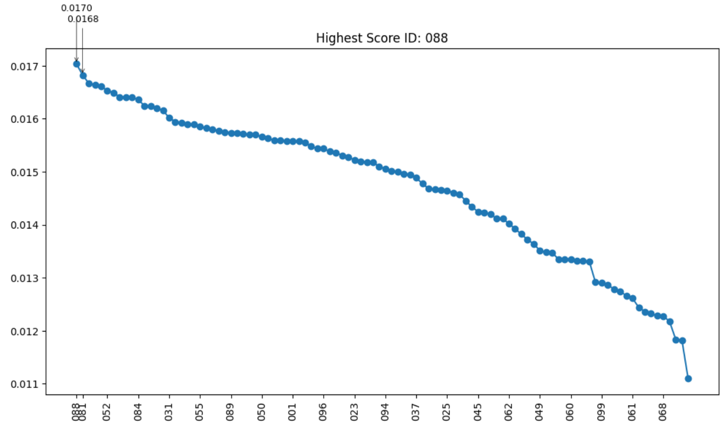
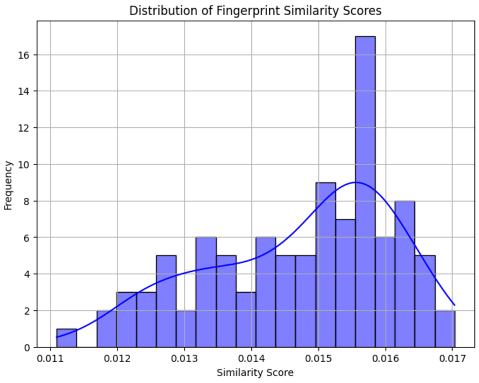
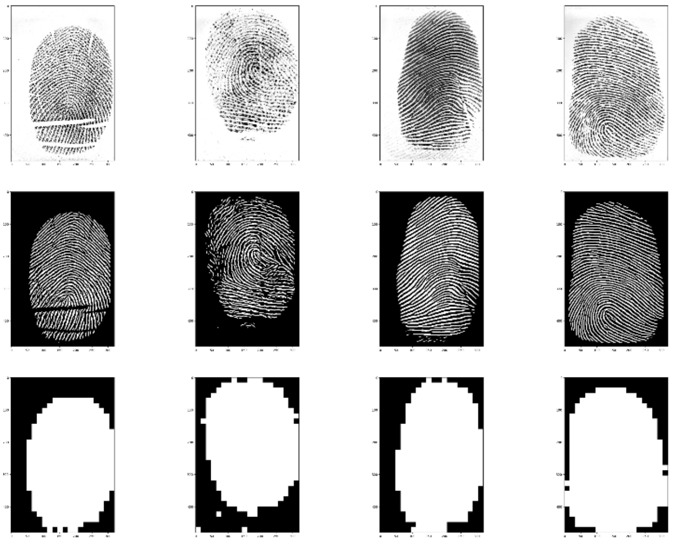
Q1:





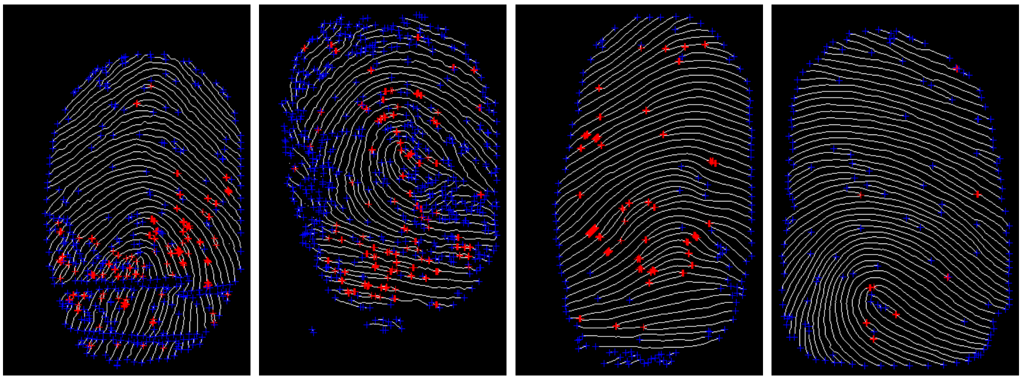
The given similarity function of comparing each fingerprint in the database with the one from the perpetrator based on the function of $\frac{1}{\text{1+average squared pixel to pixel difference}}$ is not a good measure to quantify the distance between two fingerprints. There is no strong drop in the left graph, and the distribution shows a broad spread without a clear cluster of high scores at the right tail, as can be seen in the histogram on the right. In addition, the highest values are very close to each other and the two highest values corresponding to label ID 088 and label ID 081 are 0.0170 and 0.0168 respectively, i.e. with an absolute change of only 0.0002 and a relative change of only about 1%, therefore it is not reliable enough to incriminate the suspect. A limitation is that even if the fingerprints are from the same person, small displacements, rotations or differences in pressure when placing the finger can cause significant changes in the pixel to pixel comparison, leading to low similarity scores despite the fingerprints being from the same finger. Even it can differ whether it is done by more wet or dry finger. Furthermore, any accidental marks on the fingerprint images will also affect the results, as the function is very sensitive to noise.

Q2:



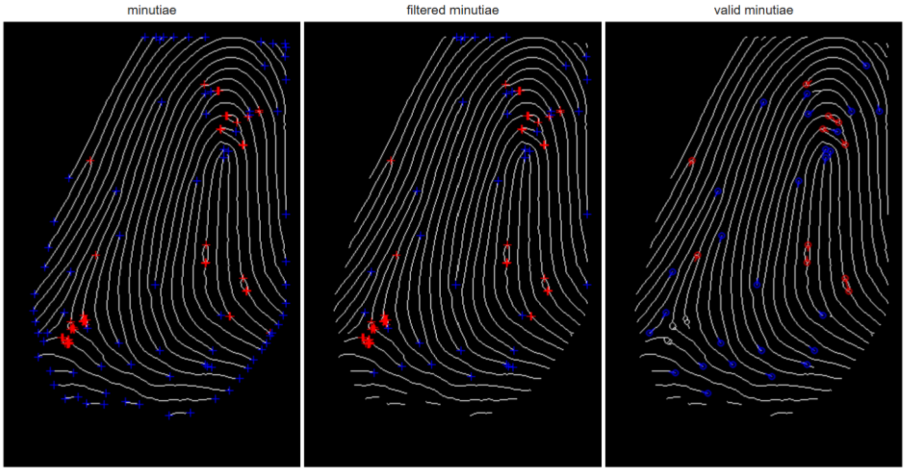
As can be seen from the graphs of the images and their enhanced versions, there are several obstacles when working with fingerprints. One of them is the presence of accidental marks on the images, as illustrated by the first image in the first column - the horizontal lines, likely caused by acquisition issues, are preserved even after enhancement. Even if these marks originate from permanent features on the finger itself, they remain problematic for fingerprint processing, as they can confuse feature extraction algorithms. Another challenge is the overall quality of the fingerprint, which can be influenced by factors such as low pressure, dry skin, dirt, or poor capture - the second image in the second column is of lower quality, and even after enhancement, the image remains more degraded than the others, with the enhancement process only partially enhancing the whiter areas. Furthermore, enhancement can exaggerate imperfections and result in oversharpening, as can be seen in the third image of the third column. In this case, there appears to be scanning-related noise - a greyish, differently oriented partial print near the bottom - which is partially captured but also partially ignored by the enhancement process. This can mislead feature extraction algorithms by introducing noise into the region of interest. In addition, some ridges may appear broken or faded due to acquisition problems, and although the enhancement algorithm addresses this to some extent, the limitation remains visible in the enhanced images. Such discontinuities can cause difficulties for algorithms trying to extract features such as minutiae points.

Q3:



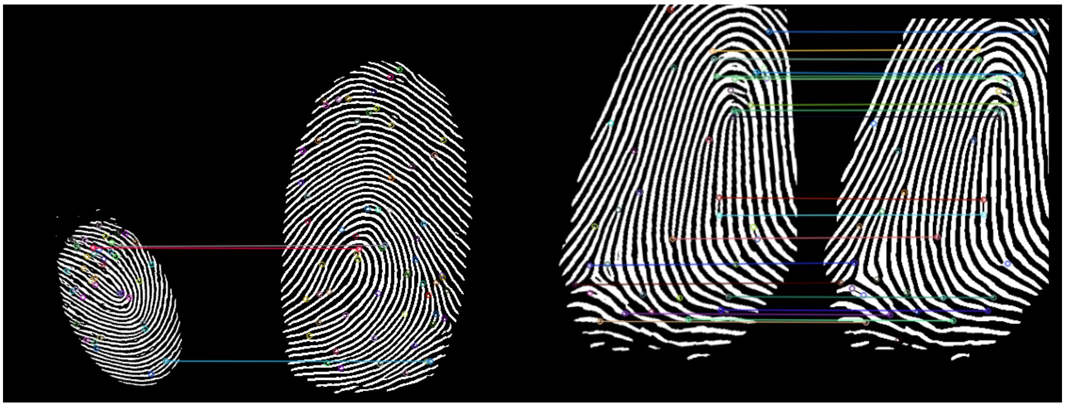
The detection algorithm works by scanning the skeletonised fingerprint image pixel by pixel, except the border pixels. For each white pixel, it extracts the pixel itself and its 8 surrounding neighbours and then counts the number of white pixels. Two white pixels means the pixel has only one white neighbour, hence it is a termination. Four white pixels means the pixel has three white neighbours, hence it is a bifurcation. The results are then stored in a list of tuples in the form of (x-coordinate, y-coordinate, True if termination, False if bifurcation). As can be observed, there are a lot of false detections of termination on the borders of the fingerprint.

Q4:



The detection of the orientations of minutiae on top of just locations helps to remove false detections of both terminations and bifurcations. As can be seen in the middle skeleton image, the application of an eroded mask removes minutiae that are too close to the fingerprint boundary, mainly eliminating false terminations. Incorporating minutiae orientations further improves the removal of false detections, not only for terminations but also for bifurcations, as seen in the right skeleton image. Orientation detection improves noise handling and results in more robust, reliable minutiae that serve as stronger features for fingerprint recognition systems.

Q5:

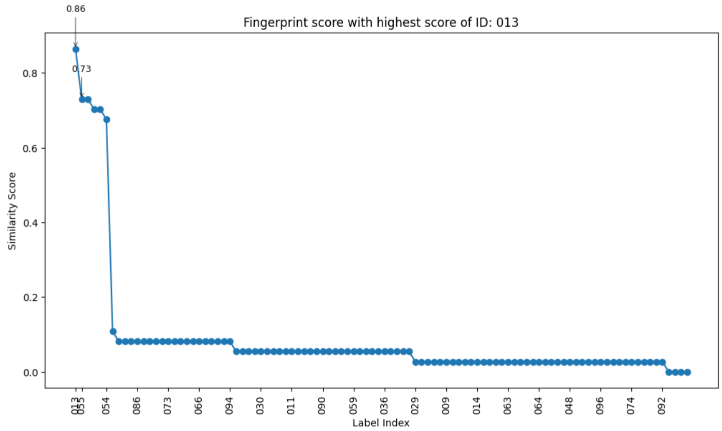


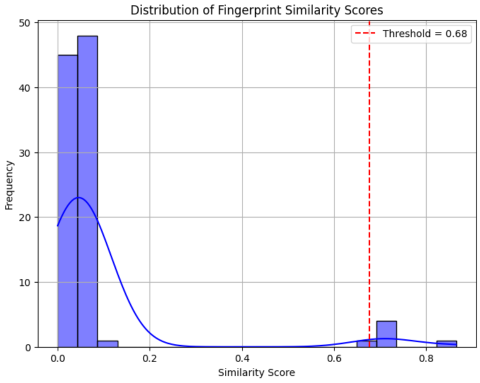
On the left comparison, there are three matches that are rather incorrect, being mapped to wrong locations of keypoints on the right fingerprint. On the right comparison, there are more matches that are mostly correct, but not all possible correct matches are taken into account. This means that not all keypoints match accurately. This result is to be expected due to the presence of noise and local ambiguities, as some regions of a fingerprint may appear locally similar even though they are from different fingers. However, as the algorithm is a hybrid strategy that performs a local structure matching followed by a consolidation stage of an affine transformation aimed at verifying if and to what extent local matches hold at the global level, the distinctiveness of the matches remains high and the influence of local ambiguities is reduced as only matches that fit into a consistent global structure are kept. Although occasional mismatches may still occur, in general we expect to find a large number of matches for fingerprints from the same person and only a very limited number of matches between fingerprints from different people, as local matches that do not fit the global transformation are discarded.

Q6:

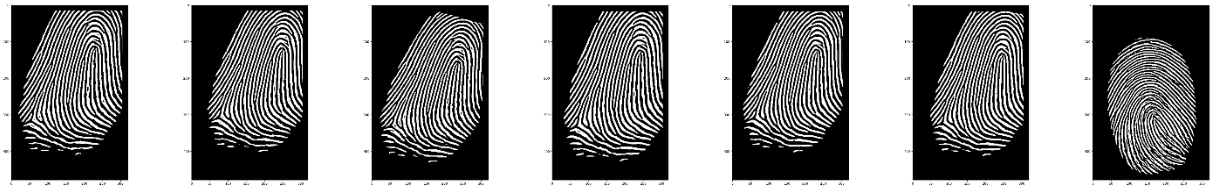
A global feature similarity function is constructed using the matching keypoints from Q5 by counting the number of keypoint pairs whose Euclidean distance is smaller than a threshold of 3 pixels, divided by the total number of minutiae extracted from the fingerprint being compared - in our case, the fingerprint of the perpetrator. The resulting similarity score lies between 0 and 1, where a score of 1 indicates very high similarity and a score close to 0 indicates no meaningful similarity.

Q7:





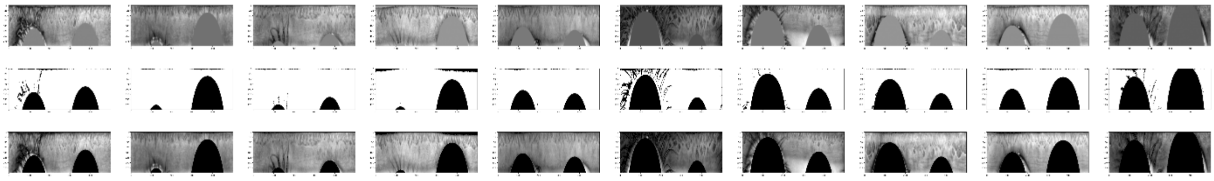
As observed, there are six top candidates for the perpetrator that have significantly higher scores compared to the rest of the fingerprints. However, the scores among these top candidates are very close to each other, indicating uncertainty about which specific match is truly connected to the perpetrator. The proposed score threshold to discriminate matching fingerprints is set at the 95th percentile of the score distribution, which corresponds to a value of 0.68, resulting in five candidates. Another approach is to use the knee point of the score curve as a threshold, leading to a threshold value of 0.094 and identifying seven candidates. Thus, no clear distinctions is achieved. This is because the six best matching fingerprints all correspond to the same fingerprint, which is indeed the fingerprint of the perpetrator:



Q8:

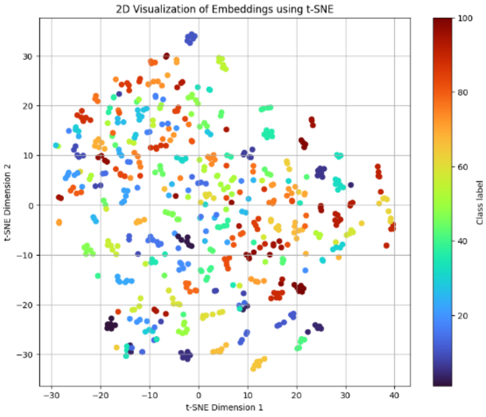
There are several difficulties associated with iris acquistion that are evident in the exercise, highlighting the need for effective preprocessing to ensure robust feature extraction. The main challenges include partial occlusion of the iris by eyelashes and eyelids. Although the iris images we work with are frontal, they can still have slight variations in angle and suffer from motion blur. In addition, because the iris is behind a curved, wet and reflecting surface, it can be obscured by specular reflections, and the illumination can vary across the surface, sometimes being too bright. In addition, the iris deforms non-elastically as pupil changes size. As a result, the similarity measures that are expected to work best are those that focus on analysing texture patterns within the iris region itself, particularly those that are able to capture the fine radial structures that are characteristic of the iris.

Q9:



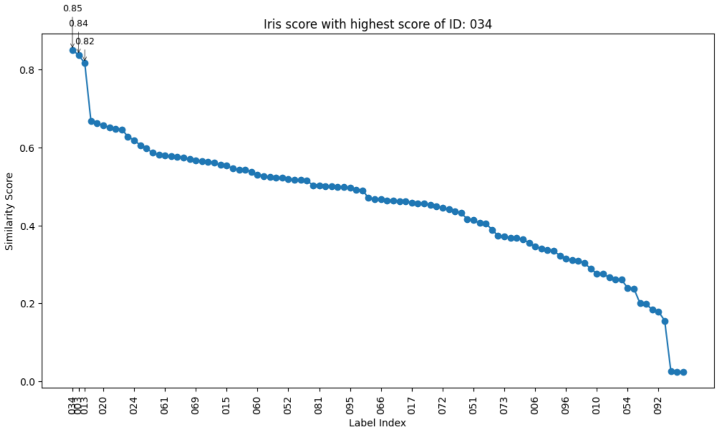
As can be seen, the segmentation process helps to remove occlusions such as eyelashes and eyelids to some extent, but some masks are either too harsh, cutting away useful iris texture, or too loose, leaving occlusions such as eyelashes visible in the enhanced images. Segmentation can also misrepresent the actual iris texture, either by removing too much or by including unnecessary non-iris regions. The enhancement process improves the visibility of iris textures, but it can also enhance noise and specular reflections caused by the wet surface of the eye, particularly around the pupil region, which are still visible although partially masked. Enhancement also helps to reduce deformation effects by providing a normalised rectangular iris segment, but variations in pupil size still affect the texture to some extent. In addition, the transformation into a rectangular shape introduces edge distortions from the unrolling process, particularly near the top of the images, where intensity transitions and stretching distortions occur.

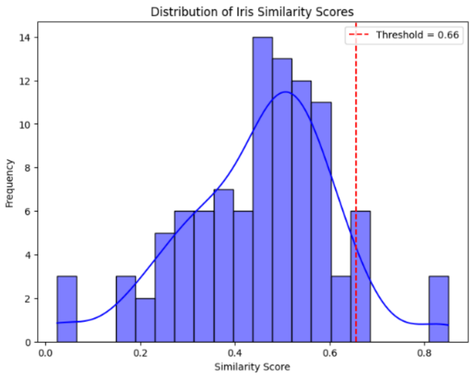
Q10:



A dimensionality reduction method used is T-distributed Stochastic Neighbor Embedding (t-SNE). It converts similarities between data points to joint probabilities and tries to minimize the the relative entropy (Kullback-Leibler divergence) between the joint probabilities of the low-dimensional embedding and the high-dimensional data. In the resulting 2-dimensional space, embeddings of training images with the same class label are located very close to each other, indicating that the embeddings preserve similarity for the same identity and successfully capture important image features, as embeddings of different images of the same label are clustered together. However, there is still some overlap between embeddings of different labels, meaning that the separation between different iris patterns is not entirely perfect, and some iris images are considered similar by the model.

Q11:





A global feature similarity function is constructed by calculating the Euclidean distances between the perpetrator's iris embedding and the embeddings of the training database iris images, and then averaging these distances for each class label. Since a lower distance indicates a higher similarity, a transformation is applied to convert distances into similarity scores, defined as:

$$\text{similarity score}=e^{\frac{-\text{average distance}^2}{2*\tau^2}}, \text{ where } \tau=2.$$

Thus, the lower the distance, the higher the similarity score, with the score ranging between 0 and 1. Alternatively, simpler transformations, similar to the one used in Q1, could have been applied, but here the decay of the score is more precisely controlled by using the parameter τ.

Based on the results, it can be observed that the top three scores, corresponding to labels 034, 003, and 013, are very close to each other and significantly higher than the scores of other labels. Following a similar approach to Q7, a threshold is set at the 95th percentile of the score distribution, corresponding to a score of 0.66, thus including the top 5 labels. This indicates that iris scans do not provide a clear suspect either.

Q12:

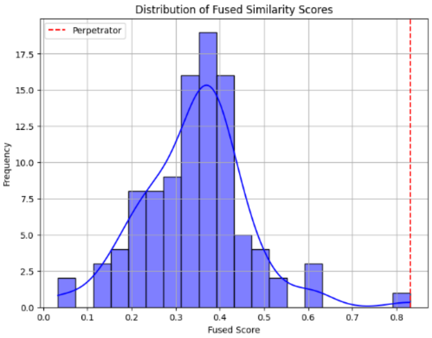
The scores are fused at the score level by summing the fingerprint and iris scores for each label, weighting the iris score slightly higher since the fingerprint data was duplicated and thus considered slightly less reliable.

multimodal (fused) score=0.7*iris score+0.3*fingerprint score.

The top 5 highest fused scores are summarized in the following table:

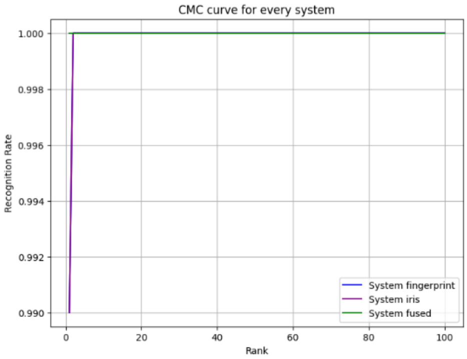
Label id	Fingerprint score	Iris score	Fused score
013	0.864865	0.816642	0.831109
055	0.729730	0.565106	0.614493
034	0.054054	0.850546	0.611598
003	0.054054	0.837081	0.602173
019	0.702703	0.462057	0.534250

The distribution of the fused similarity scores is shown below:

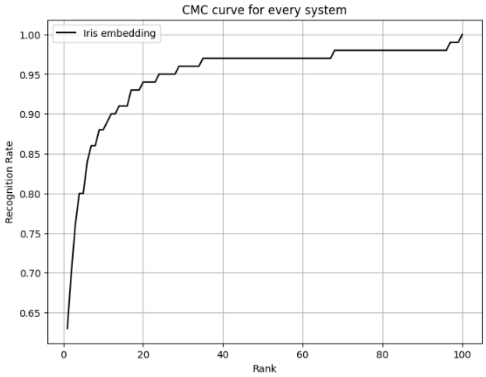


The predicted perpetrator is the label with the highest fused score, which is label 013 with a fused score of 0.831, resulting in a gap of 0.22 compared to the second best score. This significant gap indicates a relatively confident prediction. In addition, label 013 is the only one above the thresholds for both fingerprint and iris scores determined in Q7 and Q11 respectively, which further strengthens the confidence in the prediction. We therefore predict label 013 as the perpetrator.

Since the perpetrator's ground truth label is not available, in order to assess the system using the metrics implemented in the previous assignment, we instead construct a similarity score matrix of size 100 by 100 by comparing the labels in the database to each other. For fingerprints, due to the matching method used, we expect perfect similarity scores of 1 along the diagonal. However, for the iris embeddings, the scores depend on how well the network has learned to represent identity, so the diagonal scores are not necessarily perfect. To evaluate performance in an identification scenario, the CMC curve is used.



If we use the same training data of the iris and the given fingerprint as for the perpetrator identification, we get 99% and 99% Rank 1 recognition rates for fingerprint and iris respectively, and 100% for the fused scores. The reason we do not get 100% for the fingerprint is that the same label as the true label also gets a score of 1. To better evaluate the iris embeddings independently, we also evaluated the system on validation data.



Here, the Rank-1 recognition rate for iris embeddings drops to 63%, meaning that the correct identity is recognised within the top 1 score 63% of the time. However, recognition improves rapidly as the rank increases, with a recognition rate of almost 90% for the Rank-10. In addition, to further test the iris similarity score on the validation data, we also calculated the similarity score to the perpetrator to see if the predicted label of 013 received a high score, which it did, receiving a score of 0.91 as the second highest score.

Label id	Iris validation score
003	0.999334
013	0.909799
002	0.756276

For additional questions 7. And 8. are selected.

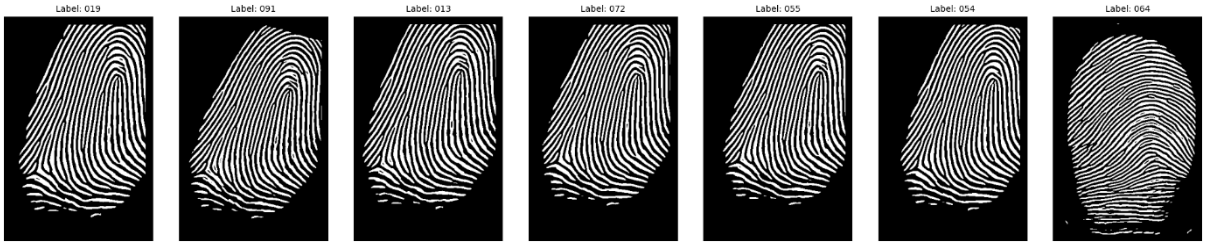
7.

To obtain similarity scores directly from image data without feature extraction, a Siamese neural network is used to learn similarity relationships between fingerprint image pairs. The model consists of a CNN encoder that extracts embeddings from each input fingerprint, followed by a fully connected layer that maps the extracted CNN features into a lower-dimensional embedding space. During training positive pairs are generated by creating augmented versions of the same fingerprint image using rotations, since only a single entry per identity is available for fingerprints:

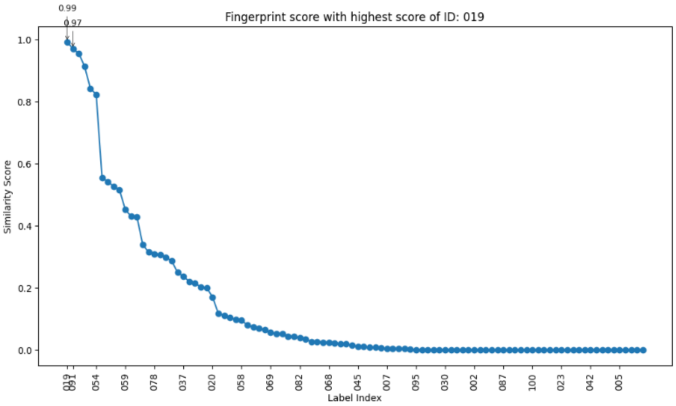
Negative pairs are constructed using fingerprint images from different labels, assuming that no duplicates are present - although, as shown in Q7, this assumption is not entirely accurate. Each image in a pair is passed independently through the CNN encoder. The pairwise distance between the two embeddings is computed using Euclidean distance. For positive pairs, the contrastive loss encourages small distances, while for negative pairs it encourages large distances. After training, the model is used to predict the similarity between the perpetrator's fingerprint and the fingerprints from the database by generating embedding distances. The distance is then transformed into a similarity score using a formula similar to the one used in Q11, with τ=1.5.

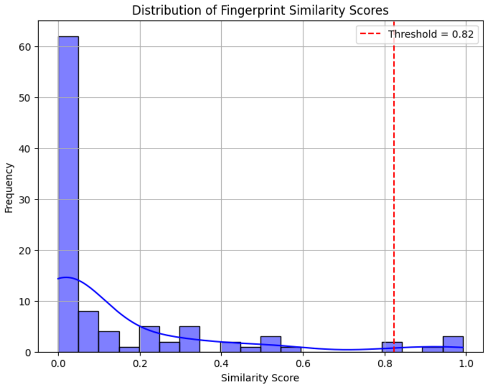
$$\text{similarity score}=e^{\frac{-\text{distance}^2}{2*\tau^2}}\text{, where }\tau=1.5$$

The result is that we get the same top six best matching fingerprints as in Q7 of the label id of 019, 091, 013, 072, 055 and 054, which all correspond to the same fingerprint, which is indeed the perpetrator's fingerprint.



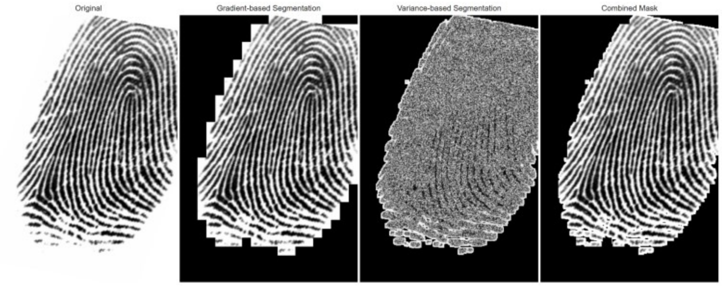
The threshold is again set as the 95th quantile of the distribution, which is equal to 0.82. We can see a clear separation between the top 6 scores and the rest. However, the gap between the 6th best score (last correct match) and the 7th label (first incorrect match) is slightly smaller compared to Q7, indicating a slight deterioration in the separation.





8.

The segmentation is improved here for fingerprint segmentation by combining the previous method based on the magnitude of the local gradient with the new method of local grey scale variance segmentation. The new method computes the local variance within a fixed 8 by 8 pixels window across the image, and pixels with variance above a set threshold of 300 are considered part of the fingerprint region, as high variance is expected to correspond to ridge patterns. A binary mask is then created by keeping the high variance areas. Finally, a combined mask is generated by retaining only those areas confirmed as fingerprint regions by both the gradient-based and variance-based segmentation methods.



To evaluate the combined mask, it is applied to the fingerprint database, and the fraction of the image considered as fingerprint region is calculated for each method:

Method	Fingerprint region
Magnitude of the local gradient mask	62.61%
Local grey scale variance mask	62.9%
Combined mask	59.4%

It can be observed that the introduction of the combined method reduces the considered fingerprint region, from 62.61% to 59.4% on average. The addition of the grey scale variance based method to the previous local gradient magnitude method helps to remove noise to some extent, as shown in the visualisation below, where the first row shows the segmentation based on gradient magnitude, the second row shows the segmentation based on grey scale variance, and the third row shows the result of the combined segmentation method.