



DEPARTMENT OF INFORMATICS

TECHNISCHE UNIVERSITÄT MÜNCHEN

Master's Thesis in Robotics, Cognition, Intelligence

Evaluation and Generalization of Capsule Networks in Neurorobotics

Jean A. Elsner





DEPARTMENT OF INFORMATICS

TECHNISCHE UNIVERSITÄT MÜNCHEN

Master's Thesis in Robotics, Cognition, Intelligence

Evaluation and Generalization of Capsule Networks in Neurorobotics

Evaluation und Generalisierung von Kapsel Netzwerken in der Neurorobotik

Author:	Jean A. Elsner
Supervisor:	Prof. Dr. Alois C. Knoll
Advisor:	Alexander Kuhn
Submission Date:	TBD



I confirm that this master's thesis in robotics, cognition, intelligence is my own work and I have documented all sources and material used.

Munich, TBD

Jean A. Elsner

Acknowledgments

Abstract

Contents

Acknowledgments	iii
Abstract	iv
1 Introduction	1
2 State of the Art	2
2.1 Deep Learning for Object Recognition	2
2.1.1 Convolutional Neural Networks	2
2.1.2 Spiking Neural Networks	2
2.2 Limits of Deep Learning Approaches	2
3 Capsule Network Architectures	3
4 Experimental Setup	4
5 Results	5
6 Discussion	6
7 Conclusion	7
List of Figures	8
List of Tables	9
Bibliography	10

1 Introduction

2 State of the Art

This chapter presents an overview of state of the art approaches to object recognition, while focusing on two families of architectures, which are motivated quite differently. Object recognition techniques based on convolutional neural networks (CNNs) currently dominate the field, achieving state of the art performance on many datasets [1, 2]. CNNs however, are only loosely based on biological neurons. Spiking neural networks (SNNs) on the other hand, try to mimic the physical properties of neurons more closely and therefore constitute biologically more plausible models [3]. Generally speaking, CNNs may be regarded as a more engineering-based approach (or top-down), while SNNs are motivated by results from neuroscience and biology (bottom-up approach).

2.1 Deep Learning for Object Recognition

Recent years have seen a surge of interest in deep learning methods, especially in the field of computer vision. While the theory behind many deep learning methods has been around for many years, their recent success is mainly due to the availability of large labelled data sets and highly parallel computing powered by GPUs. One of the specific tasks, deep learning based methods excel at, is object recognition. The objective of object recognition is to identify objects in images or videos. Potential applications of a robust image classification system reach from automated driving and image-based diagnosis to robot vision and many more.

2.1.1 Convolutional Neural Networks

2.1.2 Spiking Neural Networks

2.2 Limits of Deep Learning Approaches

3 Capsule Network Architectures

4 Experimental Setup

5 Results

6 Discussion

7 Conclusion

List of Figures

List of Tables

Bibliography

- [1] A. Diba, V. Sharma, A. M. Pazandeh, H. Pirsiavash, and L. V. Gool. “Weakly Supervised Cascaded Convolutional Networks.” In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2017), pp. 5131–5139.
- [2] W. Ouyang, X. Zeng, X. Wang, S. Qiu, P. Luo, Y. Tian, H. Li, S. Yang, Z. Wang, H. Li, K. Wang, J. Yan, C. C. Loy, and X. Tang. “DeepID-Net: Object Detection with Deformable Part Based Convolutional Neural Networks.” In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39.7 (July 2017), pp. 1320–1334. issn: 0162-8828. doi: 10.1109/TPAMI.2016.2587642.
- [3] A. J. Schofield, I. D. Gilchrist, M. Bloj, A. Leonardis, and N. Bellotto. “Understanding images in biological and computer vision.” In: *Interface Focus* 8.4 (2018). issn: 2042-8898. doi: 10.1098/rsfs.2018.0027. eprint: <http://rsfs.royalsocietypublishing.org/content/8/4/20180027.full.pdf>.