Strategies for Practical Deep Learning

Doctoral Dissertation submitted to the
Faculty of Informatics of the Università della Svizzera Italiana
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

presented by
Lukas Tuggener

under the supervision of
Jürgen Schmidhuber and Thilo Stadelmann

11 2024

Dissertation Committee

Luca Maria Gambardella Università della Svizzera Italiana, Switzerland
Rolf Krause Università della Svizzera Italiana, Switzerland
Benjamin F. Grewe Eidgenössische Technische Hochschule Zürich, Switzerland
Martin Jaggi École Polytechnique Fédérale de Lausanne, Switzerland

Dissertation accepted on 29 11 2024

Signiert von:
Jürgen Schmidhuber
E68CA477620F49F...

Research Advisor

Jürgen Schmidhuber

Signiert von:
Thilo Stadelmann
414F744C42464AF...

Co-Advisor

Thilo Stadelmann

PhD Program Director

The PhD program Director Walter Binder

I certify that except where due acknowledgement has been given, the work presented in this thesis is that of the author alone; the work has not been submitted previously, in whole or in part, to qualify for any other academic award; and the content of the thesis is the result of work which has been carried out since the official commencement date of the approved research program.

Signed by:
L. Tuggener
331ECD236964491...

Lukas Tuggener
Lugano, 29 11 2024

Abstract

In recent years, deep neural networks have achieved breakthrough results in diverse domains, from computer vision and natural language processing to game playing and life sciences. However, harnessing the full power of this technology in practical applications remains challenging. In this thesis, we explore strategies to address the challenges of applying deep learning to real-world pattern recognition problems. We tackle multiple practical problems, such as Optical Music Recognition (OMR), automated machine learning (AutoML), or the design of robust neural network architectures. In the context of OMR, we introduce two datasets, DeepScores and DeepScoresV2, the largest and most complete OMR datasets to date. Based on this data, we develop the first object detection method capable of handling the challenges of written music and methods to harden neural networks against the effects of degraded real-world data more than doubling detection performance on messy, degraded data. We then investigate the current state of AutoML, introduce a novel method for AutoML and extract design patterns for resource-constrained AutoML settings. In the latter parts of this thesis, we focus on the underlying issues that often cause neural networks to generalize poorly to real-world data. We first investigate the dataset dependency of modern CNN architectures. We show through an extensive empirical study that ImageNet alone is not sufficient to judge the power of CNN architectures and propose strategies for developing more universal evaluation methods. Finally, we tackle the lack of rotation invariance in modern vision systems and introduce a novel bio-inspired paradigm that significantly enhances the rotational robustness and outperforms the current state of the art by 19%.

Acknowledgements

I want to express my deepest gratitude to my supervisors Professors Jürgen Schmidhuber and Thilo Stadelmann, for guiding me throughout my academic journey and giving me the freedom to find my path while challenging me to become the best researcher I can be. Under their leadership, I was able to grow tremendously not only as a researcher but also as a person.

To my colleagues at CAI and InIT, I am very thankful for all the collaborations, discussions and the positive and stimulating environment they have created. Special thanks to Mohammadreza Amirian and Ismail Elezi for treading the treacherous waters of early PhD life with me.

Special thanks to Larissa for her love, patience, and support. She has been a tremendous source of strength for me. I thank her for standing by my side in all my hardest moments and celebrating the good ones with me. I am eternally grateful to have her by my side.

To my parents and sisters, I am thankful for their unconditional support. They provided me with the stable foundation upon which I was able to build and to this day give me a deep sense of calm and safety.

My dear friends Dominic, Pascal and Thomas I thank for their companionship, for all the laughs and heartfelt discussions we had throughout the years. They helped me get my mind off things even in the most stressful times.

Lastly, I want to thank all the healthcare workers who took care of me during the time when I could not.

Contents