Feb 23, 2025

Milestone 1: Data Processing

Attendees Liam Youm Kevin Nguyen

Attachments Milestone 1: Data Processing

Agenda

Topic	Time	File	Team member
Qualities of images in the dataset	30 min	File	Liam Youm Kevin Nguyen
Data Augmentation Methods	30 min	File	Liam Youm Kevin Nguyen
Data Cleansing	30 min	File	Liam Youm Kevin Nguyen

Exploratory Data Analysis

Because all of the images in the dataset are grayscale, we expect this characteristic to be advantageous for training the CNN model. By eliminating features related to skin color, the model can focus exclusively on learning facial expression patterns.

Furthermore, in Milestone 0, we found that the disgust class contained significantly fewer images compared to the other classes. Upon visual inspection of the images in this class, we agreed that they do not differ substantially from those in the angry class. Therefore, we plan to merge the disgust class images into the angry class.

If we were to train the model without merging the disgust class, the imbalance in data distribution and the difficulty in distinguishing it from the angry class could pose challenges. We anticipate that merging this class will have a positive impact on the model's performance.

Although merging the classes partially mitigates the data distribution imbalance, certain classes still contain more images than others. We plan to further address this imbalance by employing image augmentation techniques.

We believed that rotation and flipping augmentations would make our model the most robust. This is because we believe that our model should be able to still process and predict a person's expression any way they spin their head, which would mean rotation. Flipping, though likely not a significant augmentation, as facial expressions tend to not vary much when flipped a certain axis, may help our model detect a more diverse set of facial expressions, as it now must account for the facial expressions it trains on *and* their flipped versions. Other augmentations, such as cropping, color, and resizing, seemed unviable for one reason or another: CNNs should not be expected to predict from cropped faces, as they may not have enough key info to make a correct prediction; color is irrelevant when the image is converted into grayscale; and resizing seems unnecessary, as we don't plan on training our CNN on image with varying sizes.

Data Cleansing

Despite our images already being grayscale, `v2.Grayscale(1)` was added into the transformation function as a fix in order to properly make the images grayscale according to the PyTorch tensor. Without this, the tensor's shape for the images actually contains 3 channels rather than 1. It also serves to make any future image grayscale, if there's a case in the future where we would want to add our own images.

Our dataset included expressions for disgust and for anger. Ultimately, we chose to merge the disgust folder into the anger folder. This is because the disgust folders for training and testing contained significantly less samples than the other expressions, with a difference of over 1000 samples in some cases. When viewing the disgust expression samples, they all seemed similar to the anger expression samples. From this, and from seeing previous examples of CNNs intertwining disgust and anger expressions, we decided to place all disgust images into the anger folder.

We decided to cut a few thousand images from the happy facial expression folder. This is because, in both test and training folders, there were excess happy expression images compared to every other expression. In order to prevent and biases and to provide fair training and testing, we decided to cut some of these images to even out the images for each facial expression

We left the rest of the data features unchanged. Certain qualities, such as the mix of cartoon and real-life faces, may actually make our CNN more robust. As a result, we left them unchanged. It would also be very tedious and time-consuming to individually pick something such as cartoon faces out.