UNIVERSITY OF TECHNOLOGY SYDNEY
Faculty of Engineering and Information Technology

Regularization in Deep Neural Networks

by

Guoliang Kang

A Thesis Submitted
in Partial Fulfillment of the
Requirements for the Degree

Doctor of Philosophy

Sydney, Australia

2019

Certificate of Authorship/Originality

I certify that the work in this thesis has not been previously submitted for a degree nor has it been submitted as a part of the requirements for other degree except as fully acknowledged within the text.

I also certify that this thesis has been written by me. Any help that I have received in my research and in the preparation of the thesis itself has been fully acknowledged. In addition, I certify that all information sources and literature used are quoted in the thesis. This research is supported by the Australian Government Research Training Program.

Guoliang Kang

Feb. 2019

Production Note:
Signature removed
prior to publication.

ABSTRACT

Regularization in Deep Neural Networks

by

Guoliang Kang

Recent years have witnessed the great success of deep learning. As the deep architecture becomes larger and deeper, it is easy to overfit to relatively small amount of data. Regularization has proved to be an effective way to reduce overfitting in traditional statistical learning area. In the context of deep learning, some special design is required to regularize their training process. Generally, we firstly proposed a new regularization technique named “Shakeout” to improve the generalization ability of deep neural networks beyond Dropout, via introducing a combination of L_0, L_1, and L_2 regularization effect into the network training. Then we considered the unsupervised domain adaptation setting where the source domain data is labeled and the target domain data is unlabeled. We proposed “deep adversarial attention alignment” to regularize the behavior of the convolutional layers. Such regularization reduces the domain shift existing at the start in the convolutional layers which has been ignored by previous works and leads to superior adaptation results.

Dissertation directed by Professor Yi Yang

Center of AI, School of Software

Acknowledgements

First and foremost, I am tremendously grateful for my supervisor Yi Yang for his continuous support and guidance throughout my PhD, and for providing me the freedom to work on a variety of problems. I am grateful for Prof. Dacheng Tao, who has ever supervised me and provided me support. I am grateful for my co-supervisor Jun Li for his beneficial suggestions for my research.

I am happy to collaborate with the previous postdoc in our team Liang Zheng. Thanks for his creative guidance and suggestions for my research and academic writing. I am happy to collaborate with many creative students in our team. I am grateful for the creative discussions with them and I really appreciate the kind and useful suggestions given by them.

Thanks for all the people that ever helped me and encouraged me.

Finally, this thesis is dedicated to my parents Zhongwen Kang, Fenglan Zhang, and my wife Mingyue You, for all the years of love and support. They are always the source of my power and the reason I insist on pursuing my dream.

Guoliang Kang

Sydney, 2019.

List of Publications

Journal Papers

J-1. **G. Kang**, J. Li, and D. Tao, "Shakeout: A new approach to regularized deep neural network training", IEEE transactions on pattern analysis and machine intelligence, vol. 40, no. 5, pp. 12451258, 2018.

Conference Papers

C-1. **G. Kang**, J. Li, and D. Tao, "Shakeout: A new regularized deep neural networktraining scheme," in AAAI, 2016.

C-2. **G. Kang**, L. Zheng, Y. Yan, and Y. Yang, "Deep Adversarial Attention Alignment for Unsupervised Domain Adaptation: the Benefit of Target Expectation Maximization", in ECCV, 2018

Contents

List of Figures