Assignment 3, Literature draft

Philip Bouman (10668667)

Problem definition

Modelling fonts with convolutional neural networks

The goal of the project is to find relations between characters of different types of writing systems (scripts). In the first part of the research a deep convolutional neural network will be trained to classify (images of) characters. A deep convolutional neural network is a machine learning architecture consisting of (multiple) stacks of layers. Depending on the availability of (training)data the training/learning process can either be supervised, using phonemes to label characters or unsupervised, using just the characters. The second part of the project will focus on getting a better understanding of the representations learned by the network. This will mainly be done by investigating and visualizing features learned in individual layers and smaller combinations of layers. The results can be compared to existing image recognition networks or between methods (for example if supervised and unsupervised are both used).
(139 woorden)

Literature

ImageNet Classification with Deep Convolutional Neural Networks[4]

This article is cited a lot in image recognition research. It describes a method that achieved high results in classifying images for the ImageNet LSVRC-2010 contest. The used method achieved much better results then previous state of the art methods. A deep convolutional neural network (CNN) is used to classify a large dataset of images into 1000 different classes using supervised training. The article describes the architecture and the most important layers of the network aswell as some added alternative methods, in contrast to traditional CNNs. The task in this article is similar to my project, because it is an image classification task aswell and it is using the same machine learning method (CNN).

Convolutional neural network committees for handwritten classification[2]

Another article about CNNs, but the task in this paper is more closely related to my project. The classification task is for handwritten characters, in my project printed characters will be used. The description of the CNN is in less detail then the previous article, but this is mainly because the network is smaller and less complex. It can provide some insights about recognizing and classifying characters specifically. The described network is able to classify different scripts aswell (e.g. Chinese characters).

Feature extraction with convolutional neural networks for handwritten word recognition[1]

The method used in this article is a combination of a CNN and Hidden Markov Model (HMM) to recognize handwritten words. The article describes the process of feature extraction in detail. The HMM is used to segment words into characters or groups of characters, which are then put through the CNN. It uses partly the same method and similar data that I propose for my project and it provides more insights about feature extraction.

Human-level concept learning through probabilistic program induction[5]

This article is related to the second part of the project, the understanding of the learned representations by the CNN. The researchers compare performance of people and different computer models in five concept learning tasks. These tasks are based on handwritten characters from different scripts. For instance generating new examples from a given example. The focus of the research is to reduce the amount of data used for the model to learn, to approximate human-level concept learning capabilities. Although the used method is different (Bayesian based models) the article contains alot of information about character representations by the system. For instance the different parts and subparts a character is comprised of in the model.

A self-organized artificial neural network architecture for sensory integration with applications to letter-phoneme integration[3]

A neural network designed with structures closely related to neurophysiological findings is described in this article. Parts mimicking brain functions are integrated in the architecture, focussing on sensory integration. The model is used to explore/model letter-phoneme relations.
Although I haven't read the complete article (it's 40 pages), so far it seems usefull for both parts of my project. The structure and achitecture of the neural network is documented aswell as several methods used to achieve a more brain like structure and better performance. I'm not totally sure what to expect or find

in the second part of the project, but this paper contains alot of information about letter-phoneme relations in a model (neural network) aswell as in the brain.

References

[1] Théodore Bluche, Hermann Ney, and Christopher Kermorvant. Feature extraction with convolutional neural networks for handwritten word recognition. In *Document Analysis and Recognition (ICDAR), 2013 12th International Conference on*, pages 285–289. IEEE, 2013.

[2] Dan Claudiu Ciresan, Ueli Meier, Luca Maria Gambardella, and Jurgen Schmidhuber. Convolutional neural network committees for handwritten character classification. In *Document Analysis and Recognition (ICDAR), 2011 International Conference on*, pages 1135–1139. IEEE, 2011.

[3] Tamas Jantvik, Lennart Gustafsson, and Andrew P Papliński. A self-organized artificial neural network architecture for sensory integration with applications to letter-phoneme integration. *Neural computation*, 23(8):2101–2139, 2011.

[4] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[5] Brenden M Lake, Ruslan Salakhutdinov, and Joshua B Tenenbaum. Human-level concept learning through probabilistic program induction. *Science*, 350(6266):1332–1338, 2015.