Derek Lee Deep Learning Fall 2020

Professor Curro Quiz #10

1. What do you think is preventing main stream usage of these types of architectures?

These models are built to address specific problems that are prevalent in neural networks. However, these architectures are specifically designed to address those problems, so they cannot be easily incorporated into an existing network. These architectures are not generally applicable. In addition, they are also not revolutionary enough to justify using these architectures over other, well-established architectures. It appears that these architectures were developed as a proof-of-concept, just to show that neural networks can accomplish these sorts of tasks. These architectures are not well-suited to accomplish these tasks, because these tasks are usually accomplished with low-level operations. Even assuming that these architectures can achieve perfect accuracy, they would still be computationally inefficient.

Neural Arithmetic Logic Units generalize well when dealing with numerical information. They can learn addition, subtraction, multiplication, division, and power functions. This offers a solution to one of the problems with neural networks, which is their failure to compute basic arithmetic operations. However, this does not have much practical use. There are niche cases where this could be used, such as counting objects in images. However, the current solution to this problem is already sufficient. It is a very simple solution, which involves explicitly counting the images in a post-processing step. It does not require retraining the network and there is no risk of miscounting (excluding the possibility of a bug in the code). On the other hand, using a neural network to count introduces the possibility that the neural network will make an erroneous prediction. In addition, this would be inefficient since the network will both require training and is less efficient than the hardware (such as ALUs).

Neural Turing Machines (NTMs) and Differentiable Neural Computers (DNCs) have seemingly impressive results, but they have a similar issue as NALUs. It is interesting that read/write operations can be differentiable, and this opens new possibilities for neural networks, but the current applications are limited. These architectures merely simulate a CPU, which is less efficient than utilizing the actual CPU. It is a promising concept for a neural network to be able

to ‘learn’ to retain information through reading and writing, but the complexity of this architecture, in addition to the fact that it would be difficult to integrate into an existing architecture, makes it an undesirable option.