

Aeden Jameson
 9810912
 EE 596 Autumn 2021
 Homework 1

(1) Motivation – Why do the authors want to work on this problem

Given the authors knowledge of how constraints enhance the ability of a model to generalize. They would like to demonstrate how constraints can be integrated into the backpropagation algorithm. That is instead of reducing the number of free parameters allow them to be adaptively trained through the back propagation algorithm

(2) Contributions – What are the accomplishments they achieved in this paper (others did not achieve)?

Unlike previous attempts at recognizing handwritten digits taken from US mail the authors made several novel contributions,

- The images are fed directly to the network rather than feature vectors.
- The connections of the first few layers are trained with backpropagation rather than hand-chosen constants implemented on a neural network chip

(3) Formulations – How do they solve the problems as mentioned/discussed in the introduction or related literature?

The authors achieve similar capabilities of previous neural networks through the use of constraining back propagation by using weight sharing. This technique involves several connections sharing the same weight.

(4) Justification – Do the experiments/simulations support their claimed accomplishments?

I don't have access to their work, but the description of their setup would seem to make their results plausible.

- They cleaned and scaled their data.
- They used a training and test set with a well known loss function
- From what I know of neural networks the diagram looks plausible

(5) Your Own Thoughts – What are you most impressed with in this paper?

One of the most impressive and fascinating things was the degree to which they specified the network, choose the activation function in order to prevent the weights from growing indefinitely to achieve their results. The authors go on to mention that “Some kernels synthesized by the network can be interpreted as feature detectors remarkably similar to those found to exist in biological systems.” For me this example and others raises a question about, what is the nature of the relationship between the structure of the networks and what they’re learning, and the visual world?