UNIVERSITY OF TECHNOLOGY SYDNEY

Low Light Image Enhancement and Saliency Object Detection

by Yuanfang Zhang

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

under the supervision of Professor Xiangjian He

University of Technology Sydney
Faculty of Engineering and Information Technology

May 2022

Certificate of Authorship/Originality

I, Yuanfang Zhang, declare that this thesis is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the Faculty of Engineering and Information Technology, 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.

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of the requirements for a degree at any other academic institution except as fully acknowledged within the text. This thesis is the result of a Collaborative Doctoral Research Degree program with Northwestern Polytechnical University.

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

Production Note:
Signature removed
prior to publication.

Signature ________________

Date 5th May, 2022

ABSTRACT

Low Light Image Enhancement and Saliency Object Detection

by

Yuanfang Zhang

Low light images represent a series of image types with great potential. Their research focuses on images and videos of the environment at dusk and near darkness. It can be widely used in night safety monitoring, license plate recognition, night scene shot, special target recognition at dusk, and other emergency events that occur under light scenes. After the environment is enhanced and combined with other tasks in computer vision and pattern recognition, it can bring many results, such as saliency detection and object detection under low illumination, and abnormal detection in crowded places under low-light environment. For the enhancement of low light and low light scenes, using traditional methods often results in over-exposure and halo conditions. Therefore, using deep learning network technology can fix and improve these specific shortcomings. To achieve this goal, we have done several investigations about the current state-of-art researches on low-light enhancement and the relevant computer vision tasks. For low light image enhancement, a series of qualitative and quantitative experimental comparisons conducted on a benchmark dataset demonstrate the superiority of our approach, which overcomes the drawbacks of white and colour distortion. At present, most of the research works on visual saliency have concentrated on the field of visible light, and there are few studies on night scenes. Due to insufficient lighting conditions in night scenes, and relatively lower contrasts and signal-to-noise ratios, the effectiveness of available visual features is greatly reduced. Moreover, without sufficient depth information, many features and clues are lost in the original images. Therefore, the detection of salient targets in night scenes is also difficult and it is a focus of current research in the field of computer vision. The performance leads to vague effects when the existing methods are directly conducted, so we adopt a new “enhance firstly detection secondly” mechanism that firstly enhances the low-light images in order to improve the contrast and visibility, and then combines it with relevant saliency detection methods with depth information. Furthermore, we concern about the feature aggregation schemes for deep RGB-D saliency object detection and propose novel feature aggregation methods. Meanwhile, for the monocular vision, of which the depth information is hard to acquire, a novel RGB-D image saliency detection method is proposed to leverage depth cues for enhancing the saliency detection performance but without actually using depth data. Both of the extra depth cues and the proposed “enhance firstly detection secondly” mechanism can improve saliency detection abilities, according to the experimental results. The model not only outperforms the state-of-the-art RGB saliency models, but also achieves comparable or even better results compared with the state-of-the-art RGB-D saliency models

Dissertation directed by Professor Xiangjian He, Professor Michael Blumenstein and Doctor Wenjing Jia
Faculty of Engineering and Information Technology

Dedication

To my parents, and those who always love me and support me along the way.

Acknowledgements

My doctor study at UTS in the past three years has been a life-changing and priceless experience for me. Sydney is a lovely place. It has a golden light harbour with white sails, delicate and charming beaches, and a mild Mediterranean climate. The streets are filled with wild scents, lush forests, and soaring seagulls. Its natural beauty is enhanced by golden beaches and unspoiled bush lands. Sydney is a fantastic location for scientists to investigate science mysteries.

First and foremost, I would like to express my heartfelt appreciation to Professor Xiangjian He, my Principal Supervisor, who helped me tremendously by supplying me with required tools, valuable guidance, and inspiration for new ideas with exceptional patience and constant encouragement. His recommendations have drawn my attention to a variety of flaws and clarified many questions for me.

I am also grateful to Professor Michael Blumenstein and Doctor Wenjing Jia, for being my co-supervisors, especially to Dr Jia for her thoughtfulness and generosity in arranging study meetings in Professor He's Computer Vision and Pattern Recognition (CVPR) group. In research, Dr Jia is very proactive and detail-oriented, and her passion for collaboration and ability to contribute has bonded all of the research students into a family. With Dr Jia and the CVPR study community, I am really enjoying studying sophisticated deep learning techniques in image analysis and pattern recognition.

Then, I would like to express my gratitude to my classmates for their invaluable help with the original manuscript. They kindly made important remarks and sound recommendations to the paper's outline.

I would like to thank the teachers, writers, and colleagues at UTS for their time and commitment. Special thanks to Xiaochen Fan, Yue Xi, Xudong Song, Qingqing Wang, Edward Huang, Saeed Amirgholipour, Muhammad Usman, Hesam Hesamian, Haodong Chang, Farhan Mohammed, and Chengpei Xu, who provided me with a great deal of assistance at different times.

Finally, I would like to express my gratitude to my parents for their encouragement in allowing me to study abroad. It is very beneficial to me in completing this study and I always love them.

In one word, I would like to express my thanks to all who have assisted me in finishing this study.

List of Publications

Journal Papers

J-1. Zhang Y, Zheng J, Li L, Nian L, Wenjing Jia, Xiaochen Fan, Chengpei Xu, Xiangjian He. Rethinking feature aggregation for deep RGB-D salient object detection [J]. Neurocomputing, 2021, 423: 463-473.

J-2. Zhang Y F , Zheng J , Jia W , et al. Deep RGB-D Saliency Detection without Depth[J]. IEEE Transactions on Multimedia, 2021, PP(99):1-1. Doi: 0.1109/TMM.2021.3058788

J-3. Zhang Y, Zheng J, Fei Li, Wenjing Jia, Wenfeng Huang, Xiangjian He. Low Light Image Dedarking via Deep Semantic Fusion[J]. IEEE Signal Processing Letter (Under Review)

Contents

List of Figures

List of Tables

Abbreviation

ASPP - Atrous Spatial Pyramid Pooling

DASPP - Dense Atrous Spatial Pyramid Pooling

SOD - Saliency Object Detection

CGA - Convolutional Gated Attention

SA - Spatial Attention

CBAM- Convolutional Block Attention Module

DMSF- Dense MultiScale Fusion

NIQE- Natural image quality evaluator

PSNR- Peak Signal to Noise Ratio

SSIM- Structural Similarity

CNN - Convolutional Neural Networks

ExDARK - A Dataset for low light image enhancement

MSR - Multi-scale Retinex

LIME - A Dataset for low light image enhancement

LECARM - A Dataset for low light image enhancement

AAP - Adaptive Average Pooling

BN - Batch Normalization

ReLU - Rectified Linear Unit

UP - Upsampling

PCC - Pearson Correlation Coefficient

NL - Non Local

GT - Ground Truth

HA - Holistic Aggregation

EA - Early Aggregation

BU - Bottom-up

FGA - Factorized Gated Attention

SSF - A model for Saliency Object Detection

UCNet - A model for Saliency Object Detection

JLDCF - A model for Saliency Object Detection

NJUD - A dataset for Saliency Object Detection

NLPR - A dataset for Saliency Object Detection

SSD - A dataset for Saliency Object Detection

RGBD135 - A dataset for Saliency Object Detection

STEREO - A dataset for Saliency Object Detection

DUT-RGBD - A dataset for Saliency Object Detection

A2dele - A model for Saliency Object Detection

SOTA - State-of-the-art

Amulet - A model for Saliency Object Detection

DSS - A model for Saliency Object Detection

BMP - A model for Saliency Object Detection

PiCANet - A model for Saliency Object Detection

R3Net - A model for Saliency Object Detection

CPD - A model for Saliency Object Detection

EGNet - A model for Saliency Object Detection

PoolNet - A model for Saliency Object Detection

BASNet - A model for Saliency Object Detection

MINet - A model for Saliency Object Detection

ITSD - A model for Saliency Object Detection

DF - A model for Saliency Object Detection

AFNet - A model for Saliency Object Detection

CTMF - A model for Saliency Object Detection

MMCI - A model for Saliency Object Detection

PCF - A model for Saliency Object Detection

TANet - A model for Saliency Object Detection

CPFP - A model for Saliency Object Detection

DMRA - A model for Saliency Object Detection

S^2MA - A model for Saliency Object Detection

RGB-D - RGB and Depth

LFSD - A dataset for Saliency Object Detection

maxF - A performance index for Saliency Object Detection

STERE - A dataset for Saliency Object Detection

SIP - A dataset for Saliency Object Detection

SSD - A dataset for Saliency Object Detection

MAE - A performance index for Saliency Object Detection

PR Curve - A performance index for Saliency Object Detection