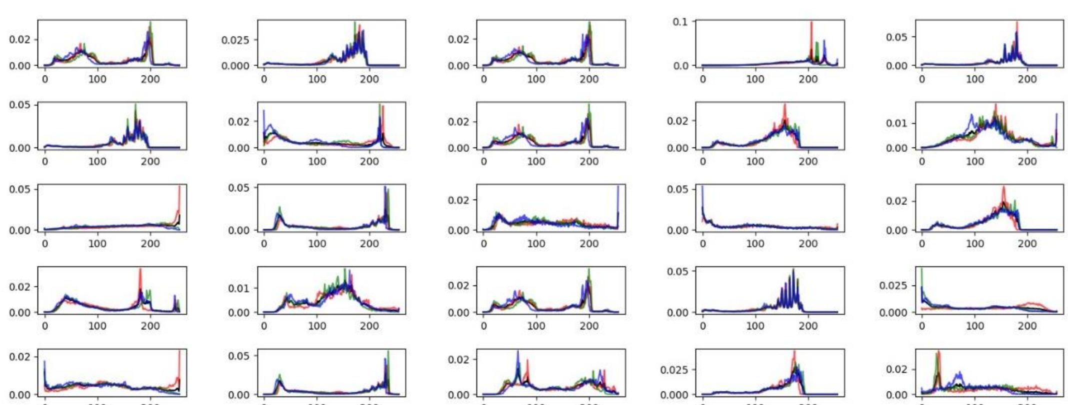


Figure 7. Pixel illumination for 25 random images

This histogram shows the distribution of pixel intensities for each color channel, offering insights into the range of lighting conditions within the dataset. Analyzing this information is essential for developing robust models capable of handling varying illumination levels.

Reference

[1]"Facial Expressions Training Data," www.kaggle.com. https://www.kaggle.com/datasets/noamsegal/affectnet-training-data?select=anger (accessed Oct. 27, 2023).

[2] "Facial Expressions Training Data," www.kaggle.com. https://www.kaggle.com/datasets/noamsegal/affectnet-training-data?select=neutral (accessed Oct. 27, 2023).

[3] "DAiSEE : Dataset for Affective States in E-Environments," people.iith.ac.in. https://people.iith.ac.in/vineethnb/resources/daisee/index.html

[4] "A Gupta, A DCunha, K Awasthi, V Balasubramanian, DAiSEE: Towards User Engagement Recognition in the Wild, arXiv preprint: arXiv:1609.01885"

[5] "A Kamath, A Biswas, V. Balasubramanian, A Crowdsourced Approach to Student Engagement Recognition in e-Learning Environments, IEEE Winter Conference on Applications of Computer Vision (WACV'16)"