# Pattern Recognition Coursework 2

Yao Lei Xu[1] (01062231) and Alejandro Gilson[2] (01112712)

*Abstract*— **Distance metrics are ubiquitous in machine learning. Given a metric that is appropriate for a dataset, the machine learning problem can be solved much more easily. For this report, we explore different distance metrics based machine learning methods in order to recognize a person's identity across different cameras. We explore some of the standard metrics such as Euclidean distances and correlation, as well as more advanced kernel and neural network based methods. The best result achieved was the cosine metric kMeans model, with 47.85% rank-1 accuracy.**

## I. THE MACHINE LEARNING PROBLEM

### A. Dataset Description

For this coursework, we are given the CUHK03 dataset [5] which contains 14096 labeled images of 1467 pedestrians taken by 2 different cameras, where each person is represented in the dataset approximately 7-8 times. In addition, we are given the features (2048 in total) extracted from each of the images via a ResNet50, as well as the indices of the feature vectors that must be used for training ($train\_idx$) and testing ($query\_idx$ and $gallery\_idx$) where the identity of each person is not repeated. The training dataset contains 767 pedestrians in total, while the testing set a different set of 700 pedestrians.

### B. Machine Learning Problem Formulation

The goal of this machine learning problem is to construct a distance metric based model that can correctly identify a pedestrian across different cameras given the features of that image. This means taking the feature vectors from the query set and create a ranklist for each of them by finding its nearest neighbors (with a given distance metric) within the gallery set (deleting the feature vectors that corresponds to same label and same camera). The performance metric used will be the standard rank-1 accuracy, which measures the percentage of correct classifications of a given method. The distance metrics used in this report will be pseudo-metrics with the following property:

1) Non-negativity: $D(\mathbf{P}, \mathbf{Q}) \geq 0$
2) Identity of indiscernibles: $D(\mathbf{P}, \mathbf{Q}) = 0$ if $\mathbf{P} = \mathbf{Q}$
3) Symmetry: $D(\mathbf{P}, \mathbf{Q}) = D(\mathbf{Q}, \mathbf{P})$
4) Subadditivity: $D(\mathbf{P}, \mathbf{Q}) \leq D(\mathbf{P}, \mathbf{K}) + D(\mathbf{K}, \mathbf{P})$

where $\mathbf{P}, \mathbf{K}, \mathbf{Q}$ are different data points and $D$ is the distance function.

## II. BASELINE APPROACH

For the baseline approach, we use a simple k-Nearest Neighbors (k-NN) algorithm using the Euclidean distance as metric (equation **??**) with k = 1.

$$D(\mathbf{P}, \mathbf{Q}) = ||\mathbf{P} - \mathbf{Q}||_2 \qquad (1)$$

This algorithm will compute distances of each of the query set feature vectors from each of gallery set feature vectors and set its nearest neighbor's label as the predicted classification. For the proposed baseline model, rank-1 accuracy is of 47%. The accuracy for different ranks are shown in the graph below: