Gavin Uhran

CSE 40535

Professor Czajka

December 8, 2022

**Semester Project Part 5: Final Report**

For my test database, I collected images of family and friends that feature themselves as a lone subject in the frame. Since my program only analyzes images in which it detects one singular face, it was important for me to only include images with one subject. Overall, the database is 30 images, which I believe to be enough to gain an accurate perspective on the performance of my project and its classifiers.

There are several differences between my test set and the training and validation datasets from the renowned CelebA dataset. To start, my test set features subjects from a wider range of ages than the subjects in the CelebA dataset. My test set ranges from pictures of my cousin as a child to pictures of my grandmother, whereas I observed the CelebA dataset primarily included images of celebrities as young adults or middle-aged. Another difference between my test set and the CelebA dataset is the orientation of the subject in the photo. While the CelebA dataset features images primarily oriented as front-on headshots of the subject, my dataset contains images of subjects who are not directly facing the camera, or who may have their head not in a perfect vertical orientation. Additionally, since the images are not headshots, the lighting of the subjects in my test set is significantly less consistent than the lighting in the CelebA dataset. I believe that these differences are sufficient to test my final project because it tests the

boundaries of how well the classifiers have been trained to make accurate predictions, regardless of age, orientation, and lighting.

On the test set, the program did not detect one, singular face in 6 out of 30 images, so the performance of the classifiers will be based on the 24 images that were processed. Additionally, it should be noted that there was no provided metadata for my test set to describe whether a subject possessed any of the features. With some of the tested attributes being subjective – such as having a big nose or bushy eyebrows – some subjects were deemed as possessing certain attributes at my discretion.

The mustache classifier made a correct prediction on 79.17% of the test set. False positives occurred on 20.83% of the subjects, and no false negatives were predicted.

The big nose classifier accurately predicted whether subjects had a big nose 95.83% of the time. False positives occurred 4.17% of the time, and no false negatives were predicted.

The bushy eyebrows classifier produced correct predictions on 83.33% of the test. False positives occurred 4.17% of the time, and false negatives occurred on 12.50% of the subjects.

The open mouth classifier correctly predicted whether subjects had their mouth slightly open in 95.83% of the images. False positives occurred 4.17% of the time, and false negatives did not occur.

Two statistics that stood out to me were the high number of falsely predicted mustaches and the number of falsely non-predicted bushy eyebrows. For the mustache misclassifications, I believe this is most likely due to improper and inconsistent lighting

on the subject. The region of the face where a mustache would be expected is between the nose and upper lip, which both protrude from the surface of the face and can cause vast differences in lighting of the mustache region. I expect this to be the culprit, and predict that the results could be improved if the dataset had consistent lighting. The bushy eyebrow misclassifications were likely due to the non-vertical orientation of some subjects' faces. The training dataset featured subjects whose faces were vertically oriented, which resulted in the eyebrow classifier being trained on eyebrows which were almost entirely horizontal. Therefore, any of the subjects in my test set whose faces were not vertical would have their eyebrows classified while not being horizontal, which would make it more difficult for them to fit the correct prediction that the classifier was trained on. This could be improved by using a training dataset that includes subjects with different facial orientations. Another solution could be analyzing the orientation of each subject's face, and rotating the sample to vertically orient the subject.