Faculty of Engineering and Information Technology
University of Technology Sydney

Rail Infrastructure Defect Detection Through Video Analytics

A thesis submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

by

Huaxi Huang

February 2022

CERTIFICATE OF ORIGINAL AUTHORSHIP

I, Huaxi Huang declare that this thesis, is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the School of Electrical and Data Engineering, 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.

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

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

Production Note:
Signature: Signature removed prior to publication.
Date: 07/02/2022

Acknowledgments

First of all, I would like to express my most profound appreciation to my principal supervisor, Prof. Jian Zhang, for his professional guidance, reliable help, and perpetual support during my pursuit of the Ph.D. degree and three and a half years of research.

I would also like to express my sincere appreciation to my co-supervisor Prof. Qiang Wu and the collaborators: Dr. Junjie Zhang, Dr. Chang Xu, and Dr. Jingsong Xu, for not only their comments on revising my manuscript but also for the insightful guidance, which have incented me to improve my research ability and broaden my research horizon.

I am grateful to my colleagues and friends at the UTS Multimedia and Data Analytics Lab: Yazhou Yao, Xiaoshui Huang, Zhibin Li, Yongshun Gong, Lu Zhang, Anan Du, Lingxiang Yao, Guofeng Mei, Wenbo Xu, Litao Yu and all other labmates. I enjoyed the time we spent together.

Finally and most essentially, I would like to thank my parents. Their selfless support and continuing encouragement helped me defeat the obstacles I encountered during my Ph.D. study.

Huaxi Huang
February 2022 @ UTS

Contents

List of Figures

List of Tables

List of Publications

Papers Published

- **Huaxi Huang**, Junjie Zhang, Jian Zhang, Qiang Wu, Chang Xu, *PTN: A Poisson Transfer Network for Semi-supervised Few-shot Learning*, in Proceeding of the 35th AAAI Conference on Artificial Intelligence (AAAI-21), pp: 1602-1609, 2021.
- **Huaxi Huang**, Junjie Zhang, Jian Zhang, Qiang Wu, Chang Xu, *TOAN: Target-Oriented Alignment Network for Fine-Grained Image Categorization with Few Labeled Samples*, IEEE Transactions on Circuits and System for Video Technology (TCSVT), vol: 32, pp: 853-866, 2022.
- **Huaxi Huang**, Junjie Zhang, Jian Zhang, Jingsong Xu, Qiang Wu. *Low-Rank Pairwise Alignment Bilinear Network For Few-Shot Fine-Grained Image Classification.* IEEE Transactions on Multimedia (TMM), vol: 23, pp: 1666-1680, 2021.
- **Huaxi Huang**, Junjie Zhang, Jian Zhang, Qiang Wu and Jingsong Xu, *Compare More Nuanced: Pairwise Alignment Bilinear Network for Few-Shot Fine-Grained Learning,* in IEEE International Conference on Multimedia and Expo (ICME-19), pp. 91-96, 2019.
- **Huaxi Huang**, Jingsong Xu, Jian Zhang, Qiang Wu, et al. *Railway Infrastructure Defects Recognition using Fine-grained Deep Convolutional Neural Network.* in IEEE International Conference on Digital Image Computing: Techniques and Application, pp: 1-8, 2018.

Abstract

Compared with the traditional railway infrastructure maintenance process, which relies on manual inspection by professional maintenance engineers, inspection through automatic video analytics will significantly improve the working efficiency and eliminate the potential safety concern by reducing physical contact between maintenance engineers and infrastructure facilities. However, the defect does not always have a stable appearance and involves many uncertainties exposed in the clutter environments. On the other hand, various brands of the same devices are used widely on the railway, which shows diverse physical models. Therefore, it creates many challenges to the existing computer vision algorithms for defect detection. In this thesis, two key challenges are abstracted with regard to video/image analytics using computer vision techniques for railway infrastructure defect detection, resulting from the fine-grained defect recognition and the limited labeled learning (few-shot learning). This thesis summarizes the works that have been conducted on utilizing different methods to solve the two challenges.

The first challenge is fine-grained defect recognition. For railway infrastructure defect inspection, damaged or worn equipment defects are usually found in some small parts compared to the whole object. That is, the differences between the defective ones and standard ones are fine-grained. How to find these subtle defects is a fine-grained recognition problem. This thesis proposes a bilinear CNNs model to tackle the defect detection problem, which effectively captures the invariant representation of the dataset and learns high-order discriminative features for fine-grained defect recognition.

Another challenge is the limited labeled data (few-shot learning). In many scenarios, obtaining abundant labeled samples is laborious. For example, in industrial defect detection, most defects exist only in a few common categories, while most other categories only contain a small portion of defects. Moreover, annotating a large-scale dataset of railway infrastructure defects is labor-intensive, which requires high expertise in railway maintenance. Thus, how to obtain an effective model with sparse labeled samples remains an open problem. To address this issue, this thesis proposes a framework to simultaneously reduce the intra-class variance and enlarge the inter-class discrimination for both fine-grained defect recognition and general fine-grained recognition under the few-shot setting. Three models are designed according to this framework, and comprehensive experimental analyses are provided to validate the effectiveness of the models. This thesis further studies the few-shot learning problem by mining the unlabeled information to boost the few-shot learner for defect/general object recognition and proposes a Poisson Transfer Model to maximize the value of the extra unlabeled data through robust classifier construction and self-supervised representation learning.