

InFoMM CDT - Scientific Computing

Image Classification Task

The Yale face database B is a test data set of the faces of 38 individuals with different expressions and under different lighting conditions. The database consists of 2414 images, all of which are 32×32 grayscale, with pixel intensities on a scale from 0 to 255.

Your task is to write a suite of code in MATLAB that can be used to test *supervised* classification on the Yale face database, using *k-nearest neighbours* as the classifier. Your code should provide the user with various choices using input arguments, such as:

- which digits to test
- the size of training/test sets
- the features to use (e.g. intensity, PCA...).

Your code should output performance metrics, such as classification rate and elapsed time, and **it should be possible for someone else to use it easily**.

Taking it further. Beyond this initial brief, the project could be extended in any of the following directions:

- visualization of clustering using PCA projection
- exploration of other (improved?) features
- use of classifiers other than *k-nearest neighbours*
- inclusion of unsupervised classification
- extension to other image databases.

Submission. Each pair/group should submit (using the Maths website) a `.zip` file of their fully documented code, including demos. Each pair/group should also make their code available as a repository on Github.