

# Research Seminar in Data Science

## Paper Reviews

Christoph Schmidl  
s4226887  
c.schmidl@student.ru.nl

June 10, 2018

### Reviewed papers:

- **Human-level control through deep reinforcement learning** by Volodymyr Mnih et al. (Presented by Valentin Koch)
- **Deep Residual Learning for Image Recognition** by Kaiming He et al. (Presented by Stephan Dooper)

## 1 Human-level control through deep reinforcement learning

### 1.1 Summary

In their paper "Human-level control through deep reinforcement learning", Volodymyr Mnih et al combine a deep neural network approach with reinforcement learning called "Deep Q-Network" (DQN) that can learn successful policies directly from high-dimensional sensory inputs. The proposed method was tested on classic Atari 2600 games receiving only pixels and the game score as inputs. By maintaining the same algorithm, architecture and hyperparameters amongst 49 different games, their network was able to beat all previous algorithms and achieve human level performance. Drawbacks of known reinforcement learning techniques have been mitigated by using experience replay and an iterative update procedure that adjusts the action-values towards target values that are only periodically updated. Although their DQN approach is showing good performance on 43 out of the 49 games with any prior knowledge, there are certain games which seem difficult for the proposed algorithm because the chain of actions without any reward seems to be too long to be learned without any prior knowledge.

### 1.2 Evidence

The authors cite 30 different research papers and describe techniques which were used before. The reference list also contains classical papers like "Shape and arrangement of columns in cat's striate cortex" by Hubel and Wiesel from 1963. Well known authors like Hinton, LeCun Sutskever and Krizhevsky can also be found. The amount of references and their topics seem to be well picked and forms a strong foundation.

### 1.3 Strengths

The main strenghts of this paper is its novelty of presenting one algorithm, one architecture and one set of hyperparameters in conjunction of two different machine learning techniques in order to learn action-reward based games. Given the fact that the algorithm is able to reach human level performance is also a significant strenghts.

### 1.4 Weaknesses

The main weakness of the paper is the comparison with "human level performance". Human performance can have a very large variance and it is difficult to say when a human player can be considered a "professional gamer" in a certain domain.

## 1.5 Evaluation

I guess this paper would get accepted to a conference or journal because it is written in a clear way and also contains supplementary material if a reader demands more background knowledge and more technical details. The main reason to accept this paper is its novelty and the main idea behind the DQN architecture.

## 1.6 Comments on the quality of the writing

The writing style was formal but also written in a clear way.

## 1.7 Queries for discussion

- How does this algorithm compare to the approach by Tuomas Sandholm and Noam Brown known for "Libratus: The Superman AI for No-Limit Poker"? I would assume that Libratus can also be applied to Atari games because it is action-reward based.

# 2 Deep Residual Learning for Image Recognition

## 2.1 Summary

Kaiming He et al address the degradation problem by introducing a deep residual learning framework. The degradation problem occurs with increased network depth. When the network starts converging the accuracy gets saturated and then degrades rapidly. In contrast to a plain network architectures like VGG16, a deep residual network is able to maintain a relatively high accuracy even when the number of layers varies from 100 to 1000. Kaiming He et al hypothesize that it is easier to optimize the residual mapping than to optimize the original, unreferenced mapping. Residual learning is realized by so called "shortcut connections". They evaluate their results based on experiments with the imagenet and cifar-10 datasets which are about image classification and the PASCAL and MS COCO dataset which are about object detection and object segmentation. Residual networks seem to yield better performance on all datasets when the number of layers start increasing into the hundreds.

## 2.2 Evidence

The authors cite 49 different research papers to back up their paper. Among the different citations are well known researchers like C.M. Bishop, I.J. Goodfellow, G.E. Hinton, Y. LeCun and J. Schmidhuber. Out of these 49 papers, only 3 are also published by K. He himself. The papers seem to be well picked and show the problems of training networks with increased depths and already known solutions on how to tackle these problems. Based on the fact that none of the cited papers includes the name "residual", I assume that this paper is the first one which came up with residual network structures.

## 2.3 Strengths

The authors explain in detail what the main problem is when it comes to training very deep neural networks. The degradation effect seems to be of main interest and the residual network structure seems to be a novel approach. In the section about "shortcut connections" the authors also mention that there seems to be a similar approach to their shortcut connection implementation which has been developed during the same time span by Schmidhuber. Schmidhuber's idea is called "highway networks". The authors state out that the main differences between their idea and the one proposed by Schmidhuber. I guess that was done because Schmidhuber is known for claiming credit for ideas which seem similar to his own but differ slightly and is known for confronting researcher publicly during conferences to explain themselves.

## 2.4 Weaknesses

I could not find any real discussion, conclusion or future work section. This made it kind of frustrating for me to read because I wanted to start to follow the main three-step approach of reading research papers.

## 2.5 Evaluation

I assume that this paper was accepted for a conference or journal because it offers a novel idea which seems promising. Maintaining the same amount of hyperparameters as so called "plain networks" but being able to make the network way deeper at the same time and keeping a high accuracy seems to be a very tempting idea to try out. Given the fact that the authors also won different competitions with their approach is another

reason why this paper is interesting. Well known datasets like imagenet, cifar-10 and coco have been chosen to validate the approach which makes it easy to reproduce.

## **2.6 Comments on the quality of the writing**

The technical detail seems to be sound and it is well written when you overlook the fact that the discussion, conclusion and future work section is missing.

## **2.7 Queries for discussion**

- The authors state that they did not use maxout/dropout for their networks with over 1000 layers although they assume that this may improve their results. Has there been further work on this assumption?