ImageNet Classification with Deep Convolutional Neural Networks - Analysis

This paper covers the development and performance of a deep convolutional neural network that can classify millions of images and describe the main object that is in each picture. These objects can vary greatly, from mites, container ships, dalmatians, mushrooms, cherries, car grills, and motor scooters. This is an impressive variety of items that a single network can recognize.

First off, it's incredible to see that the field of artificial intelligence and neural networks is recent and expanding faster than many other fields in computer science. I liked that they were able to classify a massive number of images that are extremely varied. In comparison, the ANN that we trained only had 10 output nodes and the complexity was much less. Whereas, this neural network has 6 million parameters and could identify object that many people would mess up. I didn't like that, at that point in the past, the best error percentage was 37.5% and 17.0%. This seems like a large percentage to be happy about. Without the mention that this percentage error was much better than the previous state-of-the-art network, I wouldn't be impressed. I also liked the diagram that showcases the structure of the network from a high abstract level. Putting each feature layer in a box and defining the shape of those arrays helps me to better understand the level of complexity and creativity that can be used to solve increasingly complex problems.

Something inspirational from this paper that I found was that depth is important. In fact, they mentioned that even though they had many convolutional layers, removing even one would reduce the accuracy by more than 2%. One would think that retraining the network would be able to make up for this, but indeed, the structure of a network can make or break the performance. This may lend a perspective into the amount of complexity in out own brains makes up for the intelligence that we possess. Another idea that I had during this was how networks can communicate to other portions of the network while processing. Separating the network, they were able to run two different GPUs for upper and lower portions of the network. These portions communicated only at certain parts of the process. I would love to take this idea further and connect parts of the network that perhaps don't relate and connect them to see if more "intelligence" can be gained from combining unrelated things at opportune times, much like the human brain has unrelated portions of the brain neighboring other parts.

Something I still don't like nor understand is how increasingly networks take such a large variety of parameters. At what point can the network increase in size until the network can classify all objects of define all things. Is this possible and maybe this is the thought of a general intelligence network that is able to define many objects. At this point, the cost of training, or even defining training data would be immeasurable.