Kuncheng Feng
CSC 416

Chapter 4: Who, What, When, Where, Why → Reading/Mining/Discussion Assignment

1). Why do you think that Melanie Mitchell titled this chapter “Who, What, When, Where, Why?” Please do your best to articulate what you believe she was trying to convey with this title.

These questions are what humans try to answer when we saw an image, and this chapter is about having machines try to do that.

2). TRUE or FALSE: We humans perform vast amounts of information processing in hardly any time at all, and we have very little, if any, conscious awareness of what we’re doing or how we do it.

True

3). What ability does Melanie Mitchell suggest would be one of the first things we would require for general human level intelligence?

The ability to describe the contents of a photograph or a video in a way that transforms it into a detailed story.

4). Describe the “object recognition” problem?

Recognizing a particular group of pixels in an image as a particular object category, such as “woman”, “dog”, “balloon” or “laptop computer”.

5). TRUE or FALSE: Object recognition is typically so immediate and effortless for us as humans that it didn’t seem as though it would be a particularly hard problem for computers, until AI researchers actually tried to get computers to do it.

True

6). TRUE or FALSE: Prior to the deep-learning revolution, the major job of computer-vision researchers was to develop specialized image-processing algorithms that would identify “invariant features” of objects that could be used in their recognition.

True

7). TRUE or FALSE: The ability of machines to recognize objects in images and videos underwent a quantum leap in the 2010s due to advances in the area called deep learning.

True

8). TRUE or FALSE: The “traditional” multilayer neural networks were inspired by the brain, but their structure is very un-brain-like. In contrast, the neural networks dominating deep learning are directly modeled after discoveries in neuroscience.

True

9). In just a few sentences, say something about David Hubel and Torsten Wiesel, and their contribution to the topic of this chapter.

They are two neuroscientists that radically remake our understanding of vision, and particularly object recognition in the brain. They discovered hierarchical organization in the visual systems of cats and primates and their explanation of how the visual system transforms light striking the retina into information about what is in the scene,

10). In just a few sentences, say something about Kunihiko Fukushima, and his contribution to the topic of this chapter.

A Japanese engineer who developed one of the earliest deep neural networks (cognitron and neocognitron). He reported some success training the neocognitron to recognize handwritten digits, but his methods were not extended to more complex visual tasks.

11). In just a few sentences, say something about Yann LeCun, and his contribution to the topic of this chapter.

A French computer scientist who had been inspired by Fukushima’s neocognitron, proposed ConvNets in the 1980s. (Convolutional neural networks).

12). TRUE or FALSE: Like the neocognitron, the design of ConvNets is based on several key insights about the brain’s visual system that were discovered by Hubel and Wiesel in the 1950s and 60s.

True

13). Layers in a traditional deep learning network consist of a list of simulated neurons (units). Not so with a convolutional neural network. Describe a layer in a convolutional neural network.

Each layer or unit provides input to units in the next layer.

14). What do you think is the most salient similarity between object recognition in the brain and in convolutional Neural networks?

They both have layered tasks to perform. The brain recognizes more and more complex structures and identifies the whole things, the convolutional neural networks passes value to the next layer and concludes with confidence.

15). Describe the “receptive field” of a simulated neuron (unit) in a convolutional neural network.

Small region around a location which units get its input.

16). How do you calculate the “convolution” associated with a simulated neuron (unit) in a convolutional neural network?

Multiplying each value in a receptive field by its corresponding weight and summing the results.

17). TRUE or FALSE: An “activation map” in a convolutional neural network is constructed by computing the convolution for each simulated neuron (unit) in the map with respect to some “magically determined” feature.
(For this question, think of magic as so many science fiction writers do, as “anything enabling actions beyond our current capability to understand them.”)

True

18). What analogy does Melanie Mitchell explore in the text by way of illustrating the ideas associated with “maps” in convolutional neural networks?

It’s like a very special camera that could produce separate photos for house lights, building lights and car lights.

19). TRUE or FALSE: A convolutional neural network, like the brain, represents the visual scene as a collection of maps, each reflecting the specific “interests” of a set of feature detectors.

True

20). TRUE or FALSE: Determining the number of “layers” in a ConvNet and the number of “maps” in a layer of a ConvNet is part of the art of getting these complex networks to work for a given task.

True

21). Melanie Mitchell recalls I. J. Good’s vision of a future “intelligence explosion” in which machines themselves create increasingly intelligent machines, and then proceeds to mention that with respect to convolutional neural networks we are not there yet. What do you think of the idea of using genetic algorithms as an “AI vehicle” by which to get us there? That is, what do you think of the possibility of employing a genetic algorithm to play the role of “ConvNet artist” in determining the architecture of a convolutional neural network to solve a particular problem?

I think the idea of using genetic algorithms to play the artist role is great in this problem. Since right now the structure is already figured out, it’s just trials and errors to figure out the details, and genetic algorithms are great at that.

22). Describe the “classification module” for a convolutional neural network.

An entire traditional neural network. The module takes activation maps from the highest convolutional layer, and outputs a set of percentage confidence values.

23). Describe the process of training a convolutional neural network.

Collect many images of things you want to identify and label them so, then have the training program run through it. The program will change its weight as it makes mistakes, and repeating this process will make the program getting better at it.

24). TRUE or FALSE: Even though convolutional neural networks are not constrained by a programmer to learn to detect any particular feature, when trained on large sets of real-world photographs, they indeed seem to learn a hierarchy of feature detectors similar to what Hubel and Wiesel found in the brain's visual system.

True

25). What concurrent technological revolution made possible the extraordinary ascent of convolutional neural networks from relative obscurity to near-complete dominance in machine vision?

"Big data"