Review on “Dynamic Routing Between Capsules”

Authors: Geoffrey E. Hinton, Sara Sabour, Nicholas Frost

# Short Summary

In this paper, the authors propose a dynamic routing algorithm for learning in a Capsule network. A Capsule is a group of a neurons whose activities represent properties of an object; that is, it is a vectorized alternative to traditional deep learning networks that encode both existence and parameters of an entity. By considering the relationship between capsules, such a network theoretically addresses the limitations of a Convolutional Neural Network. The proposed routing algorithm describes how inputs to a capsule are transformed into an output: the weighted sum of all prediction vectors in the previous capsule layer are squashed into vector lengths in the range [0, 1] describing the probability that the entity represented in the current capsule is present in the input. The key to this process is weighting the prediction vectors through coupling coefficients, which sum to 1 for each capsule, that describe the agreement between capsules in subsequent layers. This “routing-by-agreement” is expected to be more efficient than max-pooling between convolutional layers which loses valuable information between layers. To prove its effectiveness, a simple Capsule Network (CapsNet) architecture is proposed in this paper and compared to a baseline for classification on the MNIST dataset. Despite its infancy, the CapsNet architecture is competitive (0.25% test error) and shown to be more robust to affine transformations than the baseline (79% vs 66% test accuracy). Additionally, the inclusion of a decoder structure at the end of the network proves to be an effective form of regularization and allows for reconstruction of an input image using the activity vector. Finally, the CapsNet is applied to the task of image segmentation with the MultiMNIST dataset achieving similar classification performance as a baseline but on a comparatively more difficult task (segmenting digits with 80% overlap opposed to <4%). Tests on other well-known datasets such as CIFAR-10, smallNORB and SVHN are also briefly described and demonstrate that capsule networks are competitive with listed baselines.

# Main Contributions

* Proposed the dynamic routing algorithm for learning with Capsule nets
* Implemented a capsule network with the proposed routing algorithm and evaluated performance on various datasets
* Demonstrated that Capsule networks are more robust to affine transformations than an ordinary CNN and perform considerably better on the task of segmenting overlapping objects

# High-Level Evaluation of Paper

The algorithm proposed in this paper is fundamental leap forward in terms of efficient implementation of a capsule network. Drawing inspiration from the way humans process information, the CapsNet architecture implemented in the paper seeks to address the shortcomings of existing state-of-the-art deep learning techniques. The pseudocode discussed within the paper provides a concrete overview of the underlying theory and mathematics surrounding capsule networks. This, in conjunction with the example describing a Capsnet implementation, are tremendous aids for the reader. The use of the decoder structure as both a form of regularization and a means to visualize the activities in the capsules is also very impressive. In particular, perturbation of the image reconstruction is a testament to much a capsule network is able to learn in terms of digit variation. With all these advantages, it is surprising that the paper does not delve into the weaknesses or limitations associated with a capsule network other than its relative infancy. Furthermore, the models used as baselines appear to be a few years old. Additionally, it seems that the focus of this paper was to briefly re-introduce capsules, describe the routing algorithm and then delve into experiments. To tie everything together, a practical example detailing what the capsules represent in an image and how routing develops the relationship between layers would have been helpful.

# Discussion on Evaluation Methodology

In terms of evaluation, the paper picks simple CNN architectures as baselines of performance. The intention, as stated in the paper, is to maintain a certain degree of similarity between the proposed Capsule network architecture and CNN architecture. That said, the baselines are far from state of the art and in some cases dated by quite a few years. I believe this ties in into the lack of discussion on the weaknesses of Capsule nets suggesting that the models are not easily scalable or efficient to train. For example, in the discussion on the CIFAR-10 dataset, the authors mention a possible explanation for the lacking performance is that the backgrounds are too varied given the size of the CapsNet. This begs the question, why wasn’t a larger network used in the evaluation instead?

# Possible Directions for Future Work

As suggested in the paper, future work should aim to gain further insight into capsule networks and develop the technology. One such direction could be to investigate the routing iterations and issue with overfitting further.