UKRAINIAN CATHOLIC UNIVERSITY

MASTER THESIS

Neural architecture search: a probabilistic approach

Author:
Volodymyr LUT

Supervisor:
Yuriy KHOMA

A thesis submitted in fulfillment of the requirements for the degree of Master of Science

in the

Department of Computer Sciences
Faculty of Applied Sciences

Lviv 2020

Declaration of Authorship

I, Volodymyr LUT, declare that this thesis titled, "Neural architecture search: a probabilistic approach" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

"It's inspiring to see how AI is starting to bear fruit that people can actually taste. There is still a long way to go before we are truly an AI-first world, but the more we can work to democratize access to the technology—both in terms of the tools people can use and the way we apply it—the sooner everyone will benefit."

Sundar Pichai, CEO Alphabet Inc., May 17, 2017

UKRAINIAN CATHOLIC UNIVERSITY

Faculty of Applied Sciences

Master of Science

Neural architecture search: a probabilistic approach

by Volodymyr LUT

Abstract

In this paper we review different approaches to use probabilistic methods in existing AutoML solutions using Reinforcement Learning. We focus on providing additional knowledge about probability distribution provided to Reinforcement Learning agents solving Neural Architecture Search tasks. Based on the results of the research we come with an agent designed to model Neural Architectures for image classification tasks.

Acknowledgements

First of all, I would like to thank my supervisor Mr. Yuriy Khoma and Mr. Vasily Ganishev for all their support, patience and knowledge shared with me during this project.

I would also like to thank all my colleagues in UCU - those who already graduated, who were working on their masters during this year, and those who left UCU - for being a bright lighthouse, which navigated me towards excellence, for showing a good example and for being a strong and supportive community.

Many thanks to the UCU team - to all those people, who build and design the educational system I am proud being a part of, to all those who inspired me, shared their vision and made my time being a masters program student an unforgettable experience. I thank everyone who works every day in the UCU - no matter whether they are responsible for brilliant coffee or brilliant courses and workshops.

Special thanks goes to my friend and business partner Serhii Chepkyi for providing a big support during my studies and also for providing a big help with visual materials for this project.

Last but definitely not least - I would like to thank my closest people and my family - without your eternal love this would not be possible. This was the hardest years in my life and I'm glad and thankful that every second I've felt that I'm not alone.

Contents

List of Figures

List of Tables

List of Abbreviations

LAH	**L**ist **A**bbreviations **H**ere
WSF	**W**hat (it) **S**tands **F**or

Physical Constants

Speed of Light $\quad c_0 = 2.997\,924\,58 \times 10^8\,\mathrm{m\,s^{-1}}$ (exact)

List of Symbols

a	distance	m
P	power	W ($\mathrm{J\,s^{-1}}$)
ω	angular frequency	rad

For all the brave people who make it possible for millions of young Ukrainians to hold books in their hands instead of rifles and grenades.

Chapter 1

Introduction

AutoML (Automated Machine Learning) is becoming a new challenge for the industry, aiming to reshape the way business and individuals would use machine learning in their everyday life. This process is an automation of routine work of ML engineers and contains of problems industry regularly solves, including but not limited to data preparation, feature engineering, feature extraction, neural architecture search, hyperparameters selection etc.

As machine learning provides endless variety of automation possibilities for industries the problem of automation of ML industry itself seems natural. However, it is much trickier because the domain itself is developing rapidly. For decades ML engineers was pioneers in the new era of computer science research. As a result a new industry was shaped and this industry requires automation.

We believe that world would benefit from democratization of this new tools. ML is becoming a part of everyday routine for thousands of developers worldwide providing with ability to run models even on portable devices, IoT chips and other mass market hardware. AutoML is a big move towards in terms of variety of different applications created, tools used and research possibilities.

AutoML become a natural product for almost all big technological companies. Google, Amazon, Salesforce and others are offering AutoML products that allows non-experts to create their own ML solutions.

Still, existing AutoML techniques requires lots of computational resources and most of the research in the field is covered by tech giants nowadays.

We think that a task of neural architecture search should be optimized. We are using probabilistic approaches to show that they are able to handle an exploration and exploitation problem efficiently for reinforcement learning agents which are solving neural architecture search problems.

We are focusing on neural architecture search problem, especially on hyperparameter optimization task because historically this problem is solved mainly using exhaustive search techniques, such as grid search. Engineers often follows their empirical knowledge and try to guess optimal parameters to tune models. It is hard to optimize and automate something without knowing the nature behind it - so we are working on more effective approach than exhaustive search.

We are using reinforcement learning paradigm since it is performing well solving NAS problems.See Zoph and Le, 2016

1.1 History

Chapter 2

Background overview

The idea of using reinforcement learning agents to build neural networks is not new, however, there are not so many research projects nowadays. Mostly the reason of this is that most of research is hold by business, and business usually is not optimistic about Reinforcement Learning in production.

However, some good progress was made for recent years. In 2015, ResNet become a winner of ILSVRC 2015 in image classification, detection, and localization and winner of MS COCO 2015 detection, and segmentation. This enormous network contained of 152 layers optimized by a lot of professional engineers manually. This process is obviously expensive in terms of time and other resources. Image classification contests are constantly showing a growing amount of layers for best-performing networks (AlexNet, 2012 - 8 layers, GoogleNet, 2014 - 22 layers). Resnet has 1.7 million parameters. Each competition is turning researchers more and more towards automation of this work - and this is a place where neural architecture search becomes a new trend.

Barret Zoph and Quoc Le. in "Neural architecture search with reinforcement learning." used a recurrent network to generate the model descriptions of neural networks and train this RNN with reinforcement learning to maximize the expected accuracy of the generated architectures on a validation set. This paper is one the most cited in this field and our research is heavily based on it.

In 2019, Google researchers developed a family of models, called EfficientNets, which superpass state-of-the-art accuracy with up to 10x better efficiency (smaller and faster) using AutoML.

Amazon have two AutoML products to offer - Amazon SageMaker Autopilot for creation of the classification and regression machine learning models and Amazon DeepAR for forecasting scalar (one-dimensional) time series using recurrent neural networks (RNN). This paper is also heavily based on probabilistic approach used in DeepAR because of it's spectacular results.

A variety of methods have been proposed to perform NAS, including reinforcement learning, Bayesian optimization with a Gaussian process model, evolutionary search, and gradient descent over a past few years.

Appendix A

Code

A.1 Pseudocode

Something on the topic

Bibliography

Zoph, Barret and Quoc V. Le (2016). "Neural Architecture Search with Reinforcement Learning". In: *CoRR* abs/1611.01578. arXiv: 1611.01578. URL: http://arxiv.org/abs/1611.01578.