Siamese Neural Networks for One Shot Image Recognition

Koch, Gregory, Richard Zemel, and Ruslan Salakhutdinov. "Siamese neural networks for one-shot image recognition." *ICML deep learning workshop*. Vol. 2. 2015.

What animal is this?

One Shot Learning

▷ One-shot learning is a classification task where one example is used to classify many new examples in the future.

▷ It is distinguished from zero-shot learning, in which the model cannot look at any examples from the target classes

▷ It is easy to generalize this to k-shot learning by having k examples for each class rather than just one

Omniglot Dataset

▷ Consists of 1623 hand drawn characters from 50 different languages.
▷ For every character there are just 20 examples.
▷ Each image is a gray scale image of resolution 105x105.

Omniglot Dataset

Baseline Model - KNN

▷ Calculate the Euclidean distance of the test example from each training example and pick the closest one.

▷ 1-nn gets 22% accuracy in one shot classification on omniglot.

Siamese Networks

▷ A Siamese network is an architecture with two parallel neural networks, each taking a different input.

▷ Takes two images as input and train it to guess whether they have the same category.

▷ When doing a one-shot classification task, the network can compare the test image to other images, and pick which one it thinks is most likely to be of the same category.
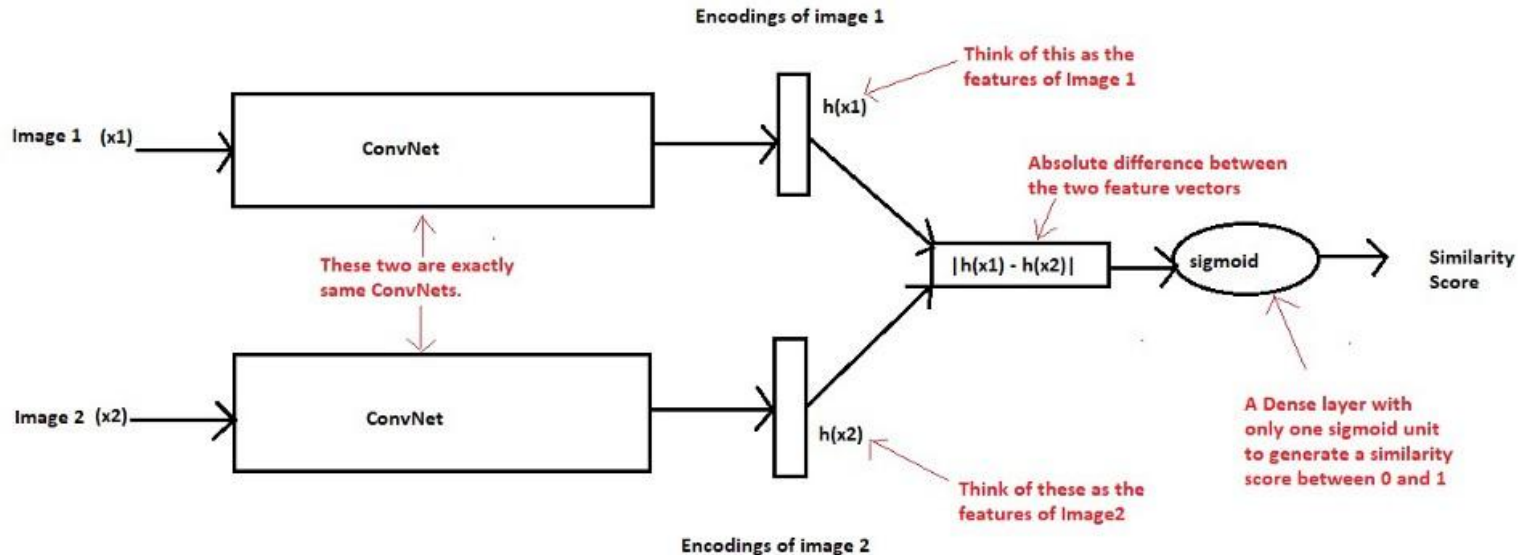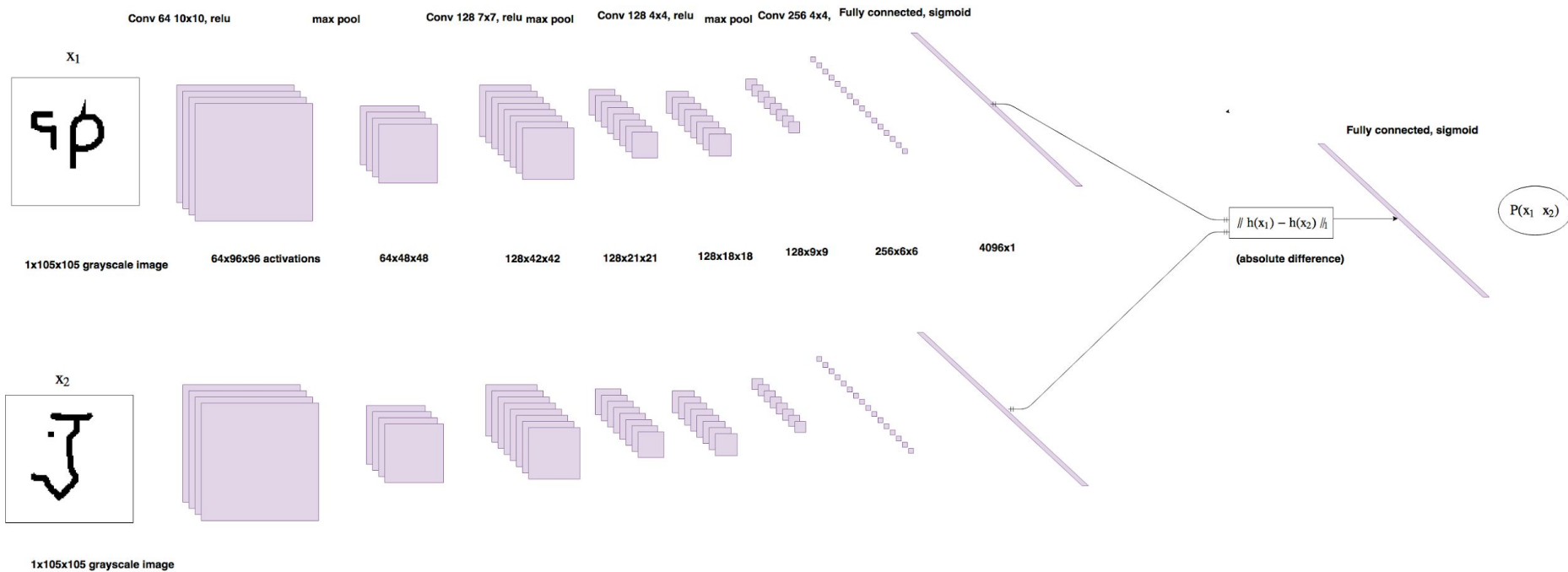
Siamese Networks
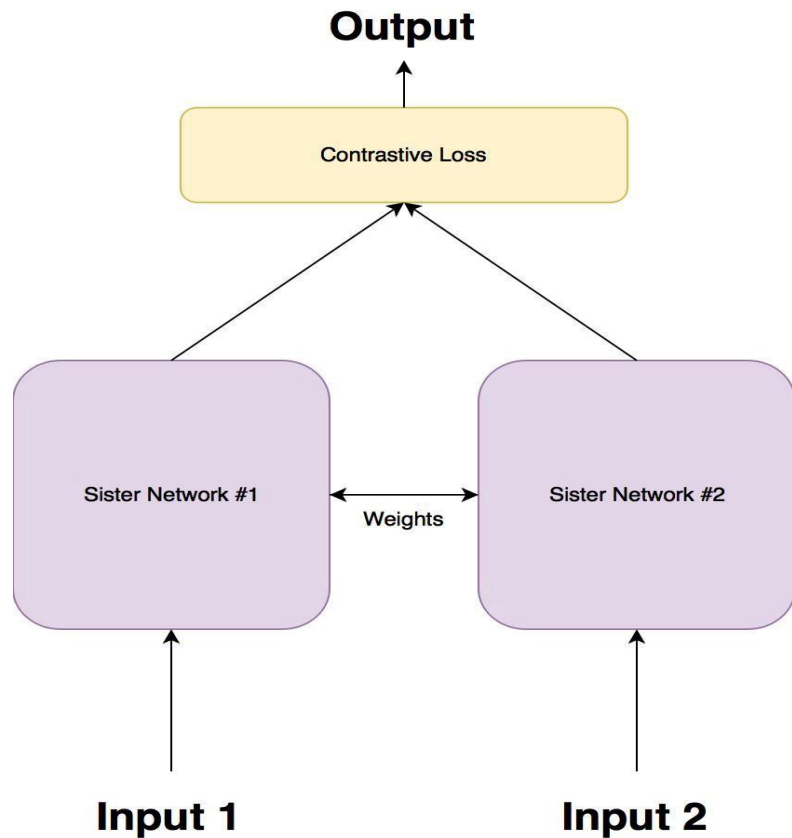
Siamese Networks - Training

1. Pass the first image of the image pair through the network.
2. Pass the 2nd image of the image pair through the network.
3. Calculate the loss using the outputs from 1 and 2.
4. Back propagate the loss to calculate the gradients.
5. Update the weights.

Siamese Networks - Testing

Conv 64 10x10, relu   max pool   Conv 128 7x7, relu max pool   Conv 128 4x4, relu   max pool   Conv 256 4x4,   Fully connected, sigmoid

x_1

1x105x105 grayscale image   64x96x96 activations   64x48x48   128x42x42   128x21x21   128x18x18   128x9x9   256x6x6   4096x1

x_2

1x105x105 grayscale image

$\| h(x_1) - h(x_2) \|_1$

(absolute difference)

Fully connected, sigmoid

$P(x_1 \quad x_2)$

Results

Method	Test
Humans	95.5
Hierarchical Bayesian Program Learning	95.2
Affine model	81.8
Hierarchical Deep	65.2
Deep Boltzmann Machine	62.0
Simple Stroke	35.2
1-Nearest Neighbor	21.7
Siamese Neural Net	58.3
Convolutional Siamese Net	92.0

Thank you!