The Renaissance of Neural Networks

Prof. Anita Wasilewska

Group No. 1

Huidong Liu, ID: 110529130

Meng Luo, ID: 110464666

Cheuk On Chung ID: 110386696

From where we've learned a lot…

- DeepFace: Closing the Gap to Human-Level Performance in Face Verification, CVPR, 2014.
- Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification, ICCV'14
- G.E. Hinton, et al., Reducing the Dimensionality of Data with Neural Networks, Science, 2006.
- Stanford CNN course: http://cs231n.stanford.edu/

Cite: http://tipshomedesign.xyz/wp-content/uploads/2015/12/good-idea-light-bulb.jpg

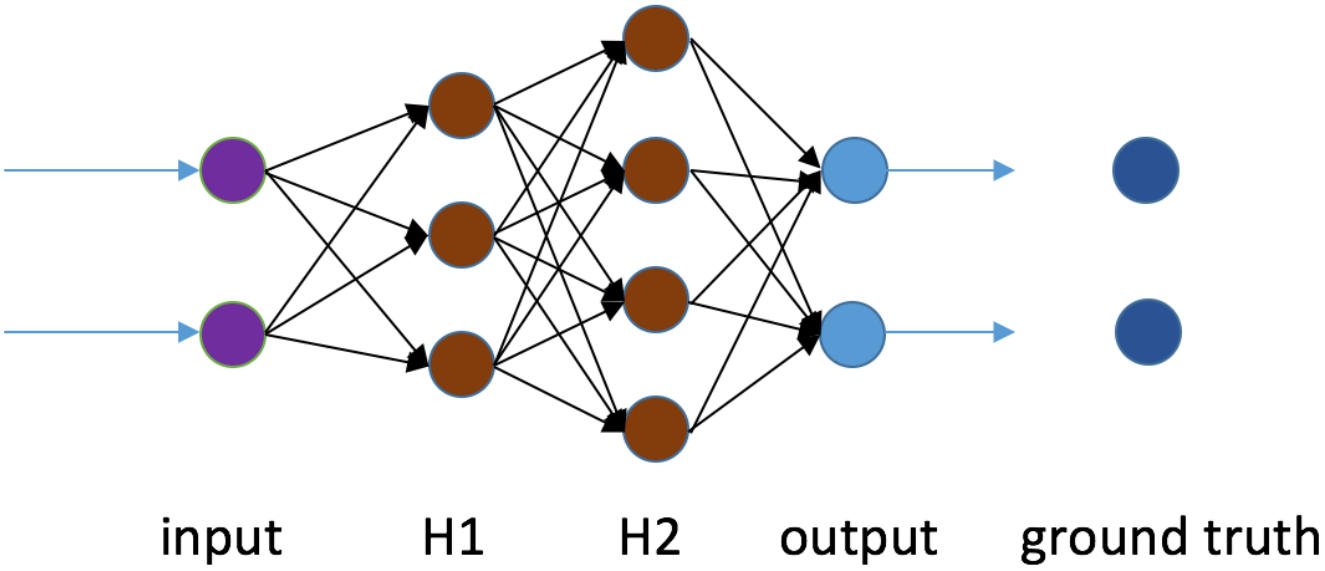
Outline

- Deep Learning: A breakthrough
- Deep Learning vs Traditional Neural Networks
- Milestone1('06): Stacked Auto Encoders (SAE)
- Milestone2('12): Convolutional Neural Networks (CNNs) applied to image recognition.

Deep Learning

Definition:

- A branch of machine learning, which is able to train multi-hidden (>1) layer Neural Networks efficiently to learn information from data.



Cite: https://en.wikipedia.org/wiki/Deep_learning

Why deep learning a breakthrough?

Face recognition

For real world face recognition (LFW), computer beats human.

- Human: 97.53%
- Deepface (facebook): 97.35%, 2014
- FaceNet (Google): 99.63%, 2015

Source:

[1] DeepFace: Closing the Gap to Human-Level Performance in Face Verification, CVPR, 2014.

[2] Labeled Faces in the Wild: A Survey, Advances in Face Detection and Facial Image Analysis, 2015

Why deep learning a breakthrough?

Image classification

ImageNet: 1 million images collected from the web belonging to 1000 objects.

- Human: 94.9%
- PReLU (Microsoft): 95.06%, 2015
- ResNet-152 (Microsoft): 95.51%, 2015

Source:

[1] Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification, ICCV'14

[2] Deep Residual Learning for Image Recognition , arXiv, 2015

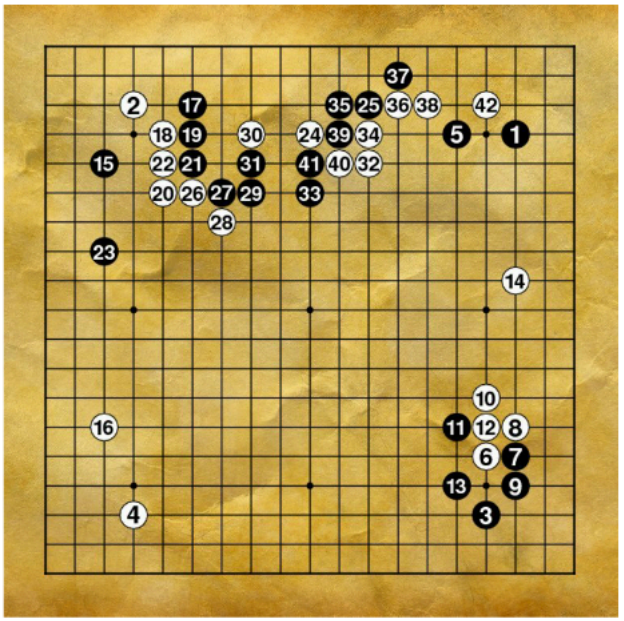
Why deep learning a breakthrough?

Playing the Game of Go

- Game of Go: 19-by-19 board.
  The most difficult game in the world.

- AlphaGo *vs* Lee Sedol
  Date: Mar 9 – 15, 2016

- AlphaGo beat Lee Sedol by 4-1



Deep Learning vs Traditional Neural Networks (before 2006)

- Traditional NN (TNN): Non-convex, hard to optimize.

  DL: layer-wise greedy training for parameter initialization.

- TNN: Needs huge amount of data to train a good model.

  DL: We have huge amount of data now.

- TNN: Computers couldn't bear huge computation at that time.

  DL: 1. Hardware are very fast now. 2. Code can run in parallel.

Outline

- Deep Learning: A breakthrough
- Deep Learning vs Traditional Neural Networks
- Milestone1('06): Stacked Auto Encoders (SAE)
- Milestone2: Convolutional Neural Networks (CNNs) applied to image recognition.

Stack Auto Encoders (SAE)

NN is non-convex. Initialization is important.

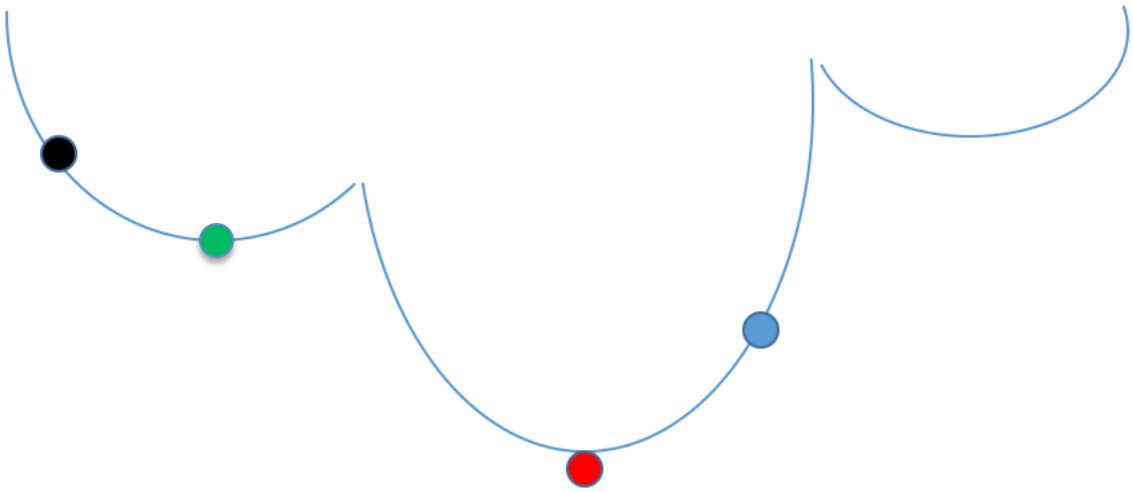
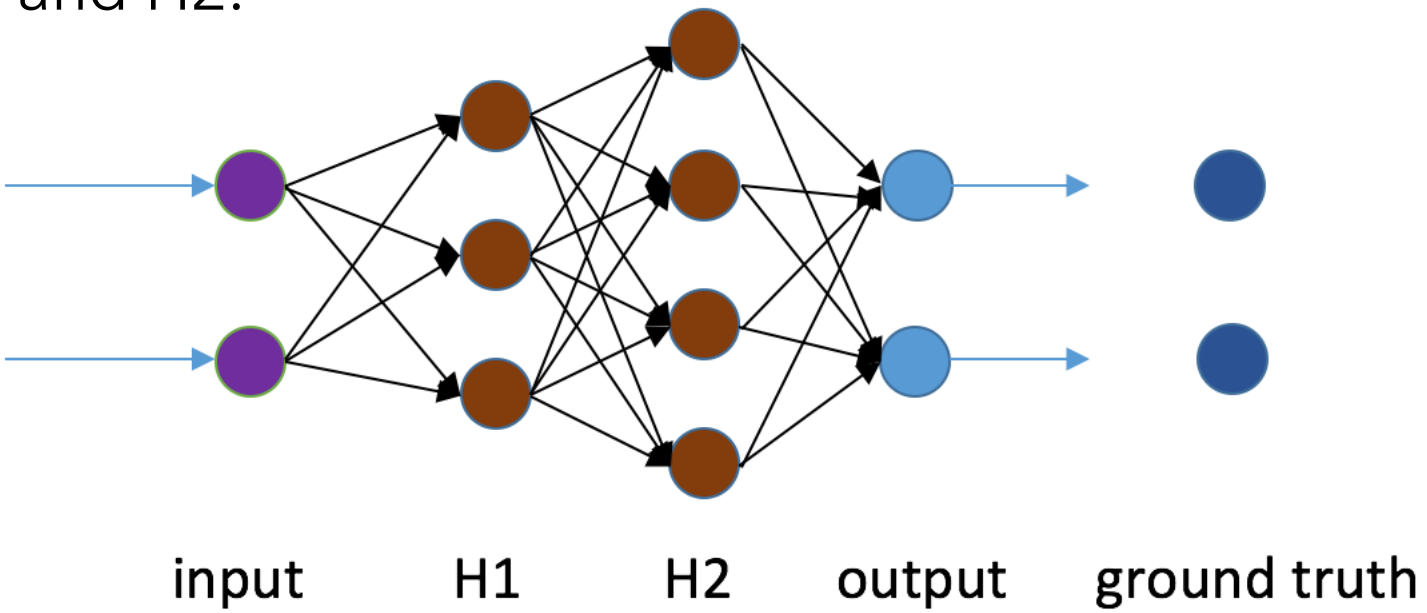


Figure 1: Illustration of an objective function of a NN.

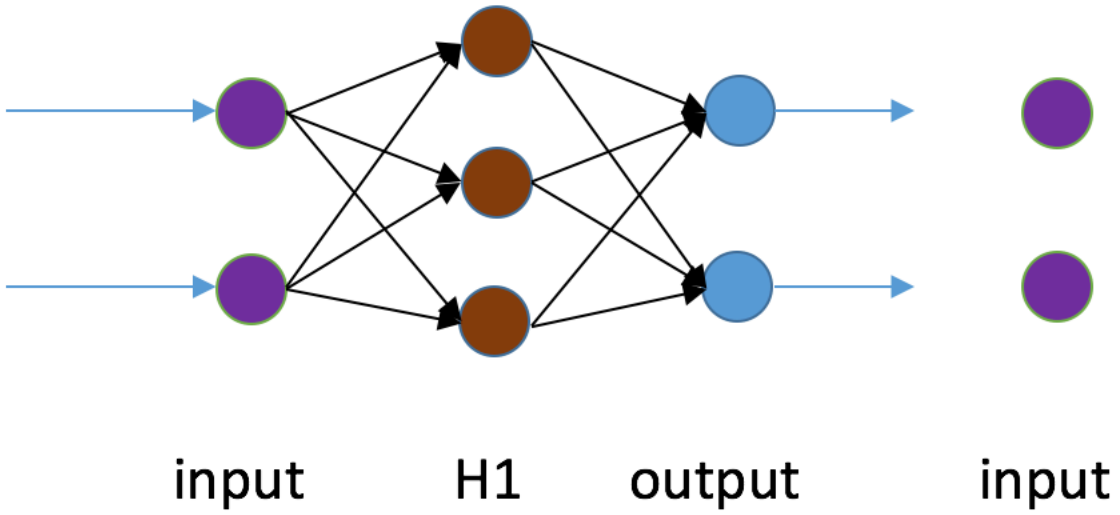
Stack Auto Encoders (SAE)

Suppose we want to train a NN with two hidden layers: H1 and H2.



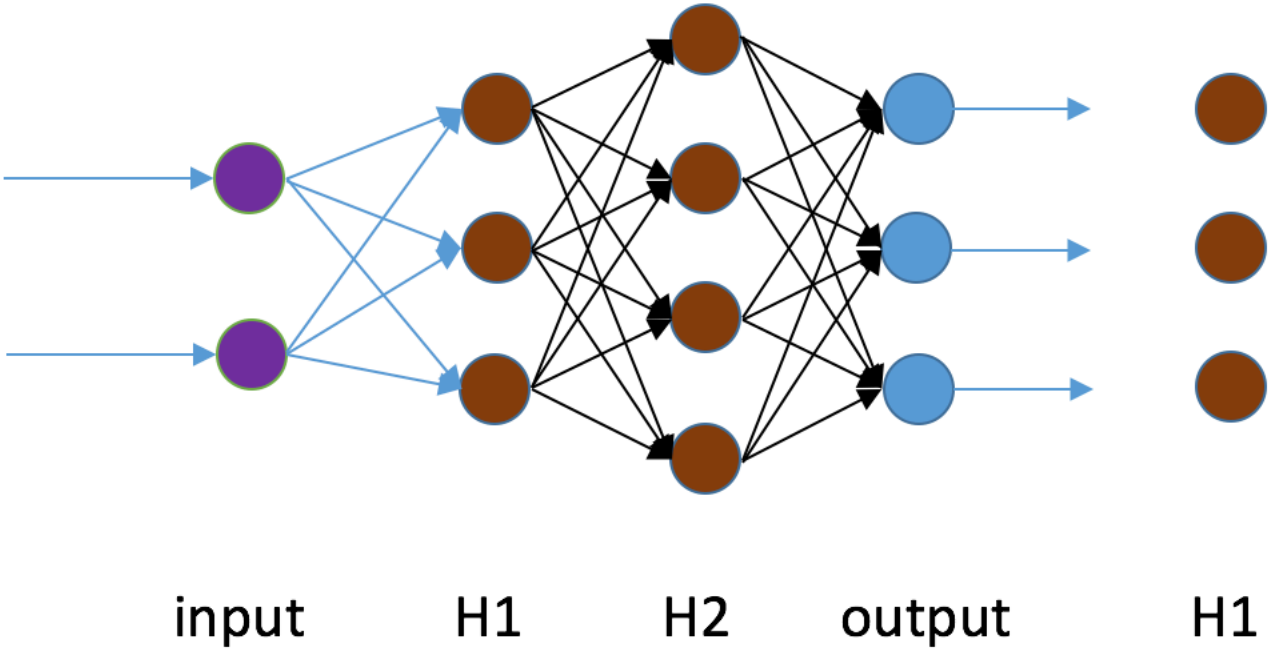
Stack Auto Encoders (SAE)

SAE performs greedy layer-wise training.



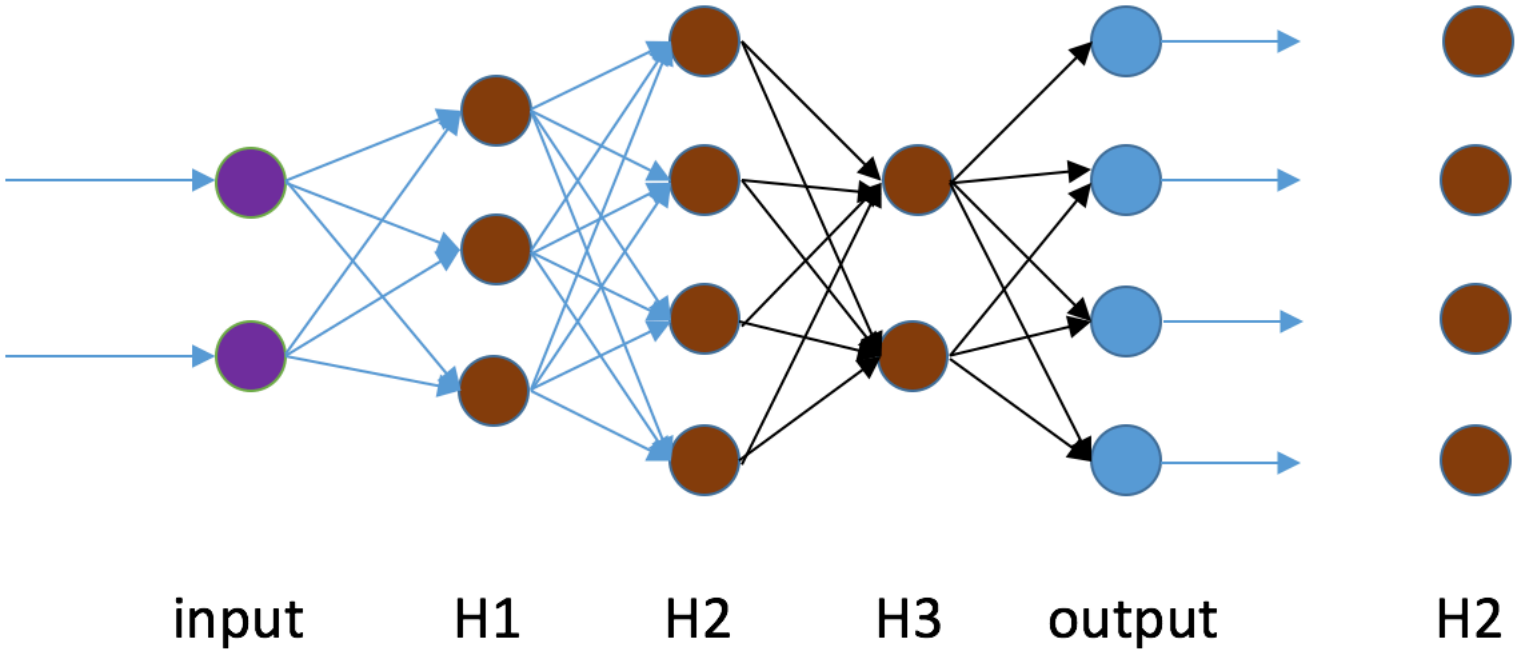
Stack Auto Encoders (SAE)

When weights in previous layers are trained, they are frozen.



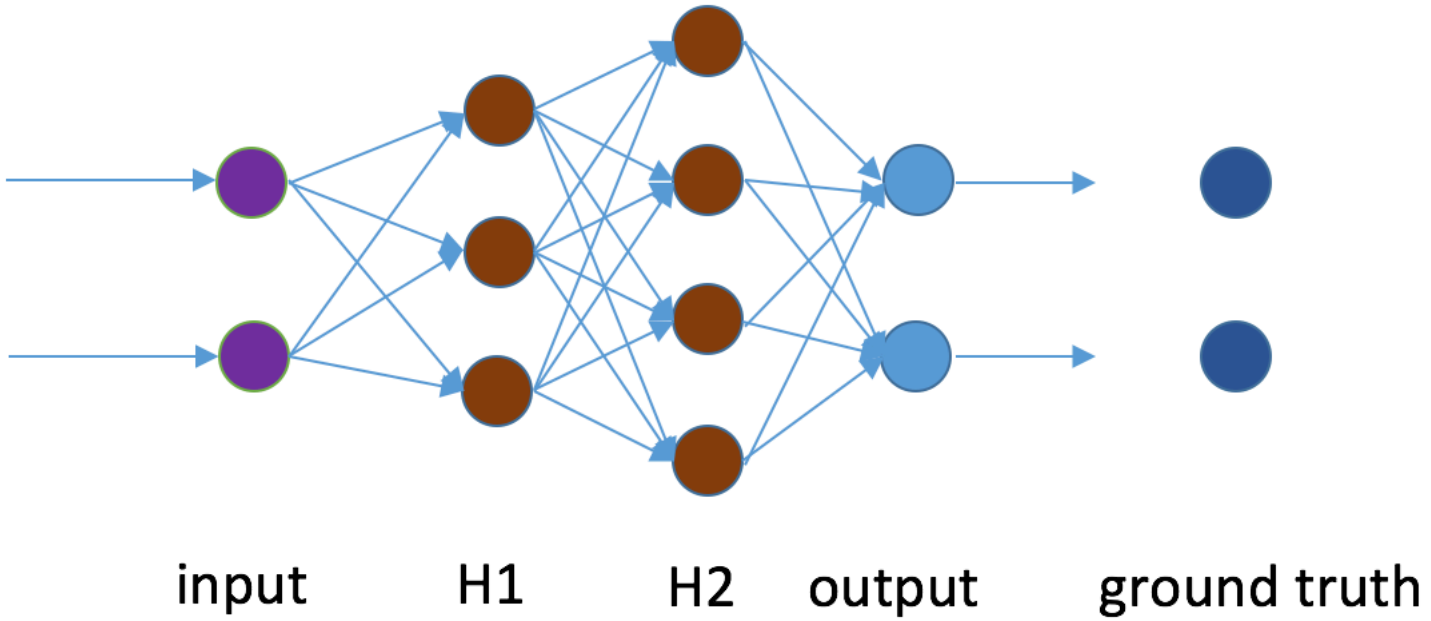
Stack Auto Encoders (SAE)

Move to the next layer and repeat this process.



Stack Auto Encoders (SAE)

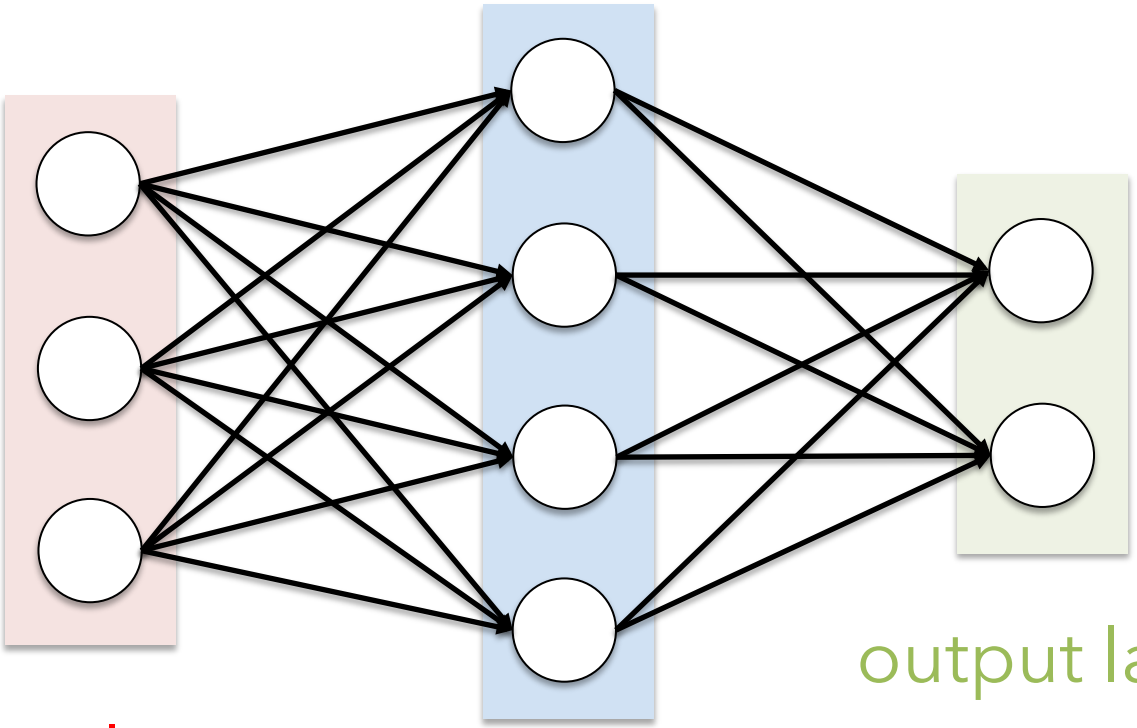
When weights in all layers are trained, they are used as initialized weights when training the whole network.



Outline

- Deep Learning: A breakthrough
- Deep Learning vs Traditional Neural Networks
- Milestone1: Stacked Auto Encoders (SAE)
- Milestone2('12): Convolutional Neural Networks (CNNs) applied to image recognition.

Recap of ordinary Neural Networks



A typical image

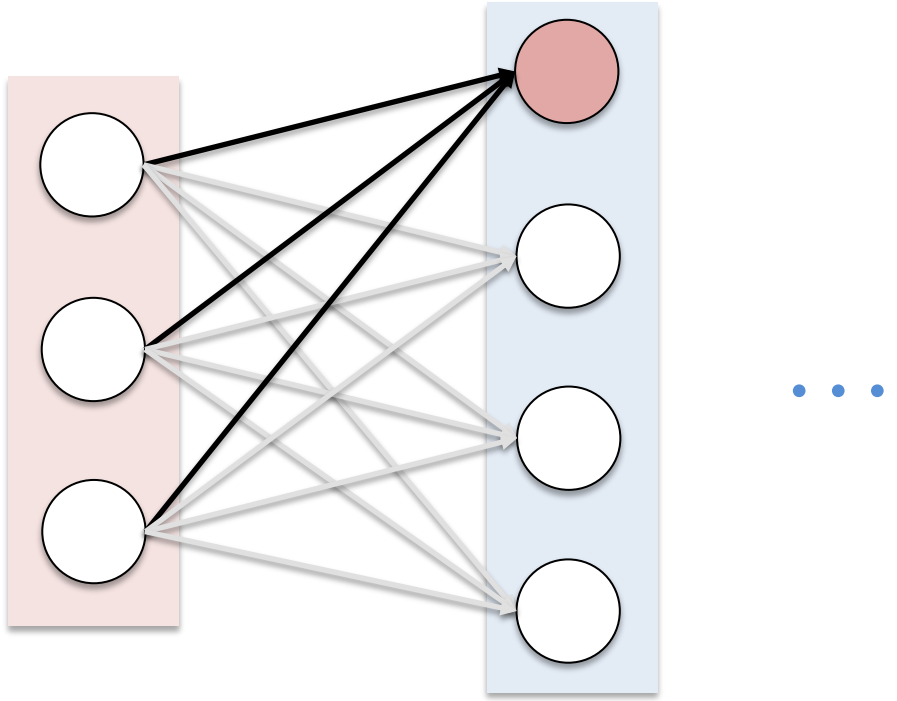
Red Green Blue 32 * 32 * 3

Cite: http://www.adsell.com/images/color-channels-RGB.jpg

How image recognition is applied…

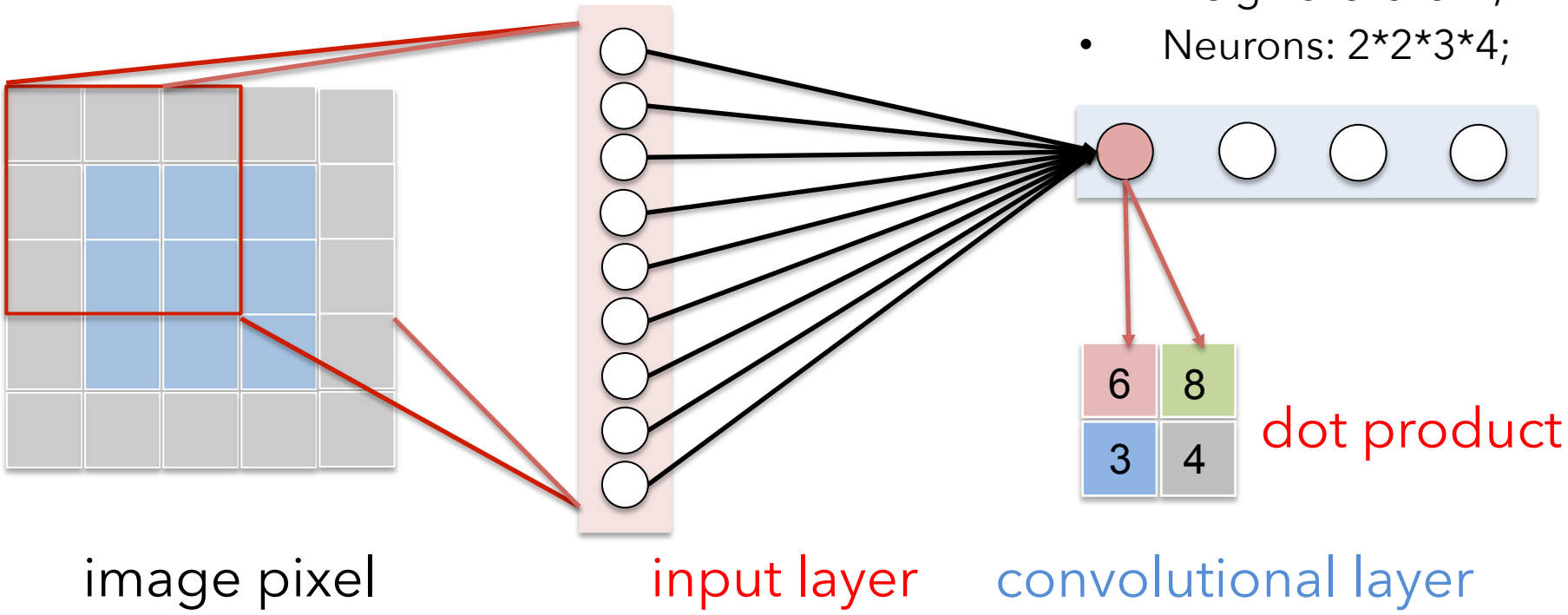
- NN: Weights for a single fully-connected neuron

  32 * 32 * 3, if the image is larger, 200*200*3? (expensive)
- Convolutional NN



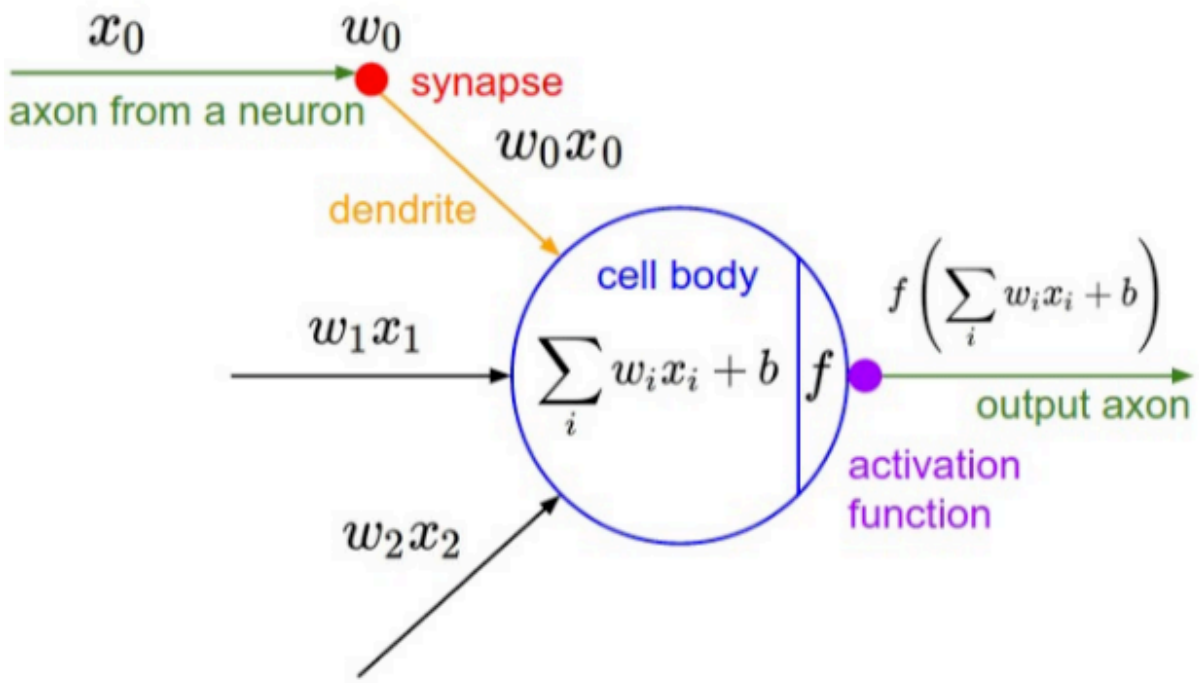
Convolutional Layer

- Input volume is Locally-connected to neurons.
- The example is a single channel.
  - Weights: 3*3*3*4;
  - Neurons: 2*2*3*4;



* Zero-padding = 1, stride = 2, Depth = 4 is a set of neurons connect to the same region.

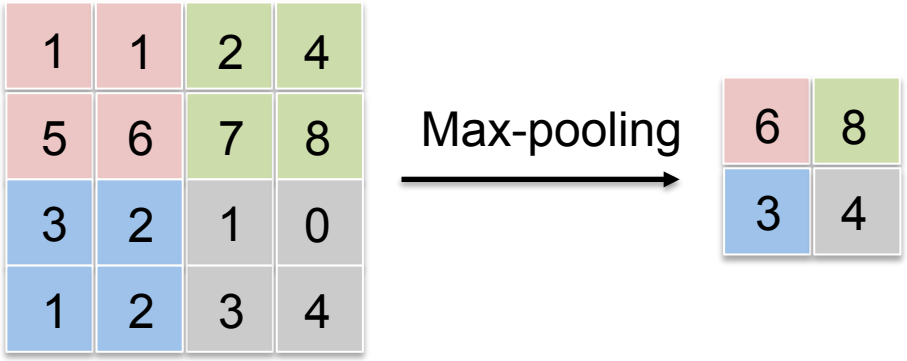
RELU layer



* Elementwise non-linearity activation function

Pooling Layer

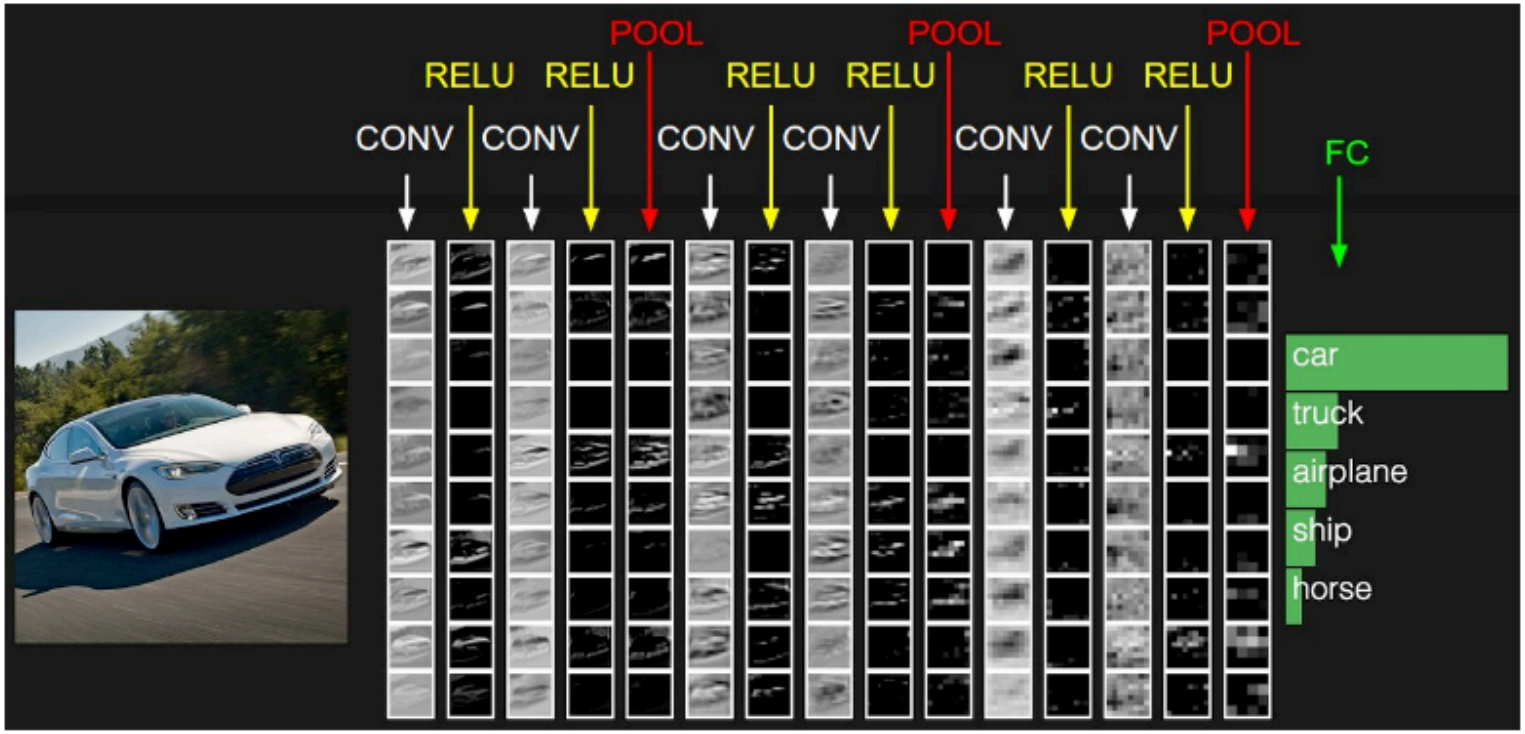
- Reduce the spatial size of the representation to reduce the amount of parameters.
- Robust to translation and rotation.
- Insert a Pooling layer in-between successive Conv-layers.



size = 2, stride = 2

CNN Architecture

- The fully-connected layer is similar to Neural Networks.



Thank you