

Biometrics System

Introduction

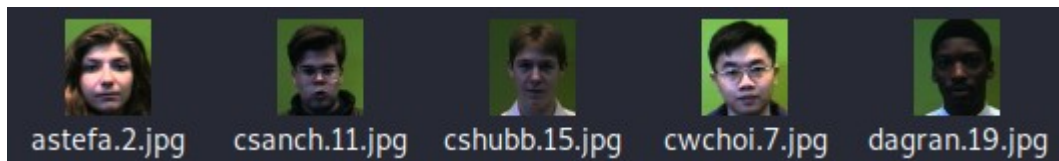
The goal of the project is to configure a biometrics system (face recognition) with the best configuration possible to a given data-set so that for future use and testing it's completely ready to use.

Data-set

The data-set is taken from (<https://cswwww.essex.ac.uk/mv/allfaces/index.html>) and a subset of images has been used for time reason.

The data-set is composed by a gallery and impostors in two separate folders and the genuine probes are obtained from the gallery, taking random images from different people (one per person).

Each persona in the data-set (gallery and impostors) has 20 images labeled with a name and an index separated with a '.' (dot).



[img1 - pattern of people name]

The impostors have the same pattern of the name as the data-set, the only difference is that it has 'IMP' at the beginning of the file name to identify that that person is an impostor and needed for the ground truth.



[img2 - Impostor's name pattern]

There are two types of data-set for testing, the first one is well made where people have a set of similar images of themselves with little variation of face expression and a second data-set where images of people are very different and they vary in age, illumination, and rotation.

The first data-set gallery is composed by ten female and ten male people; on the other hand, the impostors are composed by nine female and five male people.

All the people in the data-set have slightly different facial expressions.

The second data-set is composed by all male people. 12 people in gallery and 12 impostors.

Functions implemented

CreateGallery

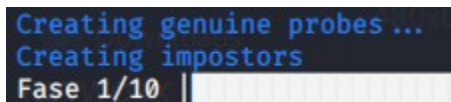
Parameter: path

This function is created to dynamically populate the probes, depending on the path it takes, path of gallery or path of the impostors directory.

From the given path it take a certain amount of images (one per person), the amount of the number of images to take is initially set and hard coded.

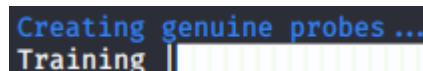
The main utility of this function is to create dynamically a potential infinite number of testing probes by combining genuine probes from gallery and impostors.

With this function is possible to easily create dynamic open set and closed set.



```
Creating genuine probes ...  
Creating impostors  
Fase 1/10 |
```

[img3 – Creation of an open set]



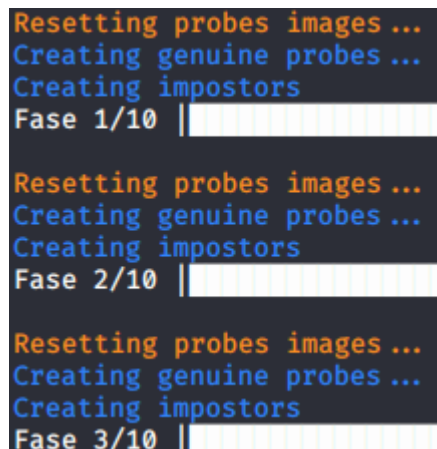
```
Creating genuine probes ...  
Training |
```

[img4 – Creation of an closed set]

ResetGallery

This function simply restore the images from the probes folder to the respective gallery or impostors folder.

Usually used before populating the probes while we want to maintain a constant and predefined number of probes, also to create other test probes.



```
Resetting probes images ...  
Creating genuine probes ...  
Creating impostors  
Fase 1/10 |  
  
Resetting probes images ...  
Creating genuine probes ...  
Creating impostors  
Fase 2/10 |  
  
Resetting probes images ...  
Creating genuine probes ...  
Creating impostors  
Fase 3/10 |
```

[img5 – Creation of open sets dynamically for testing]

DetectDatasetImgSize

This function is needed to detect if there are various type of image size.

While the system works better if all images in the data-set are the same size this function detect the most images with the same size and tries to adapt the system for them.

```
Checking dataset images size...  
Different types of image sizes has been found! it's suggested to have images of same size.  
200×180: 388 images  
600×480: 14 images  
System might be optimized for images of size 200×180
```

[img5 – Test with different size of image in gallery]

CalculateScaleFactor

This value specifying how much the image size is reduced at each image scale.

This scale factor is used to create scale pyramid as shown in the picture. Suppose, the scale factor is 1.03, it means we're using a small step for resizing, e.g. reduce size by 3%, it increase the chance of a matching size with the model for detection is found, while it's expensive.



[img6 – Scale factor effect on images]

In this system the scale factor step is 0.1, the starting scale factor is 1.1 and at every try its increased by 0.1 until it find the best scale factor for which the system identifies the highest number of faces in the images and the most match of proves with the right gallery images.

In particular this function checks first of all of the number of matches is increased or decreased then it checks the number if identified faces in the mages and at last the number of non matched people.

If their values are positive at each step or equal to previews step the scale factor keep increasing in search for the best value otherwise it stops and the the value for the rest tests.

This function creates a new probe set at every step.

```
Creating genuine probes ...
./180x200/elduns/elduns.jpg => ./180x200probes/elduns.jpg
./180x200/dioann/dioann.16.jpg => ./180x200probes/dioann.16.jpg
./180x200/cshubb/cshubb.13.jpg => ./180x200probes/cshubb.13.jpg
./180x200/jlemon/jlemon.19.jpg => ./180x200probes/jlemon.19.jpg

Training | 100.0% Complete

Ground truth: True
Matched images:
dioann.16.jpg ./180x200/dioann/dioann.15.jpg 2 38.17045698235684
elduns.jpg ./180x200/elduns/elduns.1.jpg 110 0.0
jlemon.19.jpg ./180x200/jlemon/jlemon.20.jpg 264 28.227850005677034
cshubb.13.jpg ./180x200/cshubb/cshubb.12.jpg 49 40.78917792191868

Training | 100.0% Complete

The following images may have no face:
./180x200/ksunth/ksunth.13.jpg
./180x200/ksunth/ksunth.8.jpg
./180x200/ksunth/ksunth.7.jpg

Ground truth: True
Matched images:
dioann.16.jpg ./180x200/dioann/dioann.15.jpg 2 37.71998976696026
elduns.jpg ./180x200/elduns/elduns.1.jpg 110 0.0
jlemon.19.jpg ./180x200/jlemon/jlemon.20.jpg 264 28.687553456937923
cshubb.13.jpg ./180x200/cshubb/cshubb.15.jpg 39 44.38283338074379

Best configuration for Scale Factor: 1.1
Not recognizer dataset image: 0/397
Match: 4/4
Error: 0/4
```

[img7 – Setting the scale factor by brute-forcing on the second data-set]

As the picture shows tested with the first data-set with probe number set to 4 and scale factor 1.1 all the faces in the pictures were identified and all the probes matched with their respective images in the gallery.

The check is done with the ground truth enabled so we know for sure if they match or not regardless of the set threshold.

At the second try with scale factor 1.2 in three images faces were not detected due to a high scale factor for the size of the images in the gallery (180x200) so the system take the first scale factor (1.1) and set it for the system.

```

Training | 100.0% Complete

The following images may have no face:
./480x600/S453/S453.11.t10_03.jpg
./480x600/S453/S453.11.t10_02.jpg
./480x600/S001/S001.08.t10_02.jpg
./480x600/S001/S001.08.t10_03.jpg
./480x600/S483/S483.02.t10_02.jpg
./480x600/S483/S483.04.t10_02.jpg
./480x600/S483/S483.02.t10_03.jpg

Ground truth: True
Matched images:
S171.08.t10_01.jpg ./480x600/S171/S171.04.t10_01.jpg 88 17.50325071602162
S134.04.t10_01.jpg ./480x600/S134/S134.08.t10_01.jpg 38 30.82157358464331
S243.04.t10_01.jpg ./480x600/S243/S243.06.t10_01.jpg 76 19.292069832714265
S483.05.t10_01.jpg ./480x600/S483/S483.02.t10_01.jpg 86 35.18023994341248

Training | 100.0% Complete

The following images may have no face:
./480x600/S453/S453.11.t10_03.jpg
./480x600/S453/S453.11.t10_02.jpg
./480x600/S001/S001.08.t10_02.jpg
./480x600/S001/S001.08.t10_03.jpg
./480x600/S483/S483.02.t10_02.jpg
./480x600/S483/S483.04.t10_02.jpg
./480x600/S483/S483.02.t10_03.jpg
./480x600/S044/S044.02.t10_01.jpg

Ground truth: True
Matched images:
S171.08.t10_01.jpg ./480x600/S171/S171.04.t10_01.jpg 88 17.205285097086545
S243.04.t10_01.jpg ./480x600/S243/S243.06.t10_01.jpg 76 19.334245309615262
Non matched images:
S134.04.t10_01.jpg ./480x600/S001/S001.04.t10_01.jpg 65 33.61288442162635
S483.05.t10_01.jpg ./480x600/S134/S134.08.t10_01.jpg 38 37.701517513605374

```

[img7a – Setting the scale factor by brute-forcing on the second data-set]

As we see in this case with the second data-set at the first step with scale factor 1.1 there are 7 faces are not recognized from the gallery and all the probes are correctly matched.

But in the second step with scale factor 1.2, 8 faces are not recognized so worse than before but the matching become worse, in fact 2 probes match and 2 don't.

Now the way it behaves and the criteria it uses to calculate the scale factor can be changed depending on the need of the user.

SetThreshold

This function searches for the best threshold for the given data-set.

This function calculates FRR and FAR to set the right threshold with open set and make sure that the number of impostors is equal to the number of genuine probes..

The system put its max threshold at 140 and the min threshold at 0, the actual threshold is put at the average of these value which is 70, than the system populate the open set probes and test with the ground truth enable how many how false acceptance and how many false rejection are there and depending on their value the threshold will be increase or decreased until FRR and FAR are not equal.

This process should be repeated until FRR and FAR are not equal but its being done up to 10 times due to few particular cases that goes misses during this face where few people match with themselves with a higher number of distance.

While every time a new probe set (closed set) is generated more steps we do in this phase and more accurate the threshold will be instead doing the standard test for the threshold might cause few future cases of false acceptance or false rejection.

```

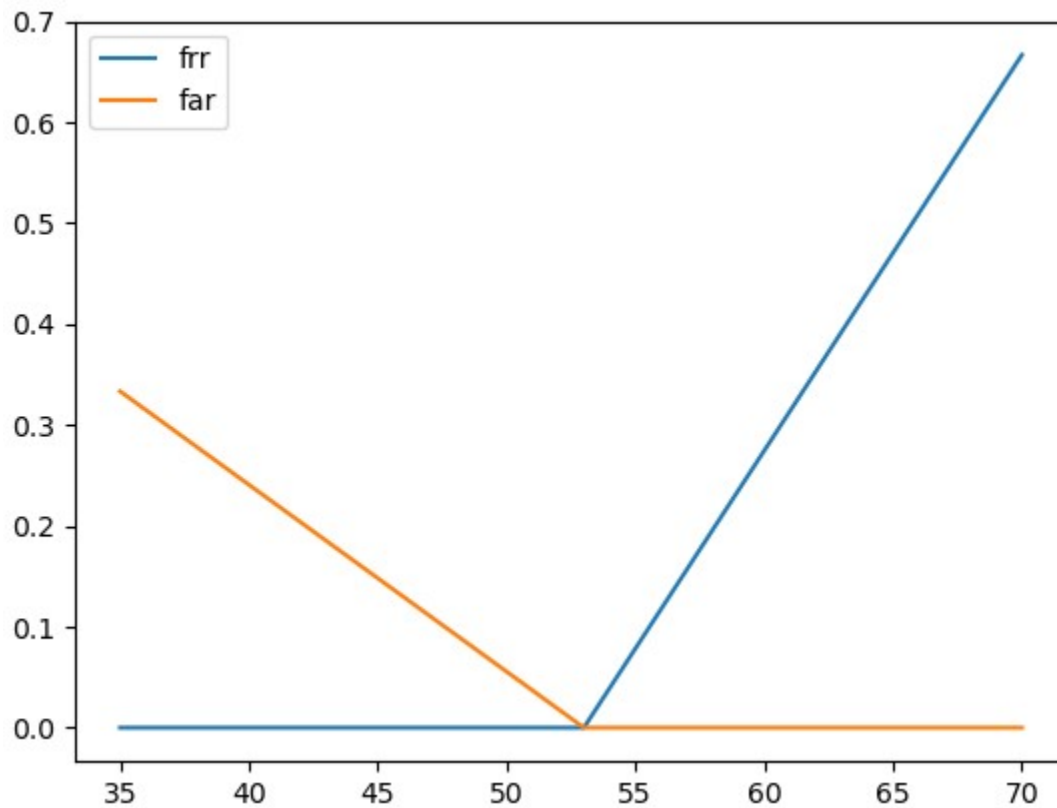
Creating genuine probes ...
./180x200/ekavaz/ekavaz.5.jpg => ./180x200probes/ekavaz.5.jpg
./180x200/ksunth/ksunth.5.jpg => ./180x200probes/ksunth.5.jpg
./180x200/astefa/astefa.8.jpg => ./180x200probes/astefa.8.jpg
./180x200/dcbowe/dcbowe.1.jpg => ./180x200probes/dcbowe.1.jpg
Creating impostors
./180x200impostors/IMP_cferdo/IMP_cferdo.9.jpg => ./180x200probes/IMP_cferdo.9.jpg
./180x200impostors/IMP_cjcarr/IMP_cjcarr.6.jpg => ./180x200probes/IMP_cjcarr.6.jpg
./180x200impostors/IMP_anpage/IMP_anpage.8.jpg => ./180x200probes/IMP_anpage.8.jpg
./180x200impostors/IMP_yfhsie/IMP_yfhsie.2.jpg => ./180x200probes/IMP_yfhsie.2.jpg
Fase 1/10 | 100.0% Complete

Ground truth: True
Matched images:
ekavaz.5.jpg ./180x200/ekavaz/ekavaz.7.jpg 285 39.936057242859015
ksunth.5.jpg ./180x200/ksunth/ksunth.6.jpg 94 52.305275731815804
dcbowe.1.jpg ./180x200/dcbowe/dcbowe.6.jpg 207 35.00617955084593
astefa.8.jpg ./180x200/astefa/astefa.4.jpg 124 28.01286837583644
Non matched images:
IMP_cferdo.9.jpg ./180x200/dagran/dagran.19.jpg 168 69.22234166734401
IMP_yfhsie.2.jpg ./180x200/asamma/asamma.20.jpg 373 71.63701292343607
IMP_cjcarr.6.jpg ./180x200/jlemon/jlemon.1.jpg 258 57.63999549845534
IMP_anpage.8.jpg ./180x200/klclar/klclar.1.jpg 334 68.64906632178862
Threshold: 70 FRR: 0.0 FAR: 0.75 Equal?: False
Resetting probes images ...
./180x200probes/ekavaz.5.jpg => ./180x200/ekavaz/ekavaz.5.jpg
./180x200probes/IMP_cferdo.9.jpg => ./180x200impostors/IMP_cferdo/IMP_cferdo.9.jpg
./180x200probes/ksunth.5.jpg => ./180x200/ksunth/ksunth.5.jpg
./180x200probes/dcbowe.1.jpg => ./180x200/dcbowe/dcbowe.1.jpg
./180x200probes/astefa.8.jpg => ./180x200/astefa/astefa.8.jpg
./180x200probes/IMP_yfhsie.2.jpg => ./180x200impostors/IMP_yfhsie/IMP_yfhsie.2.jpg
./180x200probes/IMP_cjcarr.6.jpg => ./180x200impostors/IMP_cjcarr/IMP_cjcarr.6.jpg
./180x200probes/IMP_anpage.8.jpg => ./180x200impostors/IMP_anpage/IMP_anpage.8.jpg
Creating genuine probes ...
./180x200/klclar/klclar.15.jpg => ./180x200probes/klclar.15.jpg
./180x200/astefa/astefa.16.jpg => ./180x200probes/astefa.16.jpg
./180x200/jlemon/jlemon.8.jpg => ./180x200probes/jlemon.8.jpg
./180x200/asamma/asamma.14.jpg => ./180x200probes/asamma.14.jpg
Creating impostors
./180x200impostors/IMP_yfhsie/IMP_yfhsie.9.jpg => ./180x200probes/IMP_yfhsie.9.jpg
./180x200impostors/IMP_9336923/IMP_9336923.14.jpg => ./180x200probes/IMP_9336923.14.jpg
./180x200impostors/IMP_cgboyc/IMP_cgboyc.19.jpg => ./180x200probes/IMP_cgboyc.19.jpg
./180x200impostors/IMP_cjdenn/IMP_cjdenn.14.jpg => ./180x200probes/IMP_cjdenn.14.jpg
Fase 2/10 | 100.0% Complete

Ground truth: True
Matched images:
asamma.14.jpg ./180x200/asamma/asamma.13.jpg 372 34.01118613463986
klclar.15.jpg ./180x200/klclar/klclar.2.jpg 326 27.958402219274213
astefa.16.jpg ./180x200/astefa/astefa.17.jpg 131 31.31784491746108
jlemon.8.jpg ./180x200/jlemon/jlemon.3.jpg 260 31.516007077061225
Non matched images:
IMP_cgboyc.19.jpg ./180x200/jlemon/jlemon.17.jpg 262 76.30649438593474
IMP_9336923.14.jpg ./180x200/kaknig/kaknig.8.jpg 241 76.44616143902495
IMP_cjdenn.14.jpg ./180x200/jlemon/jlemon.5.jpg 273 72.127797474643
IMP_yfhsie.9.jpg ./180x200/asamma/asamma.20.jpg 373 73.9891297464532
Threshold: 35 FRR: 0.0 FAR: 0.0 Equal?: True

```

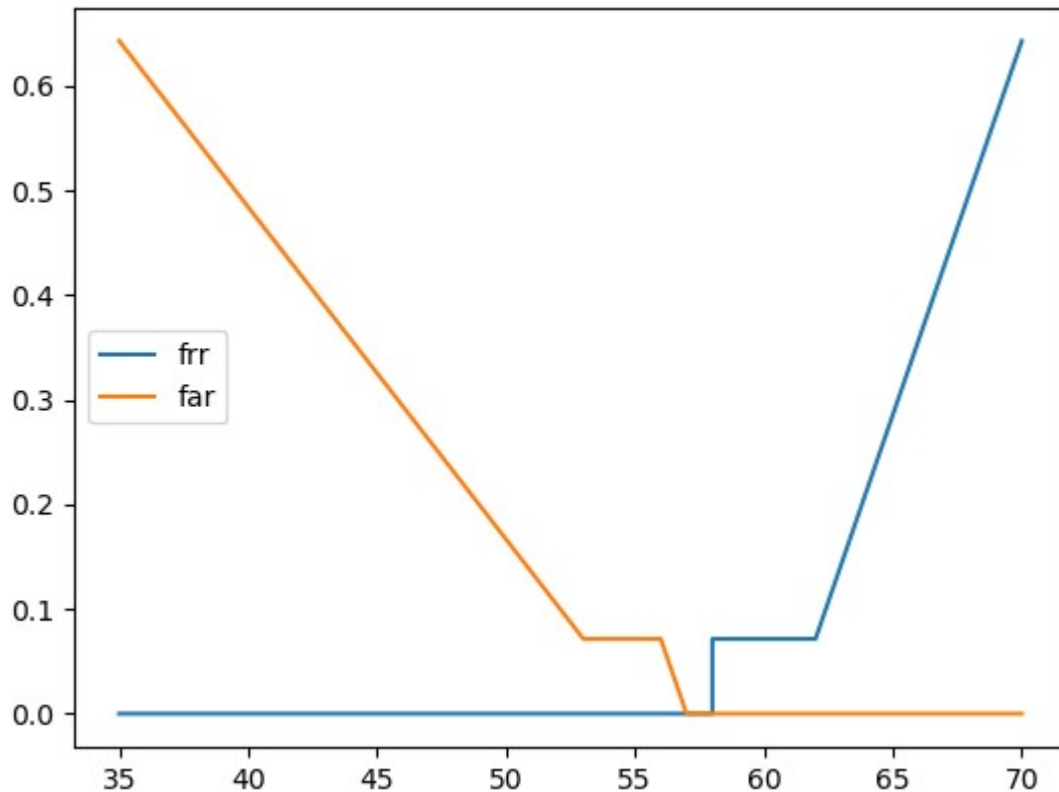
[img8 – First 2 steps in search of best threshold for the given dataset]



At the end of this process the function draw the graph of the FRR and FAR.*[img9 – FRR and FAR curve, threshold = 54]*

Notice how the curve is straight even if the intersection of FRR and FAR is found.

This is caused because of the low number of point on the graph, exactly tree point for curves has been drawn because the function found the right threshold after three steps.



[img10 – FRR and FAR curve, threshold = 58]

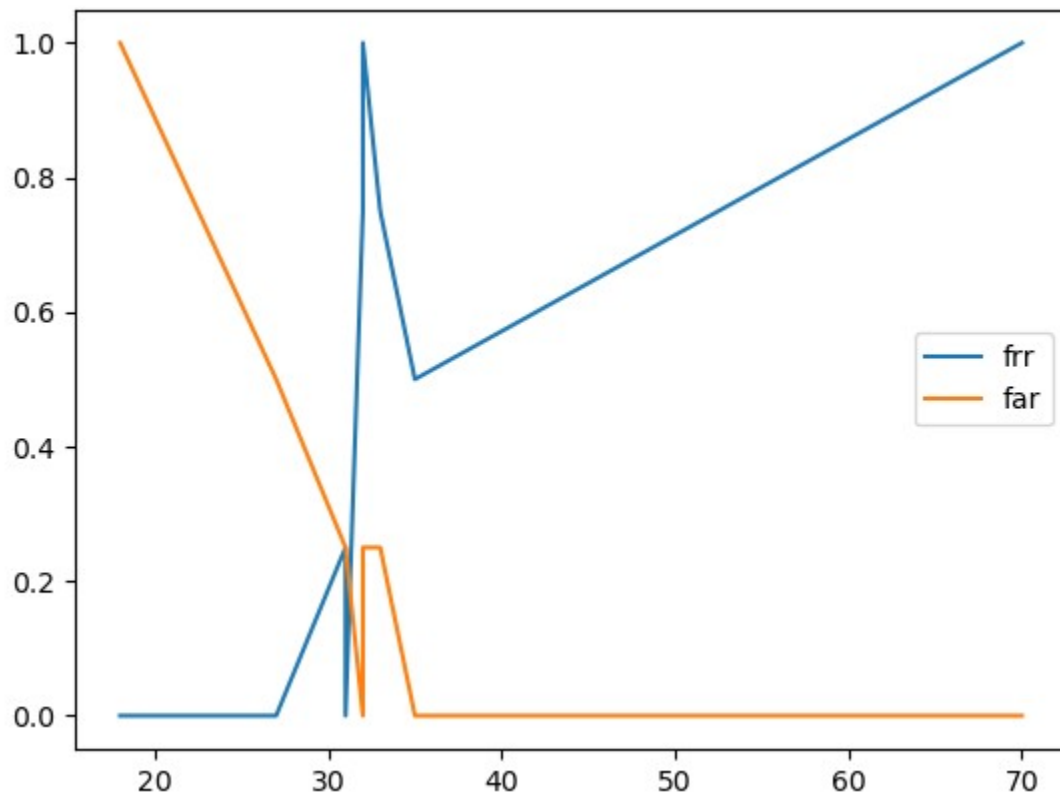
Notice how this curve is different from the previous one, it has one more intermediate and in this case the threshold is more accurate even if it's not perfect but it is more than acceptable.

Depending on the random probes the system generates the curve might vary and the threshold within it, sometime the system is more strict sometimes less.

With the following data-set the optimal threshold is 62 because there are few particular people's images that match with themselves with distance around 58 to 61 but the found threshold of 58 is acceptable because high distance might be caused due to a malformed picture and the capture can be always retaken.

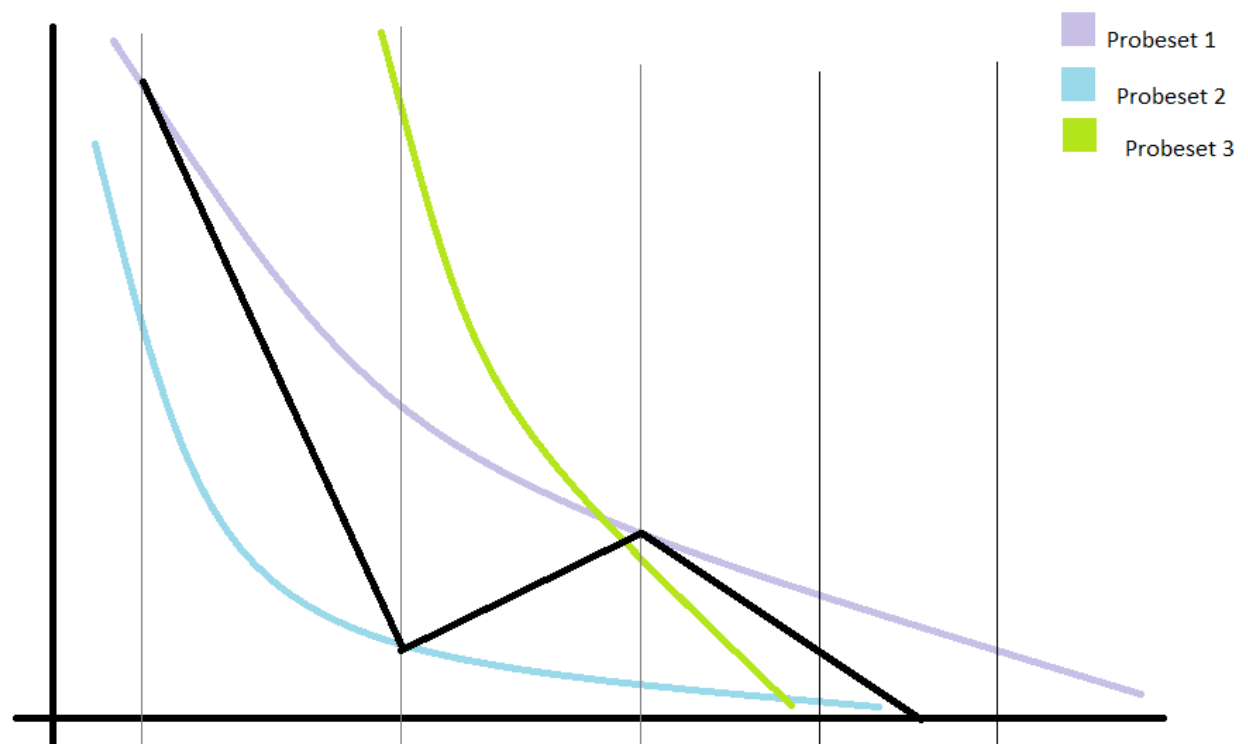
It's better to have a little bit strict system than having false acceptance.

With the second data-set we have a more messy graph due to a more heterogeneous probe set at every try.



[img10a – FRR and FAR curve, threshold 32]

This happens because every probe set that is made has a different curve at each point, The following example explains the behavior.



CMS (Cumulative Match Score)

This function calculated the probability and the ratio of CMS to a given k or a range of k (from 1 to k).

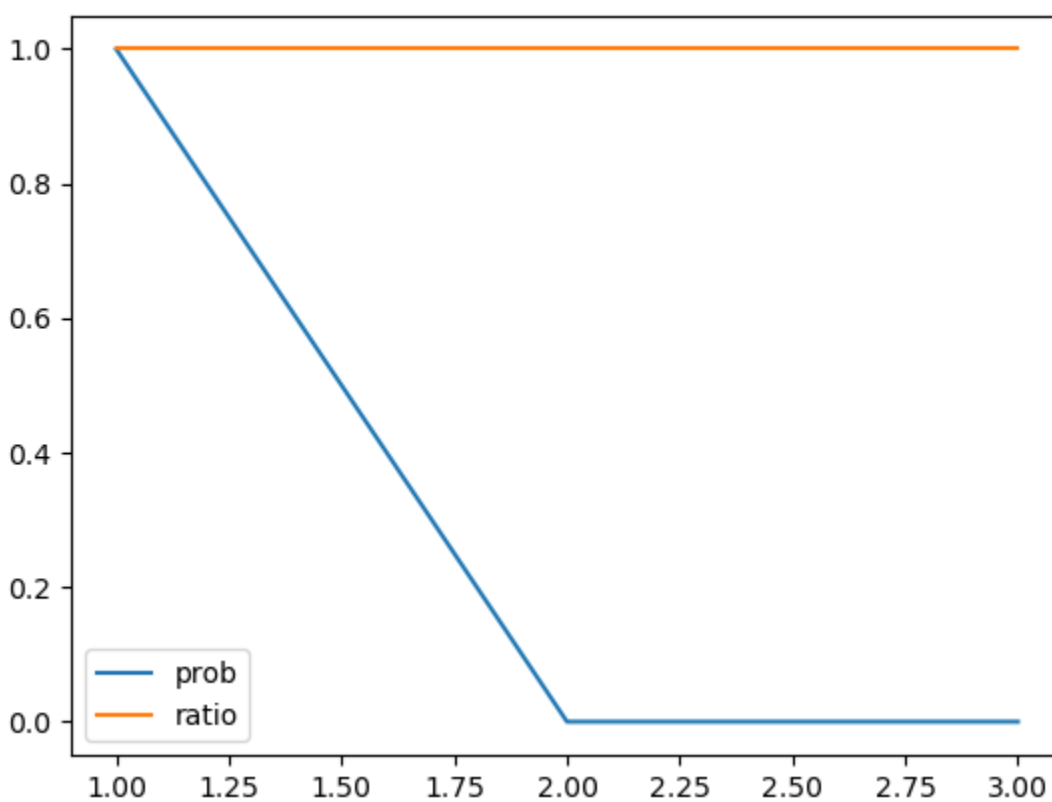
First of all we calculate the matrix (Probes X Gallery) and take the minimum distance of each person.

	A1	A2	A3	B1	B2	B3	C1	C2	C3		A1	B3	C3
Probe A	23.5	22.6	33.3	66.1	65.0	63.3	75.5	71.2	71.1		22.6	63.3	71.1

[img11 – One minimum distance per person is taken]

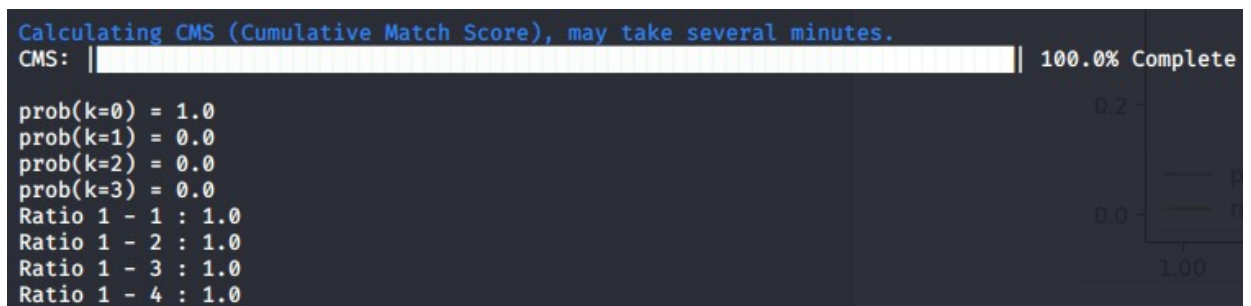
Once got the matrix we check the CMS from rank 1 to rank k and calculate it as #match at tank k / #probes at each rank.

The ratio is calculate for rank 1 to 1, 1 to 2... 1 to k and produce the graph with the first data-set.



[img12 – CMS ratio and probability with k = 4]

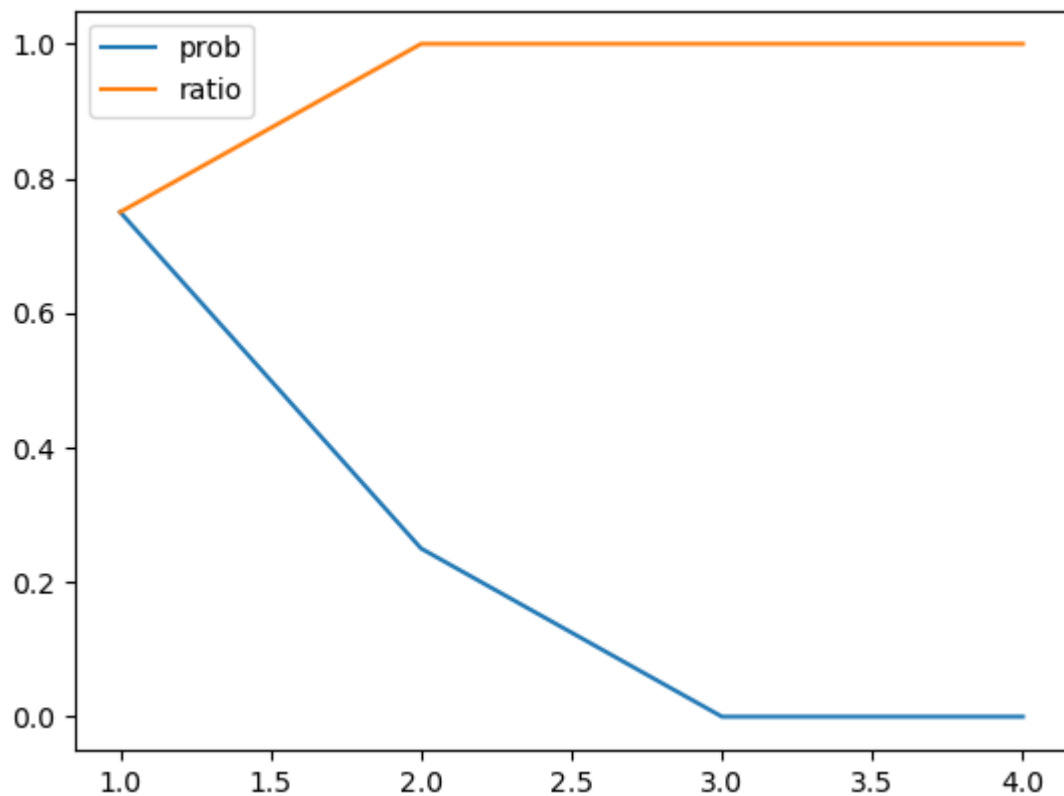
For the probability there is a 100% hit rate of the probes at rank 1 and 0% at rank 2 and 3, for the the ratio there is a 100% hit at every rank while it is the ratio between the number of individuals which are correctly recognized among the first k and the total number of individuals in the test set (probe).



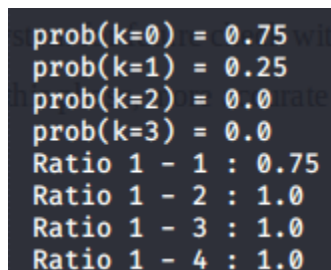
[img13 – Log of probability and ratio at each rank]

Here the img13 show the Log of the ratio and probability value at each rank.

Same test with the second data-set show a different graph.



[img12a – CMS ratio and probability with $k = 4$]



[img13a – Log of probability and ration at each rank]

Train

This function enroll the gallery images into the system for future check with probes.

The previews calculated scale factor is crucial in this phase, more accurate is the scale factor more faces will be detected and enrolled.

```
Training | 100.0% Complete
Ground truth: False
Matched images:
csanch.13.jpg ./180x200/csanch/csanch.6.jpg 224 40.481351384781
dagran.7.jpg ./180x200/dagran/dagran.8.jpg 173 36.41705161327694
kaknig.11.jpg ./180x200/kaknig/kaknig.5.jpg 249 32.84837448375788
klclar.13.jpg ./180x200/klclar/klclar.15.jpg 324 29.849577758748907
```

[img14 – Successful training]

```
Training | 100.0% Complete
The following images may have no face:
./180x200/ksunth/ksunth.13.jpg
./180x200/ksunth/ksunth.8.jpg
./180x200/ksunth/ksunth.7.jpg
```

[img15 – Training with non optimal scale factor]

TopMatch

Function that return the best match (lowest distance) between probes and an image of the gallery.

General Overview

The tool can be launched with optional parameters:

- l : enable logs from the beginning (by default its disabled)
- i : interactive mode useful for presentation (stops at each step)

The system set all the parameters at the initial phase such as scale factor, threshold and CMS (Cumulative Match Score) than let the user to make further test through a menu giving various types of options.

```
1 - Create Closed Set Gallery
2 - Create Open Set Gallery
3 - Train
4 - Predict
5 - Reset Probes
6 - Complete Test (Closed Test) (5+1+3+4+5)
7 - Complete Test (Open Set) (5+2+3+4+5)
8 - Enable/Disable Logs (True)
9 - Print configurations
0 - Exit
Select option: 
```

[img16 – Possible operations]

Once the initial phase set up all the values the used can make further test creating open set or closed set, creating as many different probe set as want and carry on further test to very the actual configuration with the help of the logs.

This tool can be the core of every biometrics system.

For example in case of development of a system that unlock a PC by taking the picture form the webcam, the picture can be sent to the tool already trained and then compare the result.

Or biometrics system for accessing room, also here the picture taken from the camera can be sent for identification or verification to the tool for the check.

The tool at the moment does only the identification but can be upgraded to do also the verification.