# Self-driving cars and their ability to handle new situations

Anders Hansson

June 5, 2018

**In recent years the development of self driving cars has seen some great progress. More and more companies and organizations have vehicles out in traffic that partly steer themselves. However, the process of reaching a level where a car can drive satisfactory by its own is a long one and involves a lot of training. Until today, the most common approach is to expose the system to as many different situations as possible as it may encounter in reality. All in order to make the car as versatile as possible. The mathematical system responsible for the steering is called an artificial neural network. A much wanted feature in such a system is the ability to generalize and use prior knowledge in new situations. In the same way as humans can tell that a golden retriever is a dog, despite never seeing a golden retriever before. In a recent master's thesis from Lund University, promising results are presented for artificial neural networks for self-steering cars. The results show that certain features of a network may increase its ability to generalize.**

The main goal of the project was to design a system that could steer itself around a simulated race track. In the training process it would not be allowed to see images similar to the ones in the race track, but only images from real-life driving. This is kind of like teaching a toddler how a giraffe looks like only by looking at photographs of giraffes. Followed by showing a cartoon of a giraffe and see if the kid still can identify it as a giraffe. The results presented in the thesis indicate that an artificial neural network trained for two tasks simultaneously may improve its performance on the simulated race track. The method to have more than one task is referred to as multi-task learning, which is a subcategory of machine learning. In the project the network was not only trained to steer a car, but also to reproduce the images fed to it by first compressing the size of them. So the network was trained to, given an input image corresponding to a wind-shield view of a car, reproduce the input image and output a steering angle. This was done by first reducing the input image in size and then based on the compressed image reproduce the full sized input image along with outputting a steering angle. Note that the reproduction of the input image was just a side task and had nothing to do with steering the car. Yet, the results suggest that the addition of the side task improved the performance of the steering.

A system that can be taught to drive in all kind of conditions but trained only on a few has a potential to be very useful. The specific case in the thesis, where the car ultimately gets good at driving around a simulated race track, is perhaps not the most interesting. However, if the opposite could be achieved that would be of great usefulness. In other words, a system that could be fed with artificial images created by a simulator and then without further training be capable of driving in traffic. Artificially created images (imagine the screen output in a race car computer game) are inexpensive to produce compared to collecting data by driving around a car. Hence, such a system would save a lot of time and resources in the development process. In theory, basically the same features of a neural network that works in one direction should also work in the other. Which means that the prominent design details presented in the thesis to work well can be useful in the opposite situation described above.

Knowledge about how systems for self driving cars can be better at generalizing can avail other areas within artificial intelligence. Ultimately, it's a piece in the puzzle to understand how the human brain works on a mathematical level. The human brain which is incredibly good at generalizing. Hopefully the results can incite further research about how multi-task learning can help generalization.