University of Toronto

CSC321 Neural Network

Assignment 1

Face Recognition and Gender Classification with KNN

Author:
Tzu Yao Chien
998758759
g2samuel

Supervisor:
Dr. Michael Guerzhoy

February 3, 2016

Contents

1 Part 1

Figure 1: Uncropped Photo With Lots Noise.

Figure 2: Uncropped Photo With Tilted Face.

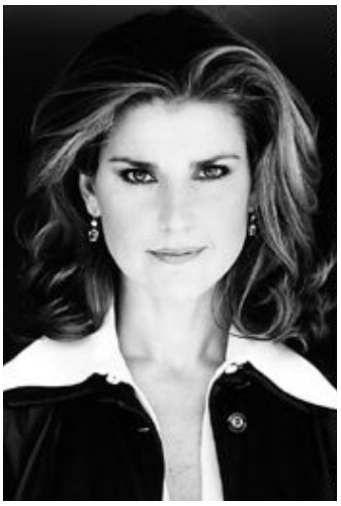

Figure 3: Uncropped Photo With Gray scale.

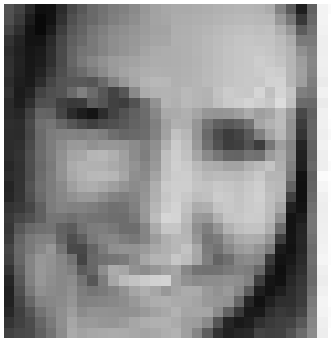

Figure 4: Cropped and Gray Scale of the Above Figure

Figure 5: Cropped and Gray Scale of the Above Figure

Figure 6: Cropped and Gray Scale of the Above Figure

After downloading the pictures from the address provided in faces.txt file, I confirm the correctness of the photos using sha256 hash. This ensures the correctness of the photos matches the exact photos thats been recorded in the database. These photos are different positions of different famous celebrities. The uncroppped pictures are all different. The difference includes colors, size, position of celebrities and all the noise in the pictures. First of all, the pictures comes in color and gray.

The size of the photo also varies as well as the proportion of the celebrities faces in the pictures are also different. Moreover, the faces are also taken from variety of angles. As shown in the examples, there are a lot of different background and watermarks in the photo. Following the suggested cropped region, I was able to cropped the faces from the original photo and change it to gray scale then save it as 32x32 size. The cropped photo are mostly consistent where none of the faces are cut in half. However, due to the angle the photos are taken, the cropped images do not aligned with each other. Although I have resized the images to make all the cropped images similar, it is clear some photos are more blurry than the others. This is due to the fact that all the photos have different size of suggested cropping region. The photos are acceptable for further classification.

2 Part2

I have downloaded and cropped 130 images for each of these actors 'Gerard Butler', 'Daniel Radcliffe', 'Michael Vartan', 'Lorraine Bracco', 'Peri Gilpin, and 'Angie Harmon. In other words, I just randomly picked first 100 photos of each actor as the training set, then use the next 10 faces as validation set, and the next 10 faces as test set. Step one, the goal is to have a dedicated training, validation, and test set to make sure our training for k does not overfit the validation set such that test set performance decrease. The test set should be independent of the training of validation percentage. The main drawback of this approach is that the variance for this approach might be high due to the fact that test set never changes. In other words, the more advance way to evaluate the performance of k should be using 10-fold cross validation where the test set, validation set, and training set changes. The result should be the average of 10 folds.

3 Part3

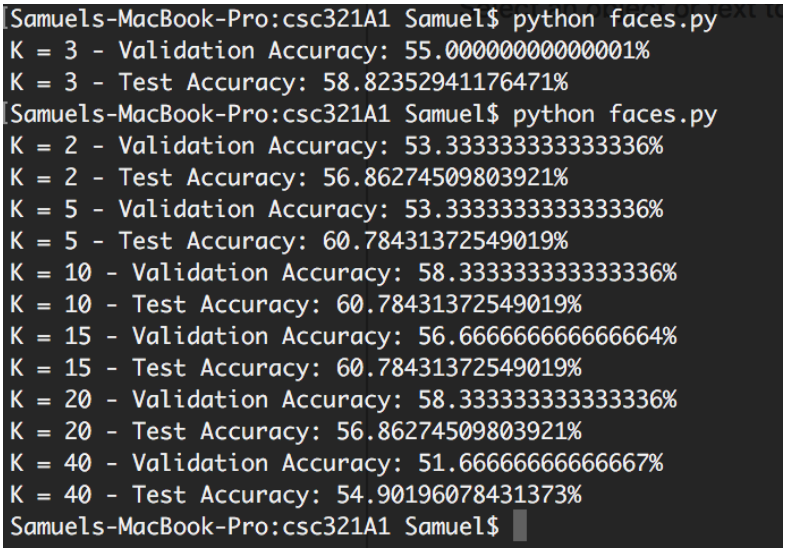

Figure 7: KNN accuracy of Validation and Test Set

In this section, I have tested a variety of k including 2, 5, 10, 15, 20, 40 and as shown in the performance output image, the accuracy of each validation and test set is recorded for each k value.

Set	K=2	K=5	K=10	K=15	K=20	K=40
Validation	53.3%	53.3%	58.3%	56.7%	58.3%	51.7%
Test	56.9%	60.8%	60.8%	60.8%	56.9%	54.9%

Figure 8: Table: Accuracy of the Validation and Test Set

From the table, we can clearly observe the trends where the accuracy of both set increases until the K=10 then it slowly decrease after the threshold. In other word, it is clear that K=10 is the best K. Keep in mind that this particular approach might have high variance since the validation and test set is constant. The best way to obtain the result with least variance is using 10-fold cross validations average result for each of the K value.

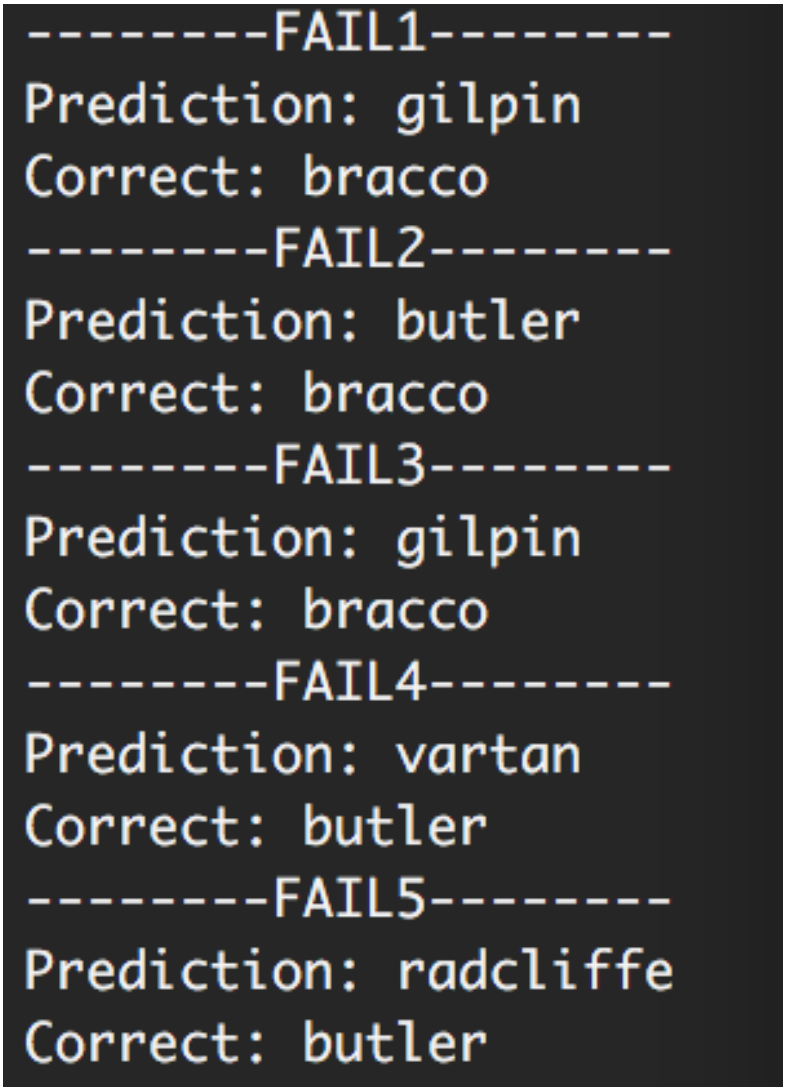

Figure 9: Failure Cases of KNN

Here I have also shown the first 5 failure case where KNN recognition is not correct as well as the the 5 Nearest Neighbors Face. As shown in the picture, I have observe that the KNN algorithm does not perform well when the pictures have similar color intensity. In addition, the tilted face or face expression can also influence the algorithm.

Figure 10: Target Image of Bracco

Figure 11: 1st Neighbor: Bracco

Figure 12: 2nd Neighbor: Gilpin

Figure 13: 3rd Neighbor: Butler

Figure 14: 4th Neighbor: Gilpin

Figure 15: 5th Neighbor: Gilpin

Figure 16: Target Image of Bracco

Figure 17: 1st Neighbor: Vartan

Figure 18: 2nd Neighbor: Bracco

Figure 19: 3rd Neighbor: Gilpin

Figure 20: 4th Neighbor: Radcli.

Figure 21: 5th Neighbor: Butler

Figure 22: Target Image of Bracco

Figure 23: 1st Neighbor: Gilpin

Figure 24: 2nd Neighbor: Radcli.

Figure 25: 3rd Neighbor: Gilpin

Figure 26: 4th Neighbor: Gilpin

Figure 27: 5th Neighbor: Gilpin

Figure 28: Target Image of Butler

Figure 29: 1st Neighbor: Butler

Figure 30: 2nd Neighbor: Butler

Figure 31: 3rd Neighbor: Vartan

Figure 32: 4th Neighbor: Vartan

Figure 33: 5th Neighbor: Vartan

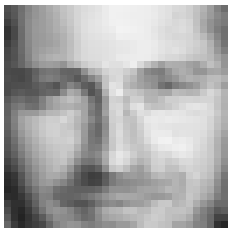

Figure 34: Target Image of Butler

Figure 35: 1st Neighbor: Radclif.

Figure 36: 2nd Neighbor: Gilpin

Figure 37: 3rd Neighbor: Butler

Figure 38: 4th Neighbor: Radclif.

Figure 39: 5th Neighbor: Vartan

4 Part4

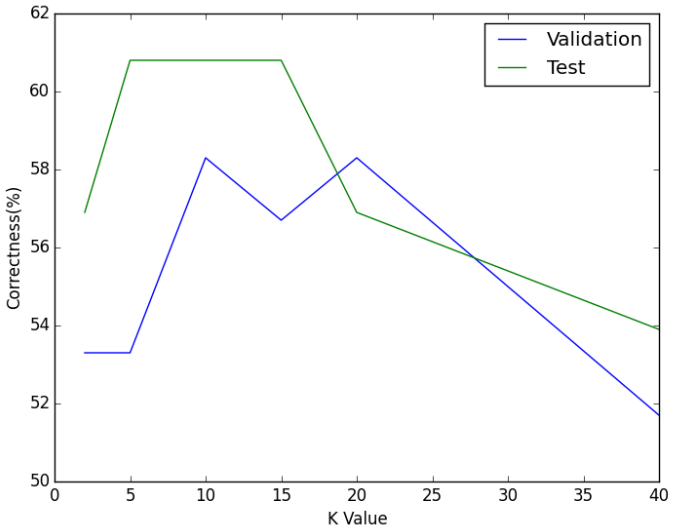

Figure 40: Performance Chart of Test and Validation Set

As shown in the chart, I have plotted the performance from part 3 and we can observe the trend easily. We can clearly observe the trends where the accuracy of both set increases until the K=10 then it slowly decrease after the threshold. The reason K=2s accuracy is less than K=10 is because when K is too small, it can easily be influence by the outlier. On the other hand, when K is too big as K = 20, the result decrease significantly because the neighbor circle is too big and it doesnt provide the fine grouping.

5 Part5

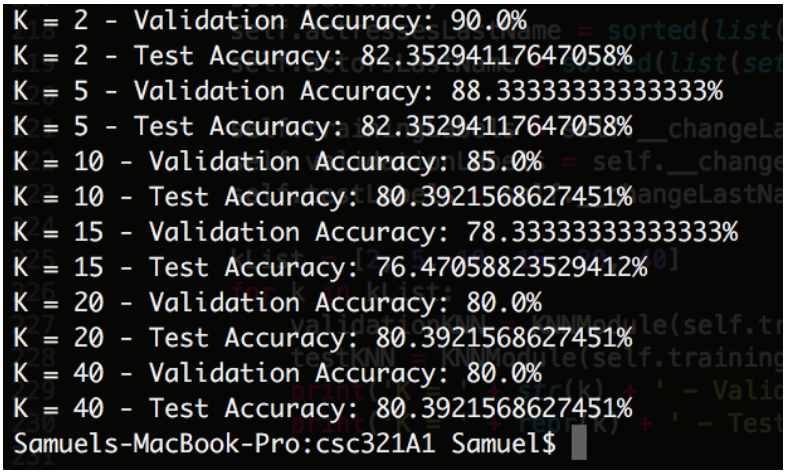

Figure 41: Performance Output of Test and Validation Set

In this section, I have tested a variety of k including 2, 5, 10, 15, 20, 40 and as shown in the performance output image, the accuracy of each validation and test set is recorded for each k value.

Set	K=2	K=5	K=10	K=15	K=20	K=40
Validation	90.0%	88.3%	85.0%	78.3%	80%	80%
Test	82.4%	82.4%	80.4%	76.5%	80.4%	80.4%

Figure 42: Table: Accuracy of Test and Validation Set

From the table, we can clearly observe the trends where the accuracy of both set started high at the K=2 then it slowly decrease after the threshold. The difference of part 5 and part 3 is that there is less classification and females generally look more alike, in other words, the data thats closer together are likely in the same gender. Therefore, due to less classifications and nature of the gender images, it require less neighbor to determine the labels. In other word, it is clear that K=2 is the best K. Keep in mind that this particular approach might have high variance since the validation and test set is constant. The best way to obtain the result with least variance is using 10-fold cross validations average result for each of the K value.

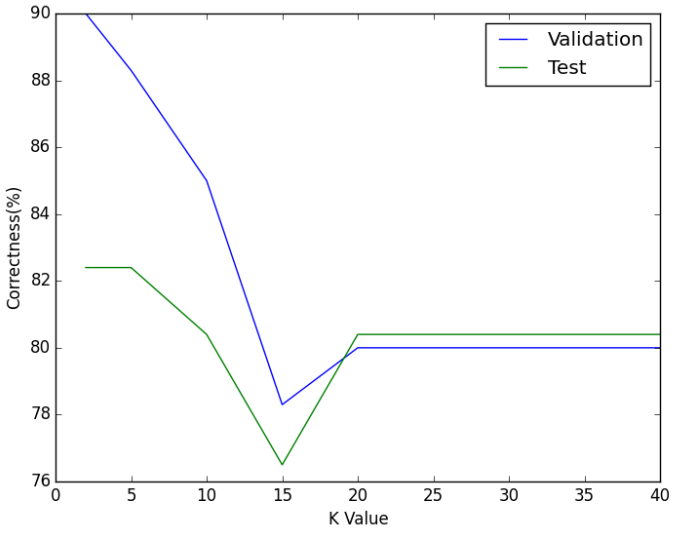

Figure 43: Chart: Accuracy of Test and Validation Set

As shown in the chart, I have plotted the performance from part 5 and we can observe the trend easily. We can clearly observe the trends where the accuracy of both set started high at K=2 then it slowly decrease after the threshold. The reason K=2s accuracy is high is because there are only two labels and photo with same genders looks like each other. In other words, there is not many outliers which results high performance for small set of K. On the other hand, when K is too big as K = 20, the result decrease significantly because the neighbor circle is too big and it doesnt provide the fine grouping.

6 Part6

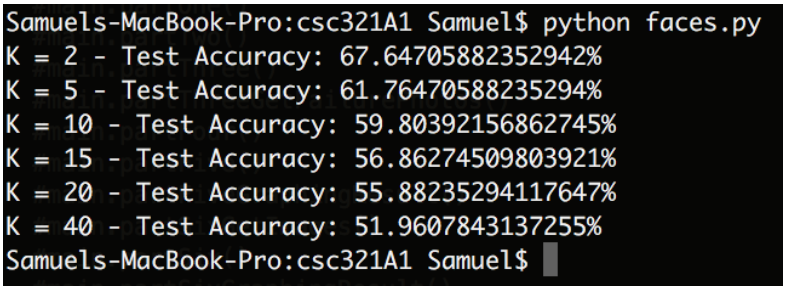

Figure 44: Performance Output of Test and Validation Set

In this section, I have tested a variety of k including 2, 5, 10, 15, 20, 40 and as shown in the performance output image, the accuracy of the people outside training set is recorded for each k value. (I take one of each persons face in allFaces.txt)

Set	K=2	K=5	K=10	K=15	K=20	K=40
GenderSet	67.6%	61.8%	59.8%	56.9%	55.9%	52.0%

Figure 45: Table: Accuracy of Test and Validation Set

From the table, we can clearly observe the trends where the accuracy of both set started high at the K=2 then it slowly decrease after the threshold. The difference of part 5 and part 6 is that the test set are the people whos not in the training set. In other words, it only works when there are two individual with same gender who looks alike. Therefore, due to the fact that all the celebrities have a pretty unique face features, the data is pretty scattered resulting better performance for smaller K. Through the test, it is clear that K=2 is the best K. Keep in mind that this particular approach might have high variance since the validation and test set is constant. The best way to obtain the result with least variance is using 10-fold cross validations average result for each of the K value.

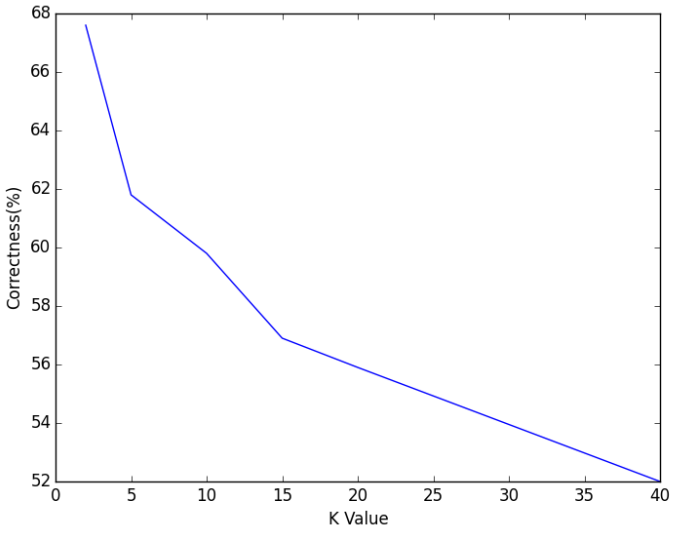

Figure 46: Chart: Accuracy of Test and Validation Set

As shown in the chart, I have plotted the performance from part 6 and we can observe the trend easily. We can clearly observe the trends where the accuracy of the set started high at K=2 then it slowly decrease after the threshold. The reason K=2s accuracy is high is because there is little variance in euclid distances since most of the photo do not look like the distinct celebrities faces. In other words, the data is very clustered together which results high performance for small set of K. On the other hand, when K is too big as K = 20, the result decrease significantly because the neighbor circle is too big and it doesnt provide the fine grouping.