Detecting Faces From Spreadsheet Data: Low Level Image Processing

Shifat E Arman
University of Dhaka
Tahmeed Abdullah
University of Dhaka
Course: CSE 402

Methodology

The following steps are followed:

1. Creating The Image Dataset.
2. Resizing The Images.
3. Preparing the Ground Truth from Resized Images.
4. Preprocessing The Data for Training.
5. Training The Dataset.

1. Creating The Image Dataset

Images of 14 persons were taken using a camera. Each person has 6 images. Hence the total number of images were 84.

Six categories of images are taken:

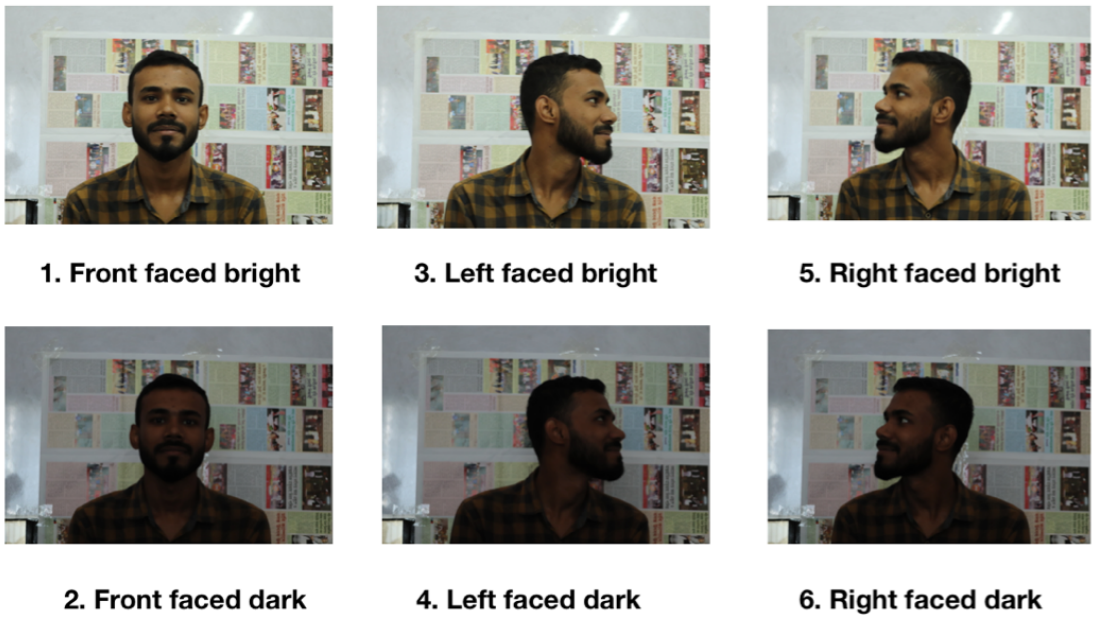
1. Front faced bright
2. Front faced dark
3. Left faced bright
4. Left faced dark
5. Right faced bright
6. Right faced dark

That means we have 3 viewpoints:

1. Front
2. Left
3. Right

And 2 illuminations:

1. Bright
2. Dark

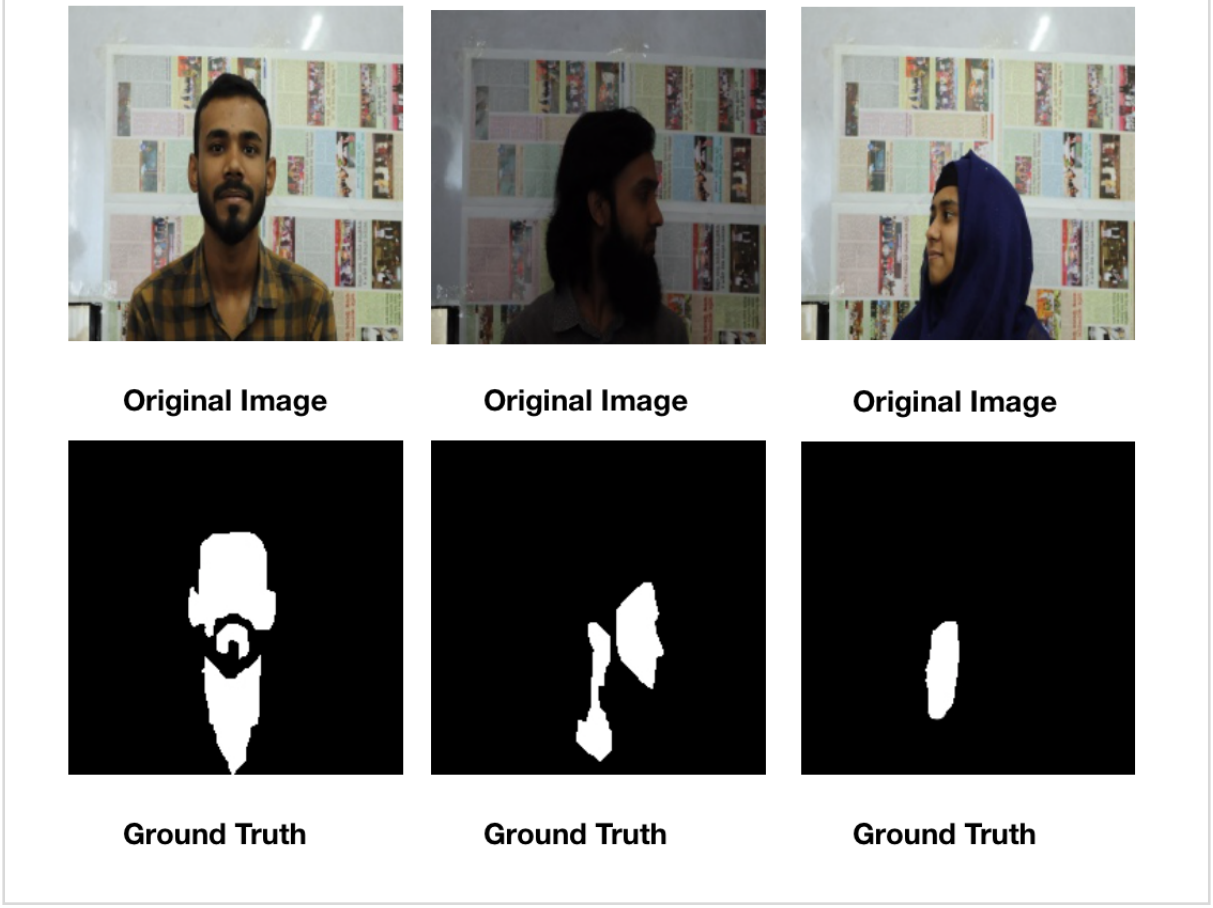


2. Resizing The Images

After capturing the images the first thing that was needed to be done was to resize the images. The images were initially of dimension 4000x6000. The problem if we keep the dimension 4000x6000 is that our model will take a lot of time to train. That is why we decided to resize them. Each image was resized to dimension 226x226. After resizing the number of pixels was reduced to almost 0.21% which is great for training the model.

3. Preparing The Ground Truth (How We Obtained The Labels?)

Images are annotated manually. The task was a bit tiresome. We have faced dilemma regarding whether we should or should not include beard and neck as a part of face, but eventually decided to omit them. We marked those part of the image which corresponds to face as white and the rest of the part of the image as black. Hence this annotated image gives us the ground truth. We labeled the part which are annotated as part of face as ‘true’ and the other parts of the image as ‘false’. Boolean datatype is chosen as weka prefers that the labels are kept boolean.



4. Preprocessing The Data for Training.

Each image has 3 channels-

1. Red
2. Green
3. Blue

Each channel has 226x226 pixels.

That means, in total, each image has 226x226x3=153228 pixels.

The pixels of each channel of the image is flattened and inserted into a spreadsheet.

At first the red channel is flattened. 226x226 intensity values of the red channel are inserted into the first column of the spreadsheet.

Then the green channel is flattened. 226x226 intensity values of the green channel are inserted into the second column of the spreadsheet.

After that the blue channel is flattened. 226x226 intensity values of the blue channel are inserted into the third column of the spreadsheet.

Lastly the ground truth image is flattened. 226x226 intensity values of the ground truth are inserted into the fourth column of the spreadsheet.

So, our spreadsheet has 4 columns for-

1. Intensity values of pixels of red channel.
2. Intensity values of pixels of green channel.

3. Intensity values of pixels of blue channel.
4. Labeling of the image from ground truth - face or not.

Each row corresponds to a pixel of an image. First 226x226 rows are pixels of the first image. The next 226x226 rows are pixels of the second image. In this way, pixels of all the images are arranged in rows. It has 226x226x84=4290384 rows for pixels.

Hence, we have successfully converted our image data into a spreadsheet.

That means we are deciding whether a particular pixel is a face or not from intensities of a single pixel in 3 channels.

5. Training The Dataset.

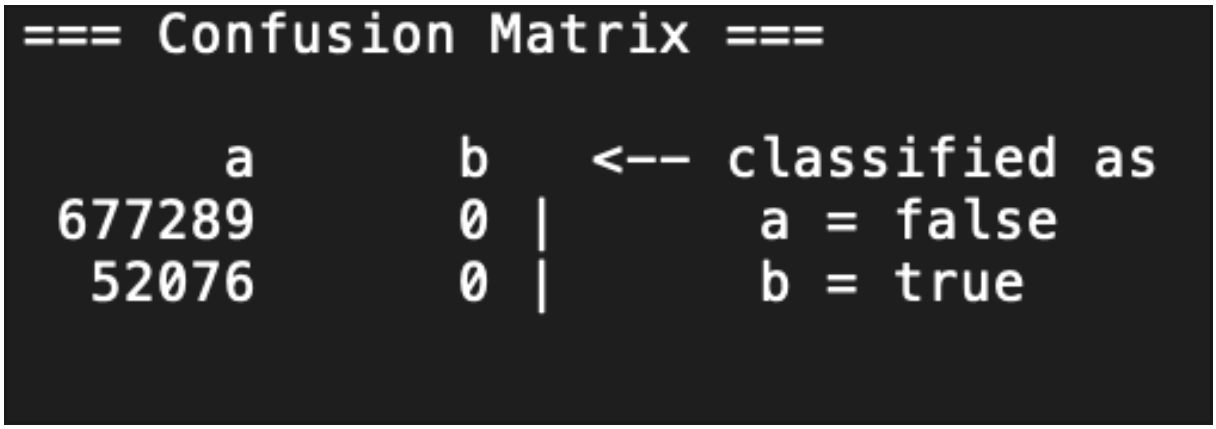
Dataset is trained using logistic regression, K-Nearest Neighbour classifier, SIFT.

Results

Result obtained from SIFT:

Result obtained from Logistic Regression:

We trained the model using logistic regression in weka and got some really bad result. The model classified every pixel as 'not face' as it is a safer option of the two.



Result obtained from KNN

The data was also trained using K nearest Neighbour classifier using number of neighbours=5 and got the following result.

```
Confusion Matrix
----------------
[[591095   6740]
 [ 43517   2206]]
             precision    recall  f1-score   support

      False       0.93      0.99      0.96    597835
       True       0.25      0.05      0.08     45723

avg / total       0.88      0.92      0.90    643558
```