Image Recognition

G. Koch, R. Zemel, R. Salakhutdinov, Siamese Neural Networks for One-shot Image Recognition, ICML deep learning workshop. Vol. 2. 2015.

Motivation

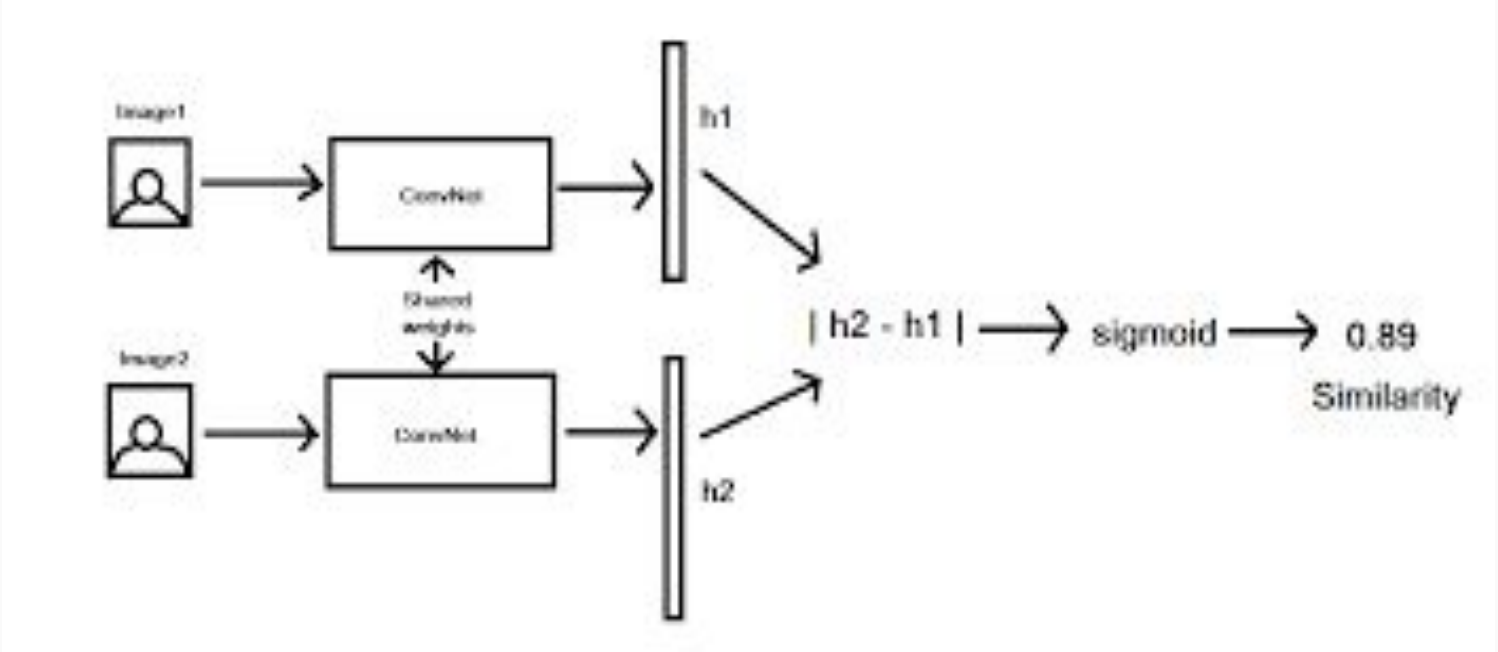
Deep neural networks are really good at learning from high dimensional data like images or spoken language, but only when they have huge amounts of labelled examples to train on. Humans on the other hand, are capable of *one-shot learning* - if you take a human who's never seen a spoon before, and show them a single picture of a spoon, they will probably be able to distinguish spoons from other kitchen utensils with astoundingly high precision.

Recently there have been many interesting papers about one-shot learning with neural nets and they've gotten some good results and one of them is using **Siamese Network.**

Idea

It is an approach to getting a neural net to do one-shot classification is to give it two images and train it to guess whether they have the same category. Then when doing a one-shot classification task described above, the network can compare the test image to each image in the support set, and pick which one it thinks is most likely to be of the same category. So we want a neural net architecture that takes two images as input and outputs the probability they share the same class.

Architecture of Siamese Network



Methodology

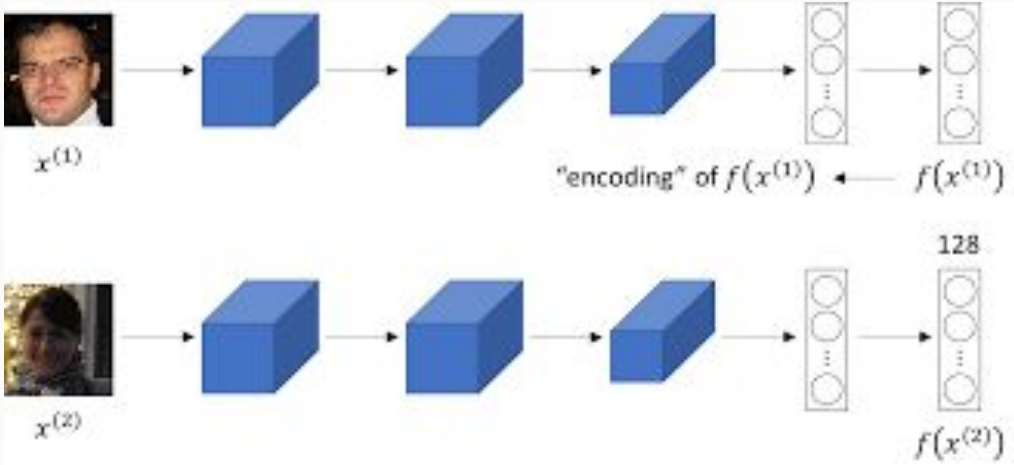
- We take two images (*Image1* and *Image2*). Both the images are fed to a single Convolutional Neural Network (CNN).

- The last layer of the CNN produces a fixed size vector (embedding of the image). Since two images are fed, we get two embeddings. (*h1* and *h2*).

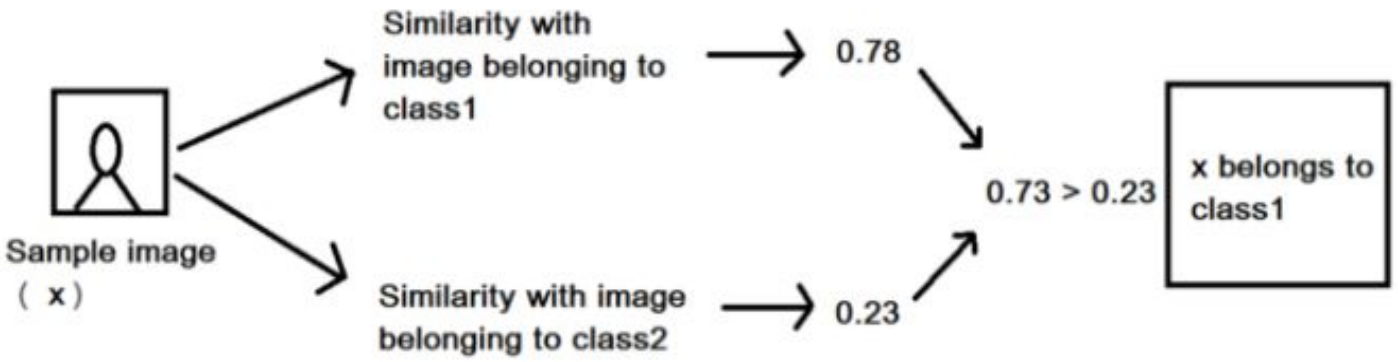
Methodology

- The Euclidean distance between the vectors is calculated.

- The values then pass through a sigmoid function and a similarity score is produced.



- If the difference between the vectors h1 and h2 is large, lower is the probability that they belong to the same class and vice versa.


Similarity with
image belonging to
class1
0.78
Sample image
(x)
Similarity with image
belonging to class2
0.23
0.73 > 0.23
x belongs to
class1

Siamese Network over CNN.

- Need only a handful of images from each class or of a person.
- They learn similarity functions. Hence, they don’t classify, but, differentiate images.
- Easy to train and have a simple architecture.

Thank You!