Capsule Networks applied to White Blood Cell Image Analysis

Andreas Forslöw[1]

[1]Lund University, andreas.forslow@gmail.com

August 22, 2018

Deep learning is a field which recently has seen enormous improvements in areas such as Image Classification, Robotics and Speech Recognition. One participant in this field is CellaVision, which applies a deep learning technique called Convolutional Neural Networks for $80-90\%$ accuracy on Image Classification of blood cells; helping research labs with automating the diagnosis process. Naturally, CellaVision is interested in finding better classification algorithms, which is why this project has researched the recent paradigm of *Capsule Networks*. The investigations show that these networks learn spatial concepts better and outperform Convolutional Neural Networks on small data sets, but otherwise perform equally well to - or worse than - Convolutional Neural Networks.

The task of image classification is to create an algorithm which "learns" how to correctly classify never before seen images. A simple example of an image classification task is to classify between images of cars, blood cells and dogs, given labeled images of what these look like. This is indeed a difficult task, as there are no clear definitions of how these look, and an image often contains a vast amount of pixels; creating a huge amount of data that has to be analysed. The way Deep Learning solves the task is inspired by the brain, as its models use structures similar to synapses and neurons, and iteratively "learns" how to convert images to correct classifications. An example Neural Network is shown in Figure 1 below. It converts an image to pseudo-probabilities of the image containing a car, a blood cell or a dog.

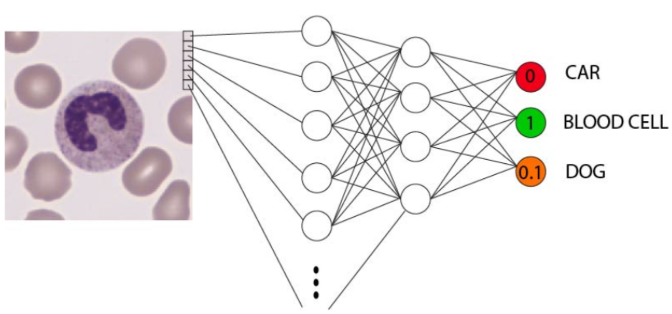

Figure 1: Neural network that classifies between cars, blood cells and dogs by assigning a probability number of it having observed each object. Here, the network thinks it observes a blood cell.

There are, of course, many different models used in Deep Learning, but this project deals with Convolutional Neural Networks, Dynamic Routing Capsule Networks and EM Routing Capsule Networks. Convolutional Neural Networks are used in CellaVision's systems today, and work by sliding rectangular filters across the images; first making neurons learn how patches of the images look, and then making consecutive neurons learn how these look in relation to each other, before the last layer of neurons converts this information to a classification.

Dynamic Routing Capsule Networks, meanwhile, increase the sophistication of the Convolutional Neural Networks in three ways. Firstly, they expand the 1-dimensional neuron to an n-dimensional vector called a capsule. Secondly, they introduce an election system, where the capsules of one layer have to agree on what the capsules in the consecutive layer should look like, before the last layer of capsules converts this information to a classification. Lastly, they learn how the image should be reconstructed based on the information that is sent to classification. The dimensional expansion allows for more information to be stored in each capsule, the election system creates more sophisticated knowledge in the model, and the image reconstruction to some extent makes the model learn how its analyses relate to the complete image.

EM Routing Capsule Networks increase the sophistication of the Dynamic Routing Capsule Network even further, as it also introduces computer graphical concepts into the model. By turning the n-dimensional capsules into 4×4 pose matrices, the network is able to learn spatial concepts better than both the Convolutional Neural Network and the Dynamic Routing Capsule Network, and to some extent perform inverse graphics computations, before converting this information to a classification.

The tests run on CellaVision's data sets of White Blood Cell images indicate that EM Routing Capsule Networks indeed do perform some sort of inverse graphics computations, as these are better at generalising spatial concepts. The Dynamic Routing Capsule Networks, meanwhile, show performance improvements when the data sets are small; indicating that they indeed are able to gain a more sophisticated knowledge of the images. However, the Convolutional Neural Networks outperform the other models when the data sets are big and complex - both in terms of accuracy and execution times. To conclude, Convolutional Neural Networks should therefore see their continued use in CellaVision's systems until the Capsule Networks are improved.