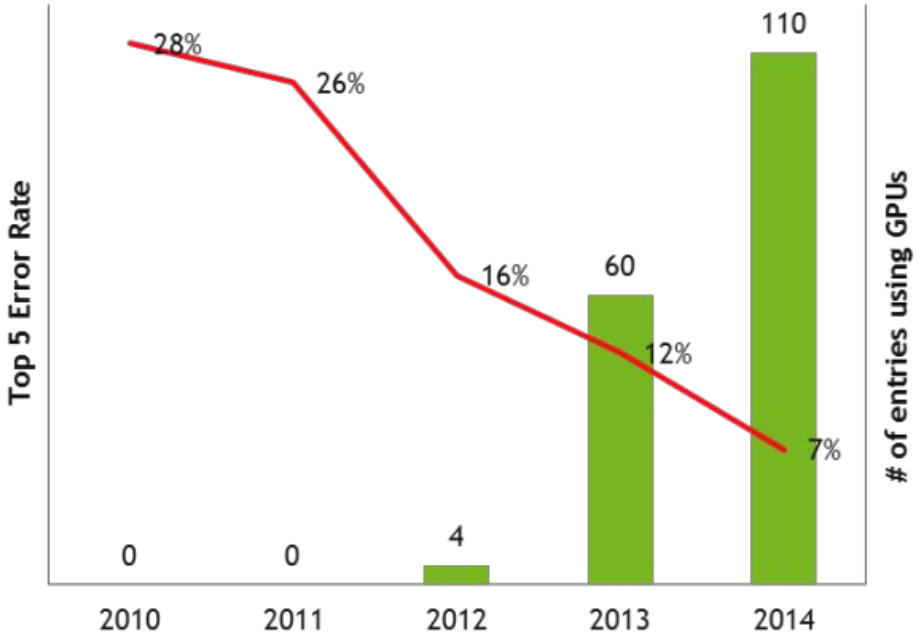
Image Classification based on CNN

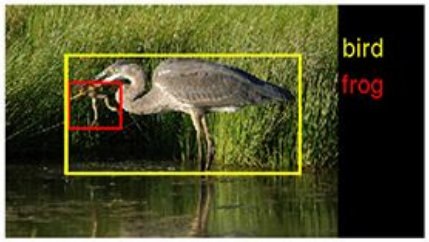
T11 - PangCloud (Tianyang Liu, Yicheng Wang, Su Pu)

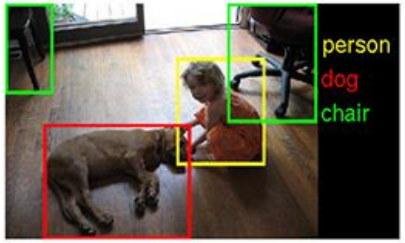
Motivation and Background

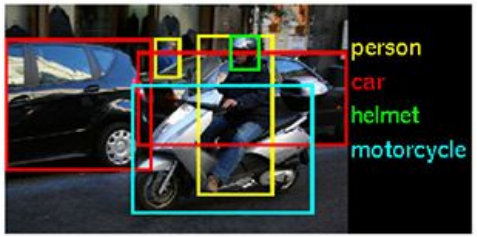
Deeper learning! Need GPU! Cloud! Cloud! Cloud!

IMAGENET









Reference: http://image-net.org/

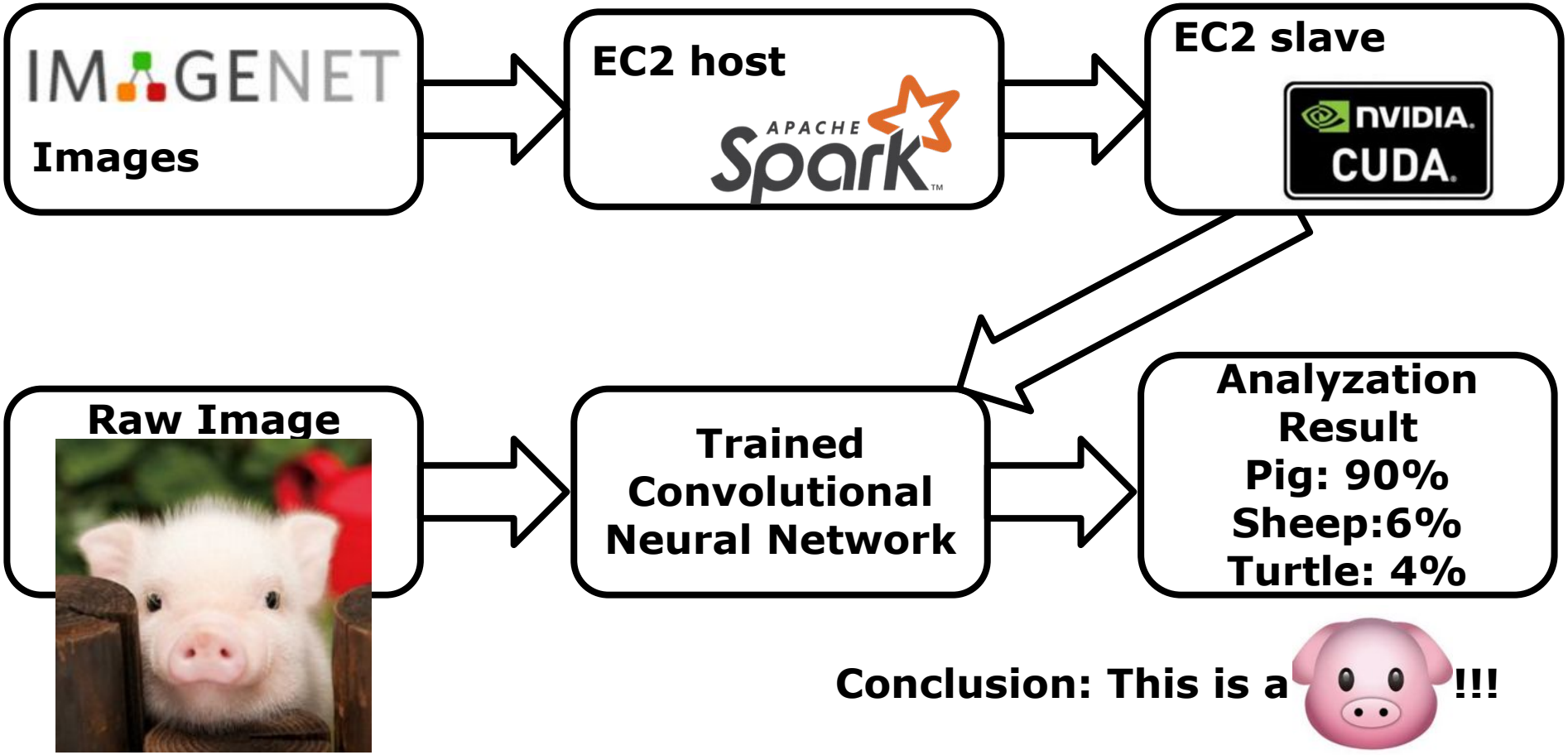
Why Cloud?

OR

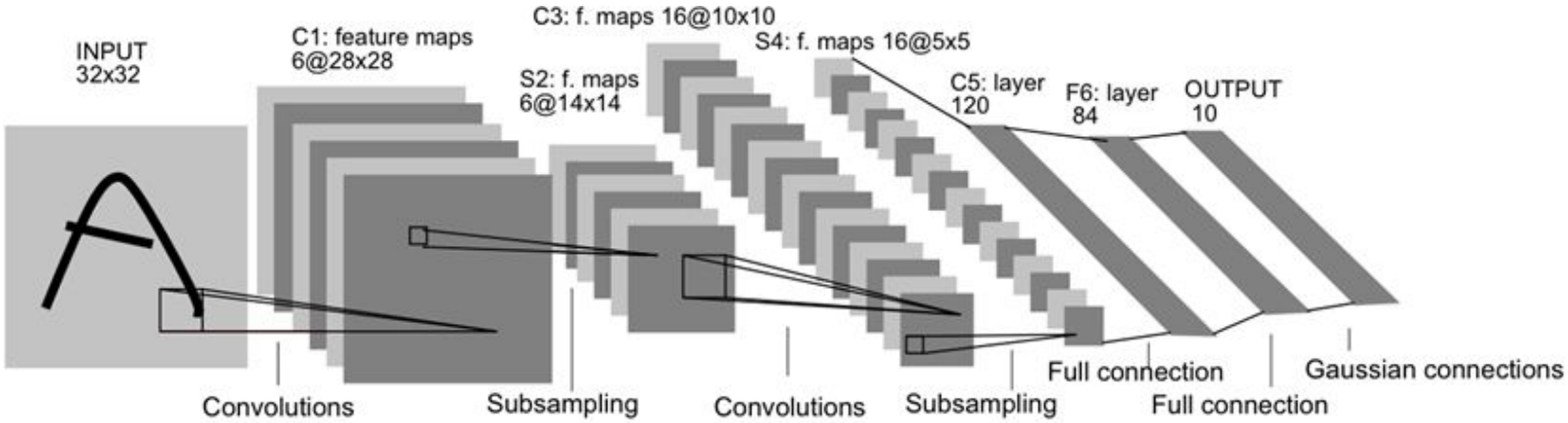
Hey Siri, what's this?

Aipoly Vision App which recognizes objects on iphone, but is limited on hundreds of objects.

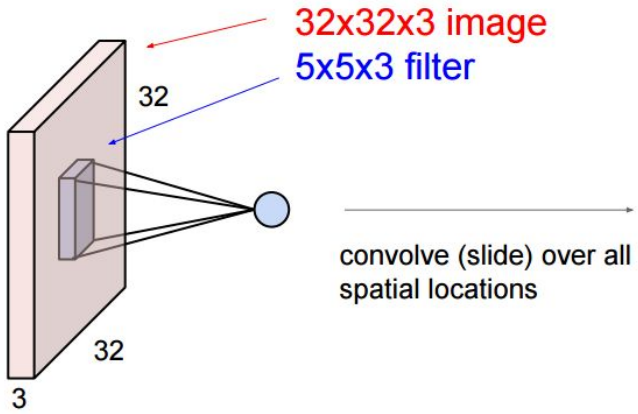
Overview Architecture



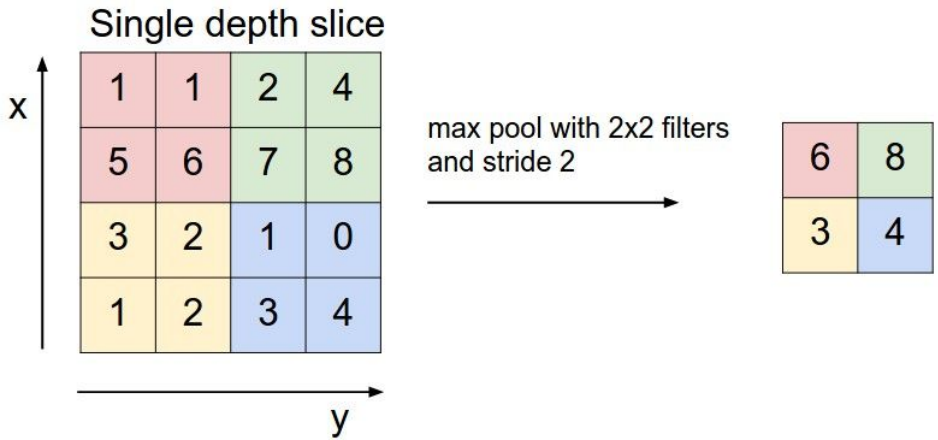
Convolutional Neural Networks



Convolution Core



Max Pooling



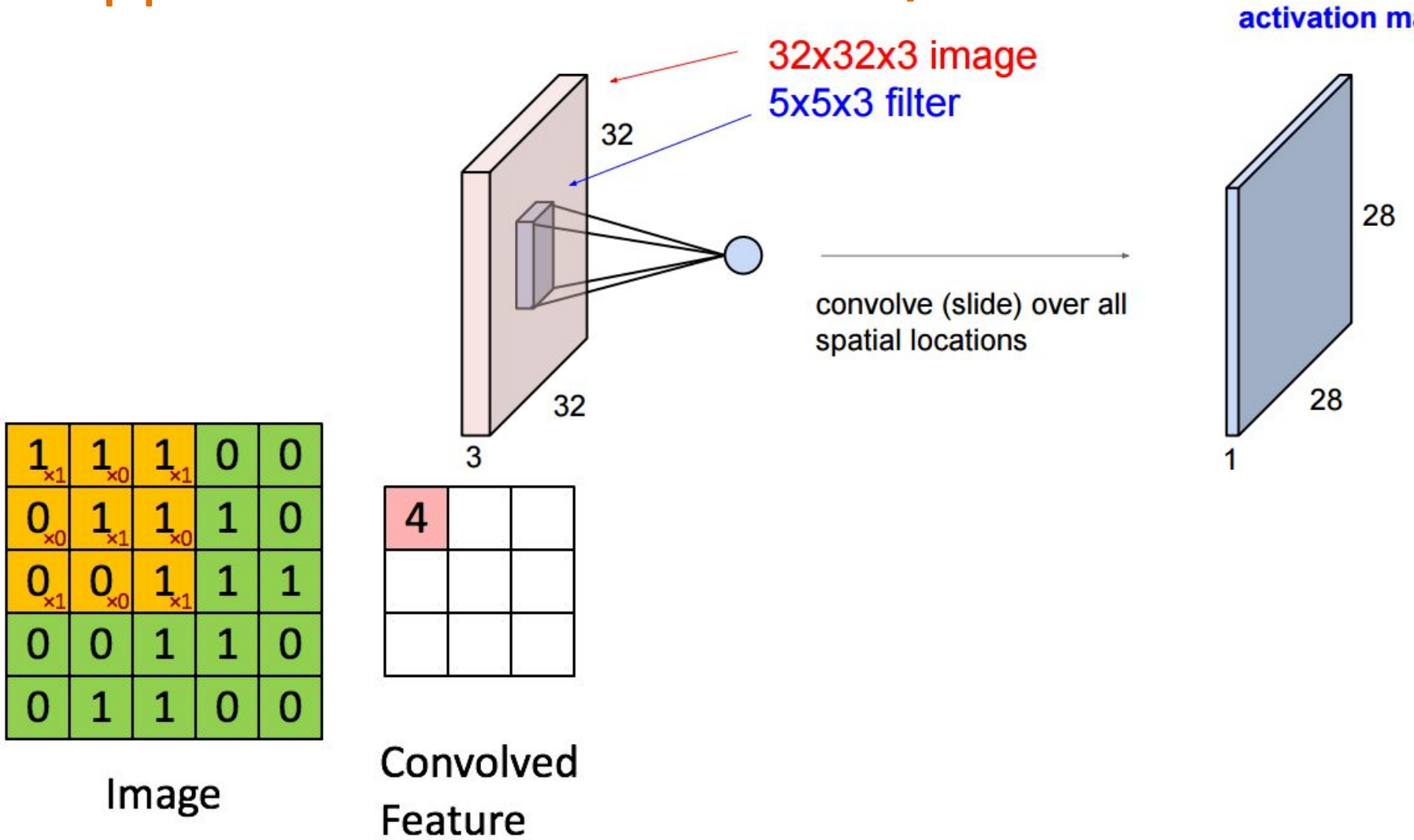
Assumption Result

Evaluation criteria: error rate

The first application of CNN in Large Scale Visual Recognition Challenge (ILSVRC) achieved 15.3% error rate, there were some improvement on the structure in the following years.

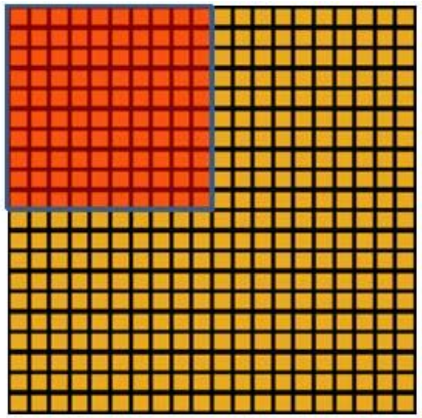
So we expect to achieve around **10%** error rate.

Appedix 1. Convolutional layer

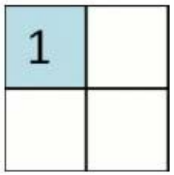


Appedix 2. Pooling layer

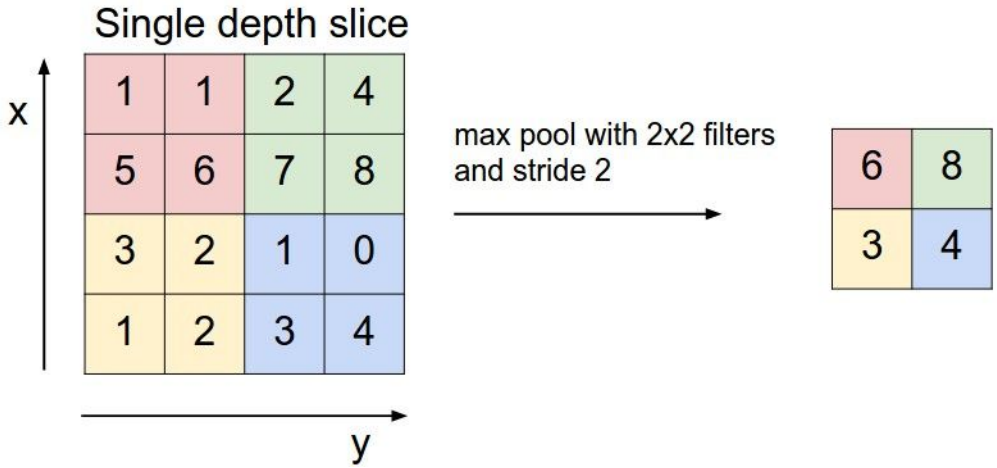
Down sample to make the representations smaller and more manageable



Convolved feature

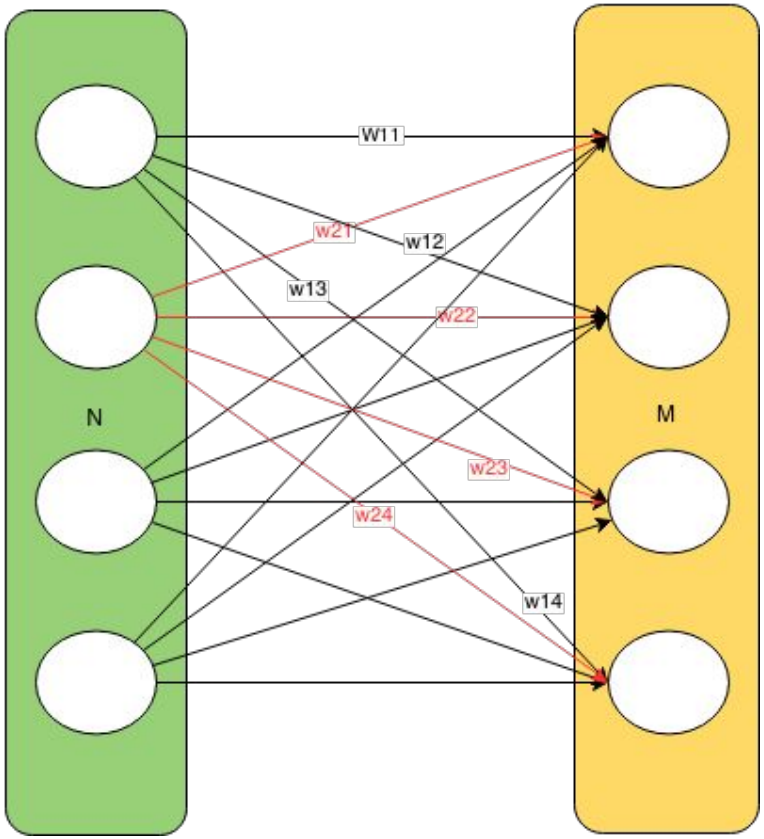


Pooled feature



Max Pooling

Appedix 3. Fully-Connected layer



Just like the connection method in ordinary NN, the neurons of one layer connect all the neurons in the previous layer, which means there are N*M parameters (N the number of neurons in present layer and M for the previous layer)