Kuncheng Feng
CSC 466

# Chapter 5: ConvNets and ImageNet

## → Reading/Mining/Discussion Assignment

1). He combined ideas from Fukushima's neocognitron with the back-propagation algorithm to create the semi-eponymous "LeNet" – one of the earliest ConvNets. Who was he? Who was his postdoctoral advisor? And, in addition to developing "LeNet", what was his contribution to the field of neural network research?

Yann LeCun and his adviser Geoffrey Hinton, along with other neural network loyalists believed that improved, larger versions of ConvNets and other deep networks would conquer computer vision if only they could be trained with enough data. They kept working on the sidelines throughout the 2000s, until in 2012 the researchers won a computer-vision competition on an image data set called ImageNet.

2). What is WordNet? Please answer by referencing its creator (the human most responsible for its existence), saying something about its use, and presenting, in just two or three sentences, a high-level description of its structure?

Created by Princton professor Geroge Miller, it was a database of English words, arranged in a hierarchy moving from most specific to most general, with groupings among synonyms. For example:

Cappuccino → Coffee → Beverage → Food → Substance → Physical entity → entity

WordNet has been used extensively in research by psychologists and linguists as well as in AI natural-language processing systems.

3). What is ImageNet? Please answer by referencing its creator (the human most responsible for its existence), saying something about its use, and presenting a high-level description of its structure, being sure to mention WordNet in doing so.

Created by Fei-Fei Li, it is an image database that is structured according to the nouns in WordNet, where each noun is linked to a large number of images containing examples of that noun.

4). Specifically, what role did ImageNet play in the advancement of computer programs dedicated to the problem of "object recognition."

ImageNet was intended to serve as a database that links nouns to images that actually illustrate that noun.

5). What is the Mechanical Turk (Amazon's Mechanical Turk), and what role did it play in building ImageNet?

According to Amazon, its Mechanical Turk service is 'a marketplace for work that requires human intelligence'. It's a service that connects requesters with workers, or people who need a task accomplished that is hard for computers and people who are willing to lend their human intelligence to a requester's task for a small fee.

The tens of thousands of workers there helped sort out irrelevant images for each of the WordNet terms quickly, at a relatively low cost.

6). Describe the nature and operation of the ImageNet competition.

It's a large-scale visual recognition challenge that aims to spur progress towards more general object-recognition algorithms. It gives competitors 1.2 million labeled training images and a list of possible categories, and tasks them to output the correctness of each input image.

7). What was the most notable thing about the 2012 ImageNet competition?

The winner (AlexNet) achieved an amazing 85 percent correctness, a big jump from previous years. Furthermore it did not use support vector machines, instead it used a convolutional neural network.

8). What was the most notable thing about the 2015 ImageNet competition?

The friendly competition became a high-profile sparring match, a Chinese internet company Baidu cheated by data snooping, they submitted with many different accounts to bypass the specific limit of 2 submissions per week.

9). Describe some commercial applications of convolutional neural networks.

Auto image blurring on personal information on Google's Street View, detect object and human emotions behind a picture, realtime facial recognition, pedestrian tracking, diagnose breast and skin cancer from medical images, and more.

10). Have ConvNets surpassed humans at object recognition?

Yes and No, on the ImageNet test the computer has an advantage at training on large datasets, thus allowing identification of very specific things that are not common knowledge. The test also accepts an answer as correct if 1 of the top 5 submitted was correct.

11). What is the relationship between “object recognition” and “visual intelligence?”

Object recognition is just describing things in an image, but visual intelligence is about telling a story from a scene, it requires understanding of many aspects of human lives and interactions, as well as experience of related situations.