

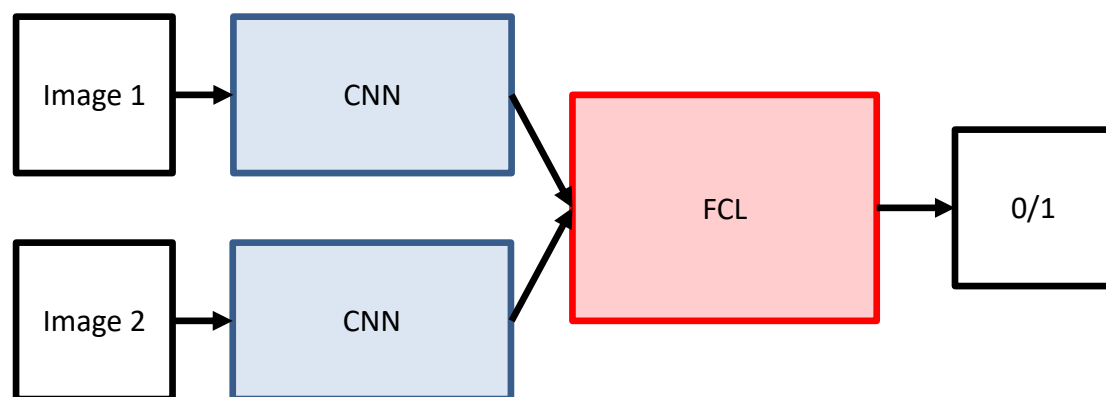
Design Spec

This project uses convolutional neural networks to extract features from two images, which are then compared by several fully connected layers to determine if two images have the same center, but from different (or identical) camera positions.

The convolutional networks that extract features from the images are identical, as order shouldn't make a difference to whether the center is the same.

Overview

The overall structure of this system looks like this



where both convolutions have the same weights. The convolutional networks and fully connected network are described further down.

Convolutions

Each layer consists of a convolutional layer and 2x2 pooling with stride 1. The convolutional networks are created with the following properties

Input

Filter size:	N/A
Input features:	N/A
Output features:	3
Pre-pooling shape:	N/A
Output shape:	(63, 63, 3)

Convolution 1

Filter size:	5
Input features:	3
Output features:	9
Pre-pooling shape:	(59, 59, 9)
Output shape:	(29, 29, 9)

Convolution 2

Filter size: 3
Input features: 9
Output features: 32
Pre-pooling shape: (27, 27, 32)
Output shape: (13, 13, 32)

Convolution 3

Filter size: 3
Input features: 32
Output features: 64
Pre-pooling shape: (11, 11, 64)
Output shape: (5, 5, 64)

Fully Connected Layers

After getting the features from the CNNs, we compare the two images with a fully connected network. Its goal is to shrink these many inputs to one output. They have the following properties

Layer 1

$1,600 \times 2 \rightarrow 1,024$

Layer 2

$1,024 \rightarrow 128$

Layer 3

$128 \rightarrow 1$

Output Layer

The output layer has size 1 and is a true/false (1/0) value that *should* be 1 iff the center of both images are the same, and 0 otherwise.

Footnotes

Last updated v0.0.1

Please note that values and descriptions may have changed since. Check the code for accurate values.