Assignment 2 | CS5332 Biometric Authentication

Sanjay Saha (A0187044A) | Sem2 AY2019/2020 | SoC, NUS
sanjaysaha@u.nus.edu

Question 1

Print the confusion matrix and overall accuracy for the identifiers in Tasks 3, 4, 5 and 7 above.

Task 3

Use scikit learn to train kNN classifiers with this training set by varying the value for k (k=3,5,7). Evaluate the classifier with the test dataset by computing the accuracy for each classifier. Plot k value vs accuracy. Calculate the confusion matrix for k=5.

k (n_neighbors)	***Accuracy***
k=3	0.7500
k=7	0.6583
k=5	0.6700

Table 1: kNN performance (k=3,5,7) with PCA features.

Accuracy values for k=3,5,7 for the k Nearest Neighbors classifiers with PCA features (D=30) are given in Table 1. A scatterplot of the k-values vs accuracy is given in Figure 1. The figure contains results (accuracy values) for more k-values (3, 5, …, 15). It is evident that we can get the best performance out of a k Nearest Neighbors classifiers using PCA features when k=3. And, confusion matrix for k=5 is given in Figure 2.

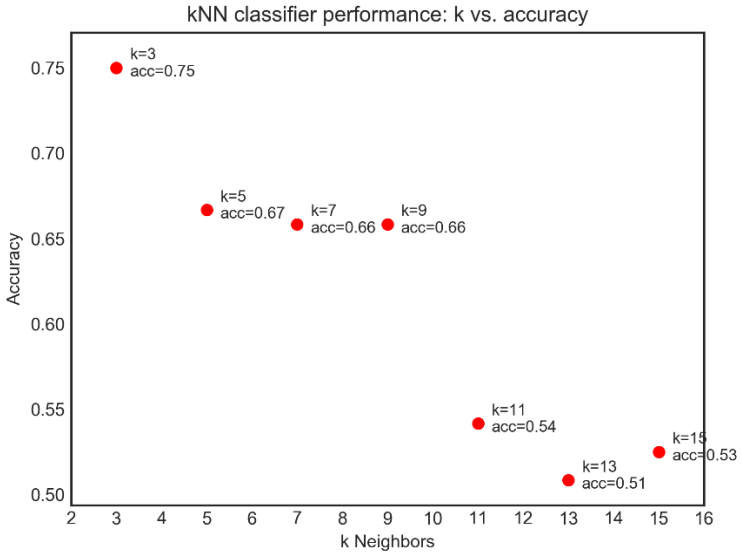


Figure 1: k vs accuracy with PCA features.

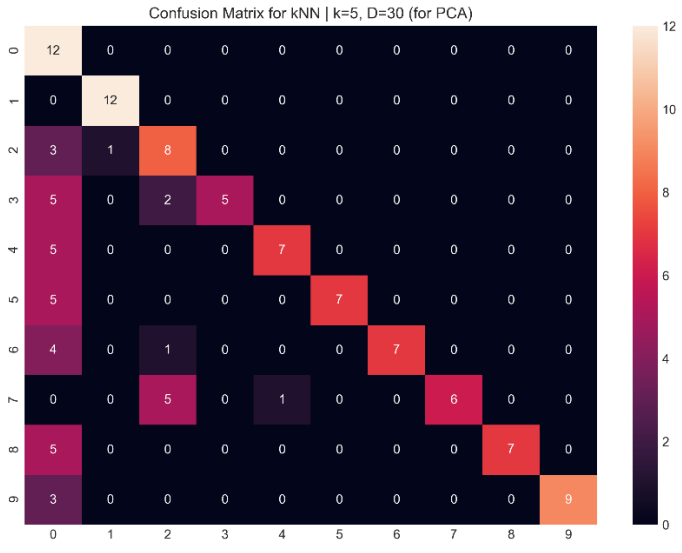


Figure 2: Confusion matrix for kNN with PCA features (k=5).

[continued to next page…]

Task 4

Use scikit learn to train a Random Forest classifier for the same training set. Try changing different values for the parameter n estimators in the Random Forest. Visualize the change in accuracy when n_estimators is changed as above. Calculate the confusion matrix for n_estimators = 100.

# of trees (n_estimators)	***Accuracy***
n=25	0.97
n=50	1.00
n=100	0.99
n=200	1.00

Table 2: Random Forest performance (n_est=25,50,100,200) with PCA features.

Accuracy values for n_estimators=25, 50, 100, and 200 for the Random Forest classifier with PCA features (D=30) are given in Table 2. A scatterplot of the n Estimators vs accuracy is given in Figure 3. The figure contains results (accuracy values) for n_estimators values (25, 50, 100, 200). All the different Random Forest classifier perform better than kNN classifiers. We can choose n_estimators=50 as it gives the highest accuracy (1.00). Confusion matrix for k=5 is given in Figure 4.

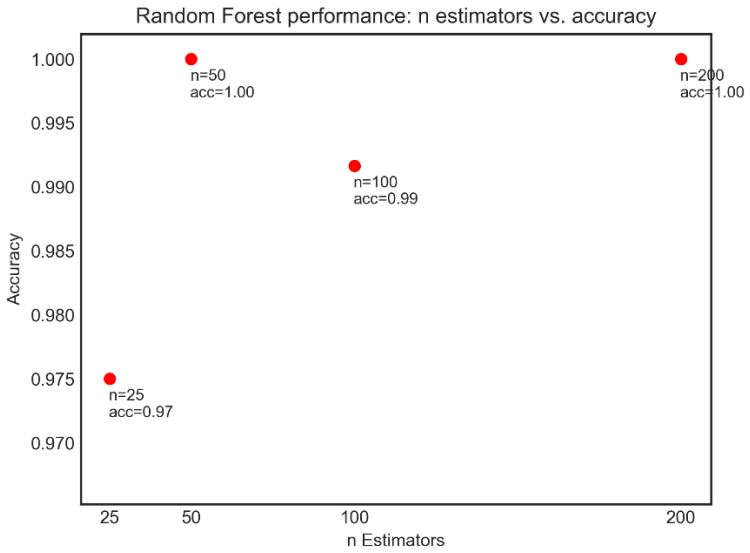


Figure 3: n Estimators vs accuracy with PCA features.

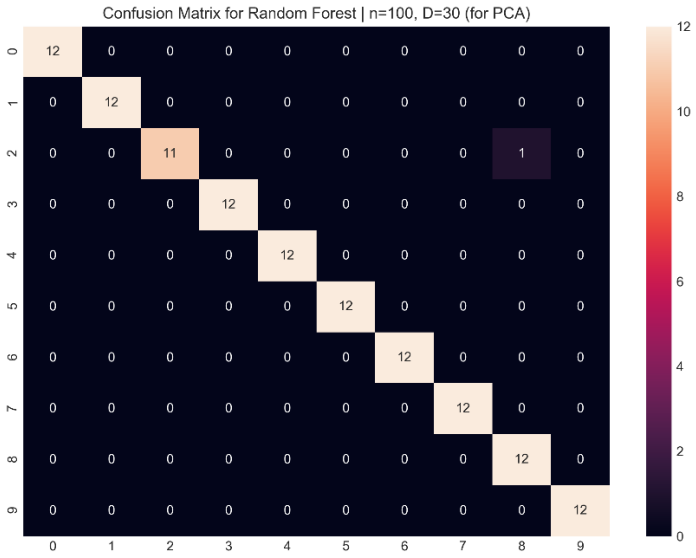


Figure 4: Confusion matrix for Random Forest + PCA features (n=100).

Task 5

Report on the other feature extraction methods you tried for F() as instructed in 3.1. Which classifiers did you use with which features? What are their accuracies?

Further experiments other than with PCA features were done on the face classification problem with the following features: LBPH, SIFT, SURF, and CNN (Convolutional Neural Network). Out of all these other features, LBPH and CNN features perform best with accuracy 1.00. Other than these two, SIFT and SURF features also perform very well with over 97% accuracy. For classification, Support Vector classifier has been used with PCA, SIFT, and SURF features other than *k Nearest Neighbors* and *Random Forest*. A list of all the results from different features and classifiers are given in Table 3 (next page).

For the given dataset, PCA features with Random Forest classifier, LBPH features, and SURF features with kNN given 100% accurate performance on test data. SIFT features with Support Vector and kNN classifiers, and PCA features with Support Vector classifier also works very well with around 99% accuracy on the test data.

[continued to next page…]

For CNN features, the python library face_recognition was used. Classification using CNN features show 100% accuracy for all the classifiers.

Feature	*Classifier*	*Parameters*	*Accuracy*
PCA	K Nearest Neighbors	k=3, D=30	0.7500
		k=7, D=30	0.6583
		k=5, D=30	0.6700
		k=9, D=30	0.6583
	Random Forest	n_estimators=50, D=30	1.0000
		n_estimators=25, D=30	0.9700
		n_estimators=100, D=30	0.9900
	Support Vector	C=1, kernel='linear'	0.9917
LBPH	Distance measure	Default (OpenCV)	1.0000
	K Nearest Neighbors	k=3, radius=9	0.7333
		k=5, radius=9	0.6917
		k=7, radius=9	0.6333
	Random Forest	n_estimators=25, radius=15	0.9250
		n_estimators=50, radius=15	0.9667
		n_estimators=100, radius=15	0.9917
		n_estimators=150, radius=15	1.0000
SIFT	Support Vector	kernel='rbf', C=10, gamma=0.00001	0.9833
	K Nearest Neighbors	k=3	0.9833
		k=5	0.9750
		k=7	0.9667
	Random Forest	n_estimators=50	0.9420
		n_estimators=100	0.9667
SURF	Support Vector	kernel='rbf', C=10000, gamma=0.1	0.9750
	K Nearest Neighbors	k=3	0.9833
		k=5	0.9917
		k=7	1.0000
	Random Forest	n_estimators=25	0.9667
		n_estimators=50	0.9750
		n_estimators=100	0.9583
CNN	Support Vector	kernel='linear', C=1	1.0000
	K Nearest Neighbors	k=3	1.0000
		k=5	1.0000
		k=7	1.0000
	Random Forest	n_estimators=25	1.0000
		n_estimators=50	1.0000
		n_estimators=100	1.0000

Table 3: Performance of classifiers for face identification

[continued to next page…]

Task 7

Extract gait features as explained above using the code given to you. And train two classifiers as above based on these features. Report on accuracy values and how they vary based on the parameters, similar to Tasks 3 and 4.

Classifier	***Parameters***	***Accuracy***
Decision Tree	default	0.8000
Random Forest	n_estimators=25	0.9750
	n_estimators=50	0.9625
	n_estimators=100	0.9750
	n_estimators=150, 200	0.9875
K Nearest Neighbors	k=3	0.9625
	k=5, 7	0.9125
	k=9	0.9000
Support Vector	kernel=rbf	0.7125
	kernel=linear	0.9500

Table 4: Performance of classifiers for gait identification

Experiments on the gait data show that Random Forest with 150 and 200 trees give the highest (98.75%) accuracy compared to other classifiers (kNN, Support Vector, and Decision Tree).

Question 2

Compare the results between kNN and Random Forrest. Which is better?

For face identification, Random Forest performs better than kNN with PCA features. Even for different values of D (10, 20, …, 100) in PCA, Random Forest performs better than kNN. The difference in their performance is clearly noticeable from the experiment results presented in the following table:

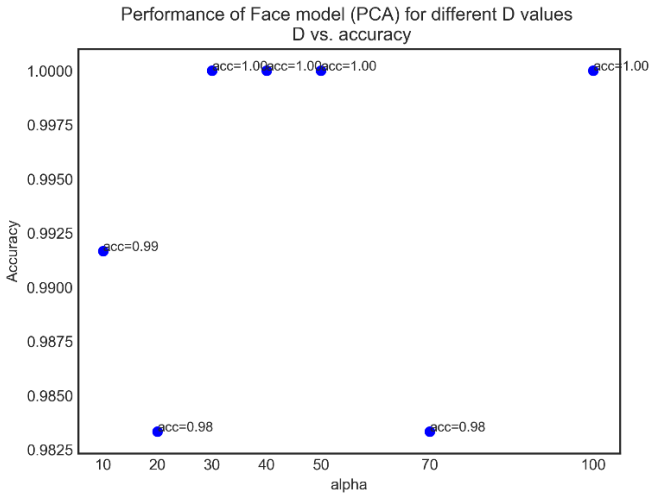
K Nearest Neighbors	k=3, D=30	0.7500
	k=7, D=30	0.6583
	k=5, D=30	0.6700
Random Forest	n_estimators=50, D=30	1.0000
	n_estimators=25, D=30	0.9700
	n_estimators=100, D=30	0.9900

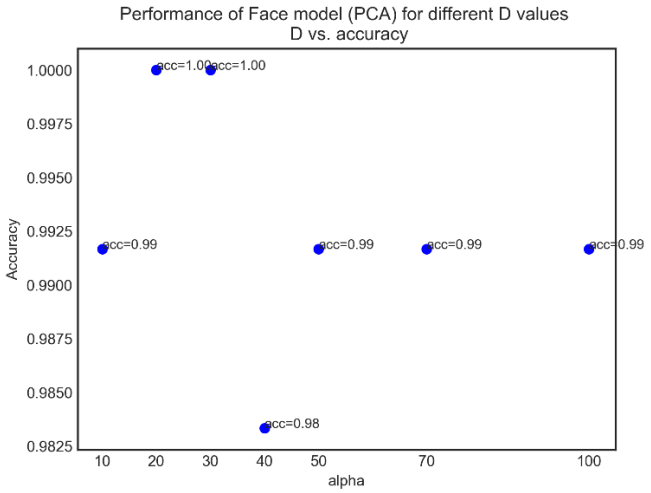
Question 3

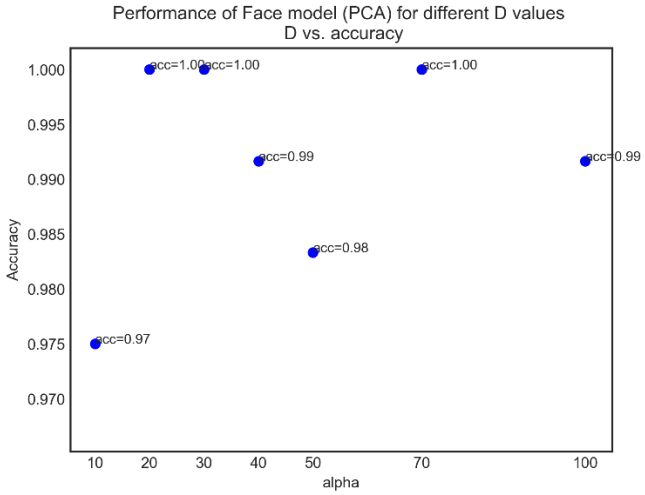
Select one of the classifiers. Using this for the PCA-based Identifier, vary D as follows: D = {10, 20, 30, 40, 50, 70, 100}. For each value of D, compute the classification accuracy. Plot accuracy versus D. Comment on this plot.

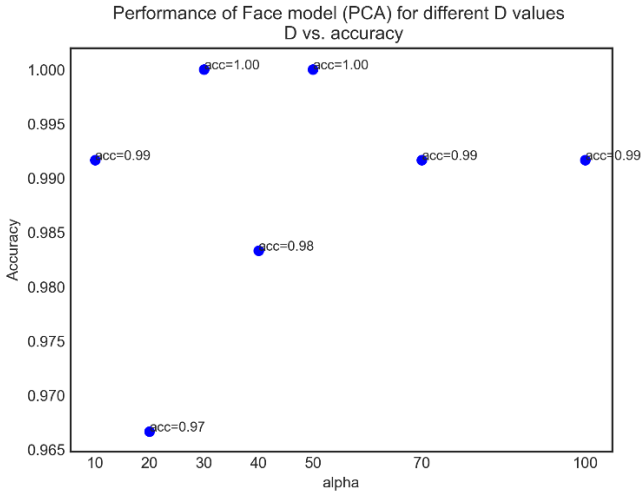
Random Forest (n_estimators=50) classifier has been selected as this works better than other classifiers for the PCA-based identifier. Results from Random Forest varies from each execution of the program. Hence, different executions of the same program with same parameters can give different results. The D vs. accuracy plot for is given in the following images:

[continued to next page…]









These four plots show the different accuracy values for different D values in four experiments. From these plots it is observable that we can often get 100% accuracy when D is 30, or 50. Event for other values like 70 and 100, the accuracy is near to 100%.

Question 4

Pick the best value of D for the PCA-based Identifier (the one that gives the highest classification accuracy). Compute the Confusion Matrix. What do you observe? Is there a row or column of non-zero values? What is this telling you? Show the relevant facial images to explain.

The best value for D is 30. From multiple runs of the experiment, it was observed that for D values other than 30, accuracy was fluctuating from 0.9750 to 1.00. However, for D=30, accuracy was 1.00 for most of the runs. Also, lower value of D will cause less computation, which would be helpful for faster training and classification.

As the best accuracy is 100%, we won't have any misclassified face image to examine for D=30. Hence, we don't have any non-zero values other than in the diagonal.

However, we can select a classifier with less accuracy (Random Forest with D=40 has 1 misclassified image). Here, two confusion matrices are shown (left: D=30, accuracy=1.00 and right: D=40, accuracy=0.9917). The confusion matrix on the right shows one face image that was misclassified: an image from class-label '0' was predicted as class-label '2'.

[continued to next page…]

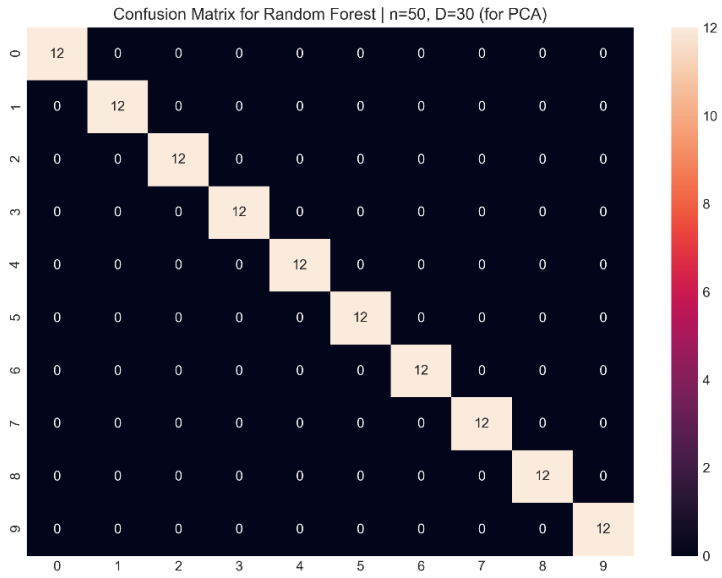


Table 5: Confusion Matrix for D=30 (RF, n_est=50)

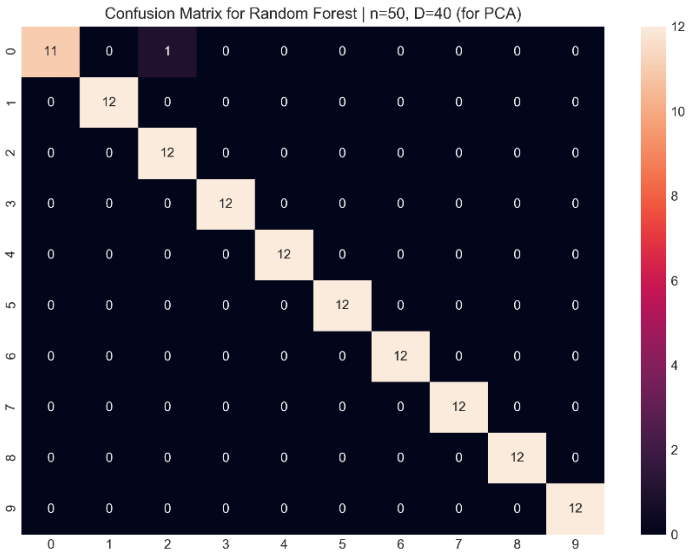


Table 6: Confusion Matrix for D=40 (RF, n_est=50)

Question 5

Let α = 0.1, 0.2, …, 0.9. Re-train your fusion-based Identifier and re-calculate its accuracy for each α. Plot accuracy versus α for different. What do you observe?

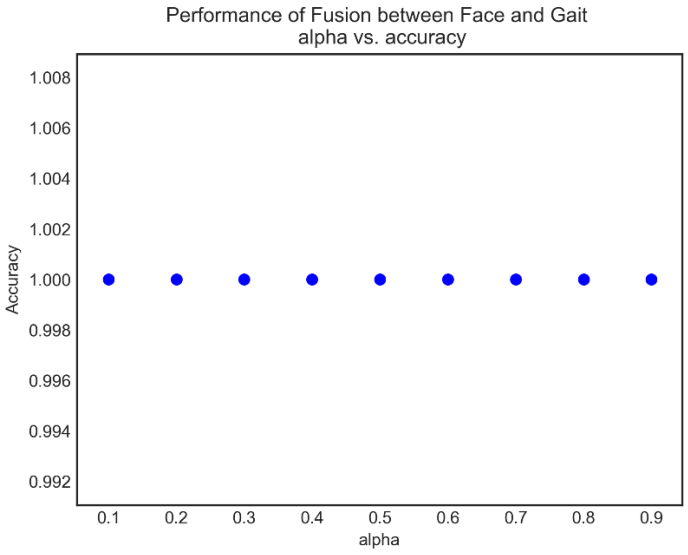


Figure 5: Fusion of best Face & Gait models

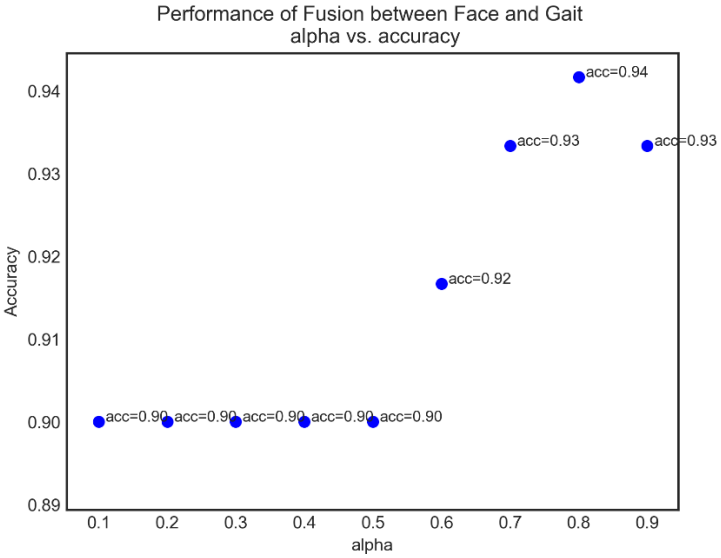


Figure 6: Fusion of weaker Face and Gait models

When the best models for face and gait were selected for fusion, the results from the fusion were 100% accurate for all the values of α (0.1, 0.2, to 0.9). The models which were used in this fusion were:

- For face: Random Forest with n_estimators = 50 (individual accuracy: 100%)
- For gait: Random Forest with n_estimators = 150 (individual accuracy: 98.75%)

Both these models perform with very high accuracy even when individually used for classification. Hence, there was no misclassification for any of the test images in the fusion. The results (α vs. accuracy) are plotted in Figure 5.

[continued to next page…]

To check the effect of fusion, two weaker models were also used in a fusion. These weaker models were:

- For face: K Nearest Neighbors with n_neighbors = 9 (individual accuracy: 65.83%)
- For gait: Decision Tree (individual accuracy: 80.00%)

The fusion of these two weaker models gave much higher accuracy than any of the two individual models. For $\alpha = 0.8$, the higher accuracy (94%) was recorded from the fusion. The results of this experiment (α vs. accuracy) are plotted in Figure 6.

Question 6

Compare between gait and face based identifiers. Which performs best? Does that mean one is superior to the other? Are there situations where this might be different? What are the short comings of the two modalities?

For the given datasets, face based identifiers perform the best (as can be observed from the results presented in the table above). However, this does not make the face based identifiers superior than the gait based identifiers for any general use cases. Performance of these classifiers mainly depend on the acquisition of good quality and quantity of data.

In a Continuous Authentication scenario for mobile devices, when the user is walking with the phone in her pocket, the data that is possible to acquire for authentication is gait as capturing face image is not possible in this case. Again, when the user is looking at the phone screen while in a sedentary position, face authentication would be the choice for this case. So, it depends on the context. However, a fusion of multiple modalities can help to gain a better performance.

Shortcomings of face:

- Face as a biometric modality is less unique, easier to get misclassified by face classifiers than fingerprint classifiers.
- Not permanent over longer period of time. Face of a person changes with age.
- Performance of face classifiers still are behind iris and fingerprint classifiers.
- Acquiring face image can be challenging depending on the lighting condition (at night/dark rooms), camera position (surveillance cameras), etc.

Shortcomings of gait:

- Gait is also not unique, different people can have similar gait.
- It is even less permanent than face. Stride length, frequency can also change with age, or people may walk faster/slower than usual depending on situation.
- Performance of gait classifiers are also not in the best category.
- Gait data is better in quality when collected from on-body sensors like the accelerometers on phone or smartwatch. This option might not be available always causing less availability of data. Visual gait is another option. However, would require clean sequential images which is a great challenge to collect when in wild.

Question 7

How does fusion of different modalities help overcome issues identified in Q6.

Each biometric modality has shortcomings in different types. When one modality is not usable, any other modality which can provide the service can come into play. For example, when face model is unable to perform well (due to dark environment, poor performance of model etc.) fingerprint model can work as the primary classifier. Again, when we need a touchless biometric modality fingerprints can be unavailable, in this case face can be the primary modality.

For the given problem, we had two modalities: face and gait. They are very different kinds of modalities in terms of the nature of data: one gives face images, other provides temporal signal of accelerometer. From the results presented in Question 5, it is clear that fusion of these two modalities help immensely to improve the classification performance. A similar fusion can be applied to resolve the issues discussed in Question 7 as well:

- When the face classifier is not confident enough of a prediction, the gait classifier can help predicting the correct identity.
- Although both modalities are not highly unique, neither their individual performances are compared to iris or fingerprint. However, a fusion of the weaker face and gait models (results in the answer to Question 5) show that accuracy can reach significantly higher with a fusion.

---------- End of report ----------