Kieran Finnegan
MM Ch4 25 Questions

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.

ANS: She is trying to articulate what a picture of something is. How our brain can process the pixels or ink into an image that will be something and what is happening in the picture. For a computer it is just a series of pixels with no meaning other than them being in the right place. In order for a computer to demonstrate AI capabilities it would have to understand what the picture is and not just hold it to being a binary number.

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.

ANS: True

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

ANS: The Ability to describe a photograph should be one of the first things we would require for general human-level AI

4. Describe the “object recognition” problem?

ANS:: It recognizes a particular group of pixels in an image as a particular object category. Taking the example of the dog, the image can have many different scenarios where the dogs are of all different sizes, can be blocked by things like humans or fences, or the lighting of the picture could obscure a dog from some portions of the background. There are many factors that can make this difficult for a computer to recognize then when a human looks at it.

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.

ANS: 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.

ANS: 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.

ANS: 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.

ANS: True

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

ANS: Their discoveries led to a hierarchical organization in the visual systems of cats and primates. Their explanation of how 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.

ANS: In the 1970’s he developed one of the earliest deep neural networks, dubbed the cognitron, and its successor, the neocognitron. He reported some success training the neocognitron to recognize handwritten digits, but the specific learning methods he used did not seem to extend to more complex visual tasks. Though neocognitron was not perfect it was an important inspiration for later approaches to deep neural networks.

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

ANS: He proposed ConvNets that were inspired by Fukushimas neocognitrons. ConvNets use an array of numbers to detect how light values in a picture correspond to values of what an image is.

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.

ANS: 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.

ANS: Given an input image, each layer successively performs its calculations, and finally at layer 4 the network has produced a set of activation maps for relatively complex features.

14. What do you think is the most salient similarity between object rectognition in the brain and in convolutional neural networks?

ANS: They both use features of an image to determine what the image is representing to determine what is in the image. Though object recognition in the brain works differently from convolutional neural networks they still achieve the same outcome

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

ANS: Each unit in a map corresponds to the analogous location in the input image, and each unit gets its unput from a small region around that location

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

ANS: Each unit in each map calculates an activation value that measures the degree to which the region ‘matches’ the unit’s preferred edge orientation (vertical, horizontal or slanted at various degrees).

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.”)

ANS: True

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

ANS: She uses the dogs fur against the grass light values to determine whether a part of an image is used within the map or not based on color values.

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.

ANS: 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.

ANS: 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 algoroithms as an “AI vehicle” by which to get us there? That is, what do you thing 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?

ANS: It will take a lot of human ingenuity to get there. I feel if the algorithm is based on detection of a specific thing we are already there. To have it be able to detect everything will take at least a few more decades before we are even close.

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

ANS: Layers 1-4 are convolution layers because each performs convolutions on the preceding layer. Given an input image, each layer successively performs its calculations, and finally at layer 4 the network has produced a set of activation maps for relatively complex features. The inputs to the classification module are the activation maps from the highest

convolutional layer. The module's output is a set of percentage values, one for each possible category, rating the network's confidence that the input depicts an image of that category.

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

ANS: First you need to collect many examples images of what is to be a 'training set.' The training programme initially sets al all the weights in the network to random values. Then the programme commences training on each image. The training programme compares the output values to the 'correct values.' Then the training programme uses the back-propagation algorithm to change the weights throughout the network just a bit, so that the net time this image is seen, the confidences will be closer to the correct values.

24. TRUE or FALSE: Even though convolutional neral 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 int the brain's visual system.

ANS: 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?

ANS: Google "search by image"