Person Re-identification:

Context:

This involves the use of dimension reduction techniques such as PCA and/or LDA, or deep learning and Siamese networks, to learn a compact sub-space in which samples can be compared.

Person re-identification is commonly evaluated using Top-N accuracy and Cumulative Match Characteristic (CMC) curves. Top-N accuracy refers to the percentage of queries where the correct match is within the closest N results and is measured by ranking gallery samples based on their similarity to the probe and determining the location of the true match within the ranked list. Ideally, the top result (i.e., the closest gallery sample to the probe) will be the same subject. A CMC curve plots the top-N accuracy for all possible values of N (from 1 to the number of unique IDs in the dataset).

The data to be used is a portion of the Market-1501 dataset.

The data has been split into two segments:

- Training: consists of the first 300 identities from Market-1501. Each identity has several images. In total, there are 5,933 colour images, each of size 128x64.
- Testing: consists of a randomly selected pair of images from the final 301 identities. All images are colour, and of size 128x64. These images have been divided into two directories, Gallery and Probe, with one image from each ID in each directory.

Using the Training dataset, a model to extract a compact representation of a sample can be trained. The resultant transform can then be applied to the Testing set to transform samples to a lower dimensional representation, at which point samples can be matched. The testing set is broken into Gallery and Probe sets. To match samples, each sample in the Probe set can be compared to each image in the Gallery set, and based on the distance between pairs of probe and gallery samples the most similar instances can be identified.

Task Descriptions:

1. Use of PCA or LDA as a non-deep learning method for person re-identification. The method should be evaluated on the test set by considering Top-1, Top-5 and Top-10 performance. A CMC (cumulative match characteristic) curve should be generated.
2. Use of the Siamese Network (using pairs or triplet network) as a deep learning method for person re-identification. The method should be evaluated on the test set by considering Top-1, Top-5 and Top-10 performance. A CMC (cumulative match characteristic) curve should be generated.
3. Comparison of the performance of using PCA/LDA and the Siamese Network by generating CKNN classifiers.