

Faculty of Engineering and Information Technology
University of Technology Sydney

Single Image Super-Resolution via Deep Dense Network

A thesis submitted in partial fulfillment of
the requirements for the degree of
Master of Engineering

by

Jialiang Shen

November 2019

CERTIFICATE OF AUTHORSHIP/ORIGINALITY

I, Jialiang Shen declare that the thesis, is submitted in fulfilment of the requirements for the award of Master of Engineering, in the Faculty of Engineering and Information Technology at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise reference 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 of Candidate

Production Note:
Signature removed prior to publication.

Date: 6/11/2019

Acknowledgments

Foremost, I would like to express my sincere gratitude to my supervisor Associate Prof. Jian Zhang and my senior fellow Yucheng Wang in Baidu for the continuous support of my master study and research, for their patience, motivation, enthusiasm, and immense knowledge. Their guidance helped me in all the time of research and writing of this thesis. I could not have imagined having better advisors and mentors for my master study.

I also would like to appreciate my co-supervisor Dr. JingSong Xu for his expert and sincere help in the project and providing me with continuous support throughout my master study and research.

I thank my fellow labmates in Global Big Data and Technology Center: Zhibin Li, Xiaoshui Huang and Muming Zhao for the stimulating discussions, and supportive helps in the research periods and thesis writing.

Last but not the least, I would like to thank my parents, for their unconditional support, both financially and emotionally throughout the whole master studying.

Jialiang Shen
June 2019 @ UTS

Contents

Certificate	i
Acknowledgment	iii
List of Figures	vii
List of Tables	xi
List of Publications	xiii
Abstract	xv
Chapter 1 Introduction	1
Chapter 2 Literature Review and Mathematical Background	9
Chapter 3 Bi-Dense Network for Image Super-Resolution . .	24
Chapter 4 Memory-Optimized Dense Network for Image Super- Resolution	39
Chapter 5 One Deep Convolutional Network for Any-Scale Image Super-Resolution	52
Chapter 6 Summary	75
Bibliography	77

List of Figures

2.1	Simplified structure of (a) DenseNet. The green lines and layers denote the connections and layers in the dense block, and the yellow layers denote the transition and pooling layer. (b) SRDenseNet. The green layers are the same dense block structures as those in DenseNet. The purple lines and element-wise \oplus refer to the skip connection. (c) RDN. The yellow lines denote concatenating the blocks output for reconstruction and the green layers are residual dense blocks. (d) DBDN. The yellow lines and layers denote the inter-block dense connection and the green layers are the intra dense blocks. The output goes into upsampling reconstruction layers. The orange layers after input are the feature extraction layers in all models. . . .	15
2.2	Comparison of two-level Laplacian Pyramids. (a) Laplacian Pyramid. The decomposition step produces two residual images R_1, R_0 by subtracting H_1 with I_1 , H_0 with I_0 to preserve the high-frequency information, which then added with the interpolated LR I_1, I_0 respectively to reconstruct the $2\times, 4\times$ frequency levels H_1, H_0 . (b) LapSRN. The residual images R_1, R_0 are progressively learnt by the networks and upsampled at each level, then added with LR images for HR images.	19
3.1	Network structure	26
3.2	Intra dense blocks	30

LIST OF FIGURES

3.3 Qualitative comparison of our model with other works on $\times 2$ super-resolution. Red indicates the best performance and blue indicates the second best. 34

3.4 Qualitative comparison of our model with other works on $\times 4$ super-resolution. Red indicates the best performance and blue indicates the second best. 35

3.5 Performance vs number of parameters. The results are evaluated with Set5 for $4\times$ enlargement. Red indicates the best performance and blue indicates the second best. 35

3.6 Discussion about transition layer, skip connection and inter-block dense connectivity in DBDN. The results are evaluated with Set5 for $3\times$ enlargement in 200 epochs 37

4.1 Our network structure with densely connected residual block and deep supervision. 41

4.2 Densely connected residual block 43

4.3 Computation graph of memory optimized network 45

4.4 Qualitative comparison.(1)In the first row image "img014"(Set14 with scale factor $\times 4$), the zebra texture is more detailed in our method.(2)In the second row shows image"img021"(B100 with scale factor $\times 3$), the black line is clearer in our method (3) In the third row image "img082" (Urban100 with scale factor $\times 2$), we can see that is two line in ours. 49

5.1 The Laplacian Frequency Representation has 11 Laplacian frequency levels ($O_{r_0}, \dots, O_{r_{10}}$), with 10 phase in the scale range of $(1, 2]$ (P_1, \dots, P_{10}). Each phase represents the difference between the successive Laplacian frequency levels. 56

5.2 (a) The overall architecture of the proposed ASDN network, multiple image reconstruction branches (IRBs) parallel allocate after feature mapping branch (FMB). The FMB adopts the bi-dense structure from DBDN and the spatial attention in IRB is the same spatial attention module from CSFM. (b) The dense attention block (DAB) in a, which combines the Intra dense block from DBDN and channel attention module from RCAN. (c) The illustration of the adopted spatial attention (SA) and channel attention (CA) modules from CSFM and RCAN. 59

5.3 The architecture of fixed-scale deep network (FSDN) for SR of multiple integer scales. FMB is the Feature Mapping Branch structure in ASDN and the Upscaling Reconstruction Branches (URB) parallel construct behind FMB for multi-scale SR, which are built by a deconvolutional layer followed with Image Reconstruction Branch (IRB). 61

5.4 PSNR comparison of ASDN with other works within continuous scale range ($\times 2, \times 8$] on Set5 68

5.5 Qualitative comparison of our models with other works on $\times 2$ super-resolution **Red** indicates the best performance, **blue** indicates the second best 69

5.6 Qualitative comparison of our models with other works on $\times 4$ super-resolution **Red** indicates the best performance, **blue** indicates the second best 70

5.7 Qualitative comparison of our models with other works on $\times 8$ super-resolution **Red** indicates the best performance, **blue** indicates the second best 70

5.8 PSNR of Laplacian frequency represented and specific scale trained HR images of scales in (1, 2] on Set5 71

5.9 PSNR of various ASDNs trained with different Laplacian frequency level densities 72

List of Tables

3.1	Public benchmark test results and Manga109 results (PSNR(dB)/SSIM). Red indicates the best performance and blue indicates the second best.	33
4.1	Public benchmark test results (PSNR(dB)/SSIM). Red indicates the best performance and blue indicates the second best.	48
4.2	The depth analysis of network	50
4.3	The memory analysis of network	51
5.1	Quantitative evaluation of state-of-the-art SR algorithms. We report the average PSNR/SSIM for $2\times$, $3\times$, $4\times$ and $8\times$ SR. Black indicates the best performance	64
5.2	Results of any-scale SR on different methods tested on BSD100. The first row shows the results of Laplacian frequency levels and the second column row demonstrates SR performance on randomly selected scales and boundary condition. Black indicates the best performance	67
5.3	PSNR of the recursive deployment and direct deployment on SR for $\times 2$, $\times 3$, $\times 4$	72
5.4	PSNR of recursive deployment and direct deployment on SR for $\times 2$, $\times 3$, $\times 4$. Black indicates the best performance	73
5.5	Investigate of channel attention (CA), spatial attention (SA) and skip connection (SC). Black indicates the best performance	74

List of Publications

Journals

- Yucheng Wang, **Jialiang Shen**, Jian Zhang (2019), ASDN: One Deep Convolutional Network for Any-Scale Image Super-Resolution. *submitted to 'Transactions on Multimedia (TMM)'. IEEE.*

Conferences

- **Jialiang Shen**, Yucheng Wang, Jian Zhang (2018), Memory-Optimized Deep Dense Network for Super-Resolution. *in 'International Conference on Digital Image Computing: Techniques and Applications (DICTA)'. IEEE.*
- Yucheng Wang, **Jialiang Shen**, Jian Zhang (2018), Deep Bi-Dense Network for Image Super-Resolution. *in 'International Conference on Digital Image Computing: Techniques and Applications (DICTA)'. IEEE.*

Abstract

Image Super-Resolution (SR) is a research field of computer vision, which enhances the resolution of an imaging system. The need for high resolution is common in computer vision applications for better performance in pattern recognition and analysis of images. However, recovering of the HR image from LR image is a highly ill-posed problem. In this thesis, the image SR problem is solved from three aspects with deep dense network models, including improving reconstruction accuracy, optimizing model training-time memory consumption, and extending effective SR scale ranges. Chapter 1 introduces the importance of image SR reconstruction and summarizes the challenges of image SR problem. Chapter 2 reviews the existing image SR methods, analyses their limitations and explains some related fundamental theories. Chapter 3 proposes a bi-dense model to improve image SR performance based on the dense connections for feature reuse. The bi-dense network does not only reuse local feature layers in the dense block, but also reuses the block information in the network to archive excellent performance with a moderate number of parameters. Chapter 4 evaluates the memory consumption of the vanilla dense model for image SR. For solving this problem, we introduce shared memory strategy into image SR by proposing a memory-optimized deep dense network. Chapter 5 discovers most of the deep SR methods are inefficient or impractical for generating SR of any scale factor, and proposes a novel Any-Scale Deep Network (ASDN), which requires few training scales to achieve one unified network for any-scale SR. In order to design such a powerful network architecture, we propose Laplacian

ABSTRACT

Frequency Representation to predict SR results of the small ratio range and Recursive Deployment for SR of any larger scale. In this way, the required training data and update periods are substantially decreased to optimize the any-scale SR network. All these algorithms are aimed to solve the single image SR problem. These algorithms are tested on many public datasets and the results on those datasets demonstrate superior performance of our approach over the state-of-the-art methods and validate the effectiveness and correctness of our methods.