Introduction to Computer Vision & Image Processing
Project 1

NOTE: I made a hypothesis that the Harris corner detector might provide better results in comparison to using SIFT features, Which later on proved to be wrong as I'm only getting an F1 score of (0.55-0.60). Whereas I was getting a score of (0.6-0.75) using SIFT features. However, I decided to stick with the Harris corner detector thinking I'll be able to optimize it. But soon time went fast & to meet the deadline I stayed with harris detector as I was afraid of breaking my code in lieu of shifting to use SIFT again. (i.e. on submission day).

Type of features used:

Harris corner detector: These detectors try to locate corners using gradient in x AND y directions but minimal in alone x OR y-direction.

Hypothesis: Having sub-optimal results for SIFT, I decided to switch to the Harris detector thinking that edges might provide a better intuition about the characters as the given sample had no scaling issues (i.e. we were provided enrolled characters of similar scale & we could've also resized the characters). But edges and corners were are to focus. So I decided to try Harris & HOG (Which I couldn't implement because of the paucity of time). By the time I ended with Harris, there were only 1-2 days left for the submission (with other course projects/home works due too).

So I decided to play safe and submit the Harris corner detector features, which was providing me a score of (0.55 to 0.60) for different hyperparameters & settings.

Connected component labeling: I accurately & correctly implemented the CCA/CCL using the two-pass method. Before that, I thresholded and converted all pixels to 0 or 1 depending on the pixel's strength (converting the image into a BLOB.)

For the 1st pass, I checked the pixel's background, north & west label's (value). And designated the class of the pixel accordingly.

For the 1st pass, I maintained an array 'conf' which noted all the conflicts & equivalencies (Parent -> Child links) for the 1st pass.

For the 2nd pass, I just had to seek back to conflicts & equivalencies using 'conf' array & change all child's link labels to their final parent's link.

To calculate character positions, I iterated over classes to get the maximum & minimum for every respective class. The rest was straightforward.

Recognition: For this part, I used previously calculated features & character positions (that I stored in the ./data folder). I then extracted the blobs for each class using character positions. I also calculated the Harris corners for the extracted blob. Now using the blob's harris corners I compare it with the harris values of each enrolled image. I compare the respective histograms using CROSS-CORRELATION and choose the best score & prediction accordingly.
I repeat this for all blobs to get the final results.