CSC3067 Group Project

Ross Brown & Rory Flynn

Training the application

Before we started on the application, a learning method and a feature descriptor had to be chosen. Since we weren’t sure which was most appropriate, we decided to implement and test several options, those being HOG and full image for the feature descriptors and SVM, K-NN and nearest neighbour for the learning methods.

Feature descriptors

HOG

HOG is fairly well known in the video analytics world as one of the best methods for identifying objects in images. This is because HOG is able to accurately detect the shape of an object (in this case someone’s face), even if the person is not looking directly into the camera. HOG is also very good at dealing with changes in lighting and scale, which is perfect for face detection on our images.

For HOG testing, a histogram is created of pixel orientations, compared with something like full image descriptor, where the image pixel array gets passed directly to the learning method. The histogram can then be used to attempt to detect objects in images, such as faces. Below is a screenshot of an example HOG generated from one of the test images:

Graphical user interface, application

Description automatically generated

A collage of a person's face

Description automatically generated

As can be seen the HOG data extractor has done a good job at picking out the faces and lines separating the faces.

Full image

Passing the full image to the learning method is another option we could use. Although this was attempted, this is not as effective as using an algorithm before passing the image to the learning method.

Full image descriptor would not allow the learning method to distinguish the foreground and background of an image and is very sensitive to variance. This would make it very hard for the learning method to differentiate between the face and the background wall of the images.

A better approach would be to use HOG to detect the faces before passing it to the learning method. This would allow the learning method to focus on the faces themselves, and not have to worry about the background. This would be more effective and would be able to more accurately identify the faces.

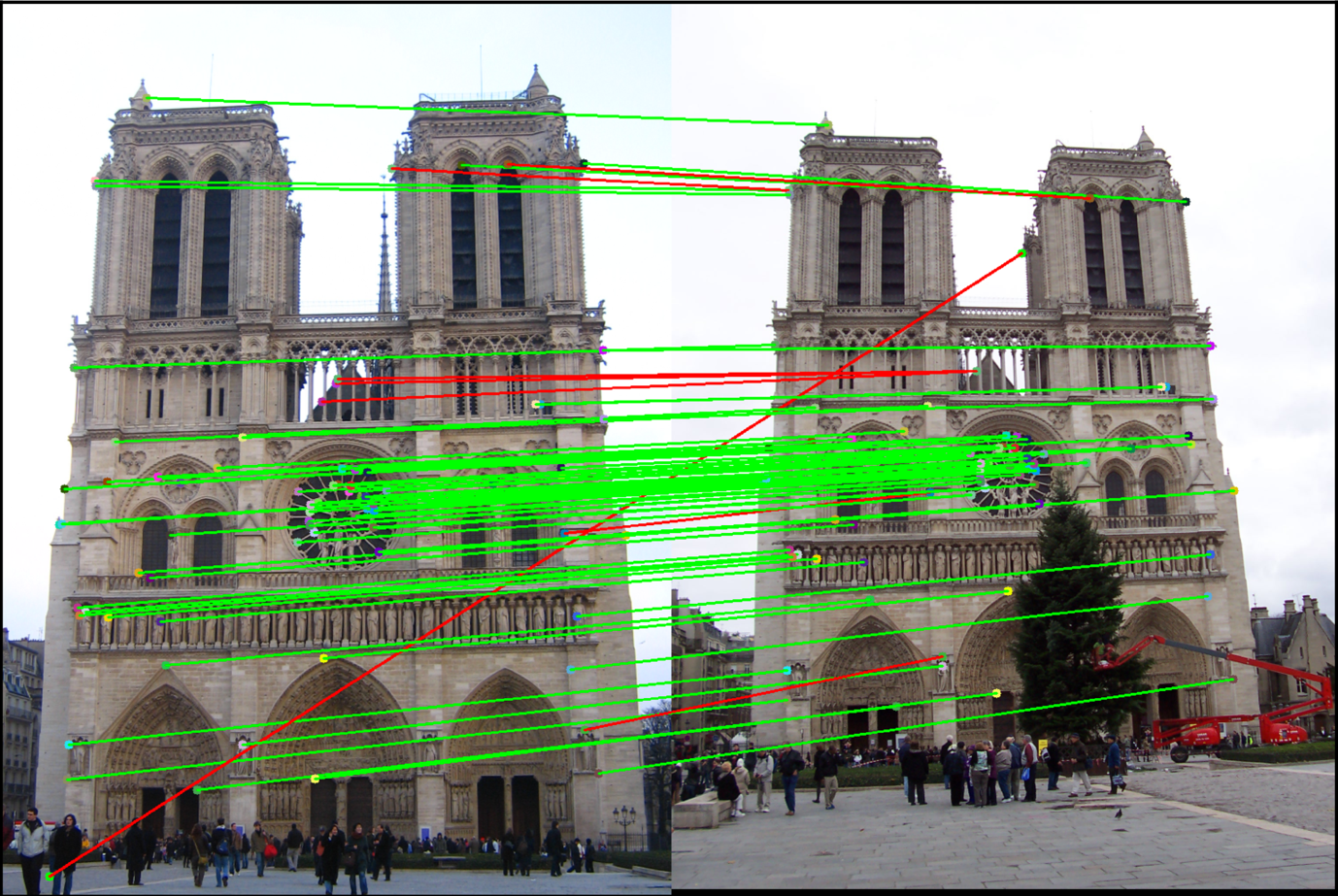
Gabor filtering

Using Gabor filtering is another feature descriptor for extracting information from an image before being passed to the learning method, in a similar way to HOG.

Gabor filters are mainly used for edge detection, which should help us in detecting the edges of faces. They do this by measuring the frequency and orientation of the image, and are very good at extracting specific objects from images. They’re also very widely used in facial recognition systems, so this method would be perfect for our use case.

MSER, SIFT, SURF and BRISK

We also briefly used MSER, SIFT, SURF and BRISK feature descriptors, which attempt to highlight key areas of images. Using these methods they can highlight points in the image which are areas of interest, or can attempt to match parts of the image. For example with SIFT, it works by matching up features in images. This is a good example visually demonstrating what the process looks like:



Ultimately, while these methods showed promise, they were incompatible with our dataset. If we were to make improvements to the system in future, these descriptors could be reconsidered.

Learning methods

SVM