UNIVERSITY
OF TECHNOLOGY
SYDNEY

# Efficient and Reproducible Automated Deep Learning

**by Xuanyi Dong**

Thesis submitted in fulfilment of the requirements for the degree of

**Doctor of Philosophy**

under the supervision of
Prof. Bogdan Gabrys and Prof. Katarzyna Musial

University of Technology Sydney
Faculty of Engineering and Information Technology

Apr 2021

# CERTIFICATE OF ORIGINAL AUTHORSHIP

I, **Xuanyi Dong** declare that this thesis, is submitted in fulfilment of the requirements for the award of **Doctor of Philosophy**, in the **School of Computer Science, FEIT** at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.

This research is supported by the Australian Government Research Training Program.

Signature: Production Note: Signature removed prior to publication.

Date: **28 Apr 2021**

# ABSTRACT

**Efficient and Reproducible Automated Deep Learning**

by

Xuanyi Dong

Deep learning has shown its power in a large number of applications, such as visual perception, language modeling, speech recognition, video games, etc. To deploy a deep learning model successfully, inevitable manual tuning is required for each component, such as neural architecture design, the choice of optimization strategy, data selection, augmentation, etc. Such manual tuning costs expensive computational resources and is labor-intensive. Moreover, this paradigm is not scalable when the model size or the data size significantly increases. Fortunately, AutoDL brings hope to alleviate this problem by making the tuning procedure automated. Despite the recent success of AutoDL, efficiency and reproducibility for AutoDL algorithms remain a tremendous challenge for the community.

In this thesis, we address this challenge in the following aspects. We comprehensively review the current state of AutoDL and set up six step-by-step objectives to further develop AutoDL. To achieve these objectives, we propose a series of efficient approaches to learning to search (1) neural architecture topology, (2) neural architecture size, and (3) hyperparameters by gradient descent. In addition to common empirical analysis on vision and NLP datasets, we build a systematical benchmark for neural architecture topology and neural architecture size. This benchmark aims to provide a fair and easy-to-use environment for our proposed algorithms as well as other AutoDL participants.

Dissertation directed by Professor Bogdan Gabrys

Advanced Analytics Institute, Faculty of Engineering and IT, University of Technology Sydney

# Acknowledgements

First and foremost, I would like to thank my academic supervisor Prof. Bogdan Gabrys. I am incredibly grateful for his kind, timely, and strong support. He guided me on automated deep learning, which is my favorite research direction; he let me know what good research work is; he also kindly shared many useful suggestions regarding career development. Since 2015, when I first got in touch with deep learning, I was lucky to be mentored by many outstanding researchers. Thanks to Dr. Junjie Yan, my mentor, when I interned at SenseTime, from whom I learned how to get a good ranking at large-scale research competitions. Thanks to Prof. Shuicheng Yan and Dr. Junshi Huang, my mentors, when I interned at Qihoo 360, where I published my first paper. Thanks to Prof. Deyu Meng, my mentor, when I visited at Xi'an Jiaotong University, from whom I learned the rigorous attitude and mathematical expression of doing research. Thanks to Prof. Yaser Sheikh and Dr. Shoou-I Yu, my mentors, when I interned at Facebook Reality Labs, from whom I learned how to perform research in a systematic way. Thanks to Dr. Yan Yan, my senior, from whom I learned many machine learning methodologies and shared many experiences when I was a junior. Thanks to Prof. Yi Yang for the scholarship support of one and a half year's study at the University of Technology Sydney and his recommendation for the internship and fellowship. Thanks to Dr. Quoc V. Le and Dr. Mingxing Tan, my hosts, when I interned at Google Brain, from whom I learned critical thinking and what is the research impact. Thanks to Mr. Daiyi Peng, my co-host during my Google internship, from whom I systematically learned the industrial system and rethought the relationship between engineer and research. Thanks to Prof. Katarzyna Musial, my co-supervisor in UTS, who provided useful suggestions for my research on AutoDL.

I would also like to thank my colleagues and friends at the University of Technology Sydney and other institutes. I want to thank Xiaojun Chang, Hehe Fan, Qianyu Feng, Qingji Guan, Yang He, Yanbin Liu, Ping Liu, Yutian Lin, Peike Li, Fan Ma, Guang Li, Guangrui Li, Jiaxu Miao, Pingbo Pan, Ruijie Quan, Tianqi Tang, Bingwen Hu, Minfeng Zhu, Yu Wu, Xiaohan Wang, Zhongwen Xu, Guoliang Kang, Yan Yan, Zongxin Yang, Fengda Zhu, Hu Zhang, Zhun Zhong, Zhedong Zheng, Liang Zheng, Xiaolin Zhang, Tianjian Meng, Adams Wei Yu, Arber Zela, Cihang Xie, Xinshuo Weng, and many others. I was very fortunate to collaborate with or discuss with them. These discussions inspired and motivated many of my research works.

I would also like to thank Data to Decision CRC, CAI FEIT UTS, AAI FEIT UTS, Baidu Scholarship, and Google PhD Fellowship for supporting my research.

Lastly, I would like to thank my father, Aimin Dong, and my wife Lu Liu for their support and love throughout the years. Especially without taking any credits, Lu helped me a lot in discussing new research ideas, designing experiments, and re-writing the papers.

Xuanyi Dong
Sydney, Australia.

# List of Publications

## Selected Journal Papers

J-[1] **Xuanyi Dong**, Lu Liu, Katarzyna Musial, Bogdan Gabrys. "NATS-Bench: Benchmarking NAS Algorithms for Architecture Topology and Size", *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)* (ERA CORE Rank A*, IF=17.861)

J-[2] **Xuanyi Dong**, Yi Yang, Shih-En Wei, Xinshuo Weng, Yaser Sheikh, Shoou-I Yu. "Supervision by Registration and Triangulation for Landmark Detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)* (ERA CORE Rank A*, IF=17.861)

J-[3] **Xuanyi Dong**, Liang Zheng, Fan Ma, Yi Yang, Deyu Meng. "Few-Example Object Detection with Model Communication", *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)* (ERA CORE Rank A*, IF=17.861)

J-[4] **Xuanyi Dong**, Yan Yan, Mingkui Tan, Yi Yang, Ivor W. Tsang. "Late Fusion via Subspace Search with Consistency Preservation", *IEEE Transactions on Image Processing (TIP)* (ERA CORE Rank A*, IF=9.34)

## Selected Conference Papers

C-[5] [NAS@ICLR-2021] **Xuanyi Dong**, Mingxing Tan, Adams Wei Yu, Daiyi Peng, Bogdan Gabrys, Quoc V. Le. "AutoHAS: Efficient Hyperparameter and Architecture Search", Workshop on Neural Architecture Search (NAS) at International Conference on Learning Representations (ICLR)

C-[6] [ICLR-2020] **Xuanyi Dong**, Yi Yang. "NAS-Bench-201: Extending the Scope of Reproducible Neural Architecture Search", International Conference on Learning Representations (ICLR) (H-index=150)

C-[7] [NeurIPS-2020] Daiyi Peng, **Xuanyi Dong**, Esteban Real, Mingxing Tan, Yifeng Lu, Hanxiao Liu, Gabriel Bender, Adam Kraft, Chen Liang, Quoc V. Le. "PyGlove: Symbolic Programming for Automated Machine Learning", *Advances in Neural Information Processing Systems (NeurIPS)* (ERA CORE Rank A*)

C-[8] [NeurIPS-2019] **Xuanyi Dong**, Yi Yang. "Network Pruning via Transformable Architecture Search", *Advances in Neural Information Processing Systems (NeurIPS)* (ERA CORE Rank A*)

C-[9] [ICCV-2019] **Xuanyi Dong**, Yi Yang. "One-Shot Neural Architecture Search via Self-Evaluated Template Network", *IEEE Conference on International Conference on Computer Vision (ICCV)* (ERA CORE Rank A*)

C-[10] [ICCV-2019] **Xuanyi Dong**, Yi Yang. "Teacher Supervises Students How to Learn from Partially Labeled Images for Facial Landmark Detection", *IEEE Conference on International Conference on Computer Vision (ICCV)* (ERA CORE Rank A*)

C-[11] [CVPR-2019] **Xuanyi Dong**, Yi Yang. "Searching for A Robust Neural Architecture in Four GPU Hours", *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (ERA CORE Rank A*)

C-[12] [CVPR-2018] **Xuanyi Dong**, Shoou-I Yu, Xinshuo Weng, Shih-En Wei, Yi Yang, Yaser Sheikh. "Supervision-by-Registration: An Unsupervised Approach to Improve the Precision of Facial Landmark Detectors", *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (ERA CORE Rank A*)

C-[13] [CVPR-2018] **Xuanyi Dong**, Yan Yan, Wanli Ouyang, Yi Yang. "Style Aggregated Network for Facial Landmark Detection", *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (ERA CORE Rank A*)

C-[14] [CVPR-2017] **Xuanyi Dong**, Junshi Huang, Yi Yang, Shuicheng Yan. "More is Less: A More Complicated Network with Less Inference Complexity", *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (ERA CORE

Rank A*)

# Contents

# List of Figures

# List of Tables

# Abbreviation

ML: Machine Learning

DL: Deep Learning

NN: Neural Network

CNN: Convolutional Neural Network

RNN: Recurrent Neural Network

NAS: Neural Architecture Search

HPO: Hyperparameter Optimization

AutoML: Automated Machine Learning

AutoDL: Automated Deep Learning

RL: Reinforcement Learning

ES: Evolutionary Strategy

FLOP: FLoating Point Operation

GPU: Graphics Processing Unit

LSTM: Long Short-Term Memory

AOS: Alternative Optimization Strategy

SVD: Singular Value Decomposition

PPO: Proximal Policy Optimization