UTS UNIVERSITY OF TECHNOLOGY SYDNEY

PERSON RE-IDENTIFICATION IN THE WILD: FROM SHORT-TERM TO LONG-TERM

by Yan Huang

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

under the supervision of Qiang Wu

University of Technology Sydney
Faculty of Engineering and Information Technology

05/2021

Certificate of Original Authorship Template

CERTIFICATE OF ORIGINAL AUTHORSHIP

I, *Yan 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.

Signature: Production Note:
Signature removed prior to publication.

Date: 21/05/2021

ABSTRACT

PERSON RE-IDENTIFICATION IN THE WILD: FROM SHORT-TERM TO LONG-TERM

by

YAN HUANG

The task of person re-identification (re-ID) is to confirm the identity of a person in visual traces captured by different cameras. Person re-ID "in the wild" is a highly demanded technology and quite challenging due to the lack of data diversity, dramatic background variation between different domains, uncontrollable clothing change, and influence caused by the shortage of lighting. According to the scale of time gaps when footages are captured, these challenges are exposed into two different scenarios: 1) Short-Term person re-ID (ST-reID), and 2) Long-Term person re-ID (LT-reID). ST-reID addresses the time gap of several minutes. This scenario is mainly to deal with variations of illumination, viewpoint, and pose. Thus, data diversity is the primary concern. In addition, when footages are taken in different environments, ST-reID often encounters a large background shift issue. On the contrary, the time gaps between two footages in LT-reID can be several hours or even longer. Thus, a person has a great chance to change clothing. As another widely seen case, after a long-time gap, in order to take footages to re-identify a person when s/he reappears at night, infrared cameras are required.

For ST-reID in the wild, the diversity of training data is essential to ensure a re-ID system can tolerate variations of illumination, viewpoint, and pose, *etc.* In addition, a model trained on one domain can lack certain generalization when it is applied to a new domain. This thesis will deeply study the two challenges exposed in ST-reID. Corresponding solutions are provided by using generated data to compensate for the limited data diversity. Also, a background shift suppression

model is proposed to deal with the background shift issue for cross-domain ST-reID.

For LT-reID in the wild, it is worth investigating approaches to tackle the clothing change issue. This thesis will introduce new clothing change datasets to the community. Corresponding solutions are given to tackle the clothing change issue. In addition, under tight security surveillance in LT-ReID, how to recognize the same person who appears under a RGB camera (in the daytime) and an infrared camera (reappear at night) is an immediate problem. A modality-biased training issue is unveiled for the infrared-visible LT-reID task, and corresponding solutions are given. This thesis will provide useful insights into diverse person re-ID issues in the wild from the short-term scenario to the long-term scenario to support practical usages in the real world.

Dissertation directed by Associate Professor Qiang Wu
School of Electrical and Data Engineering

Acknowledgements

Throughout the past few years during my PhD study I have received a great deal of support and assistance from my supervisor, my colleagues, and my family.

First, I would like to thank my principle supervisor A/Prof. Qiang Wu whose expertise was invaluable in formulating the research questions and methodology. His insightful feedback on my researches helped me to sharpen my thinking and brought my work to a higher level. During my PhD study, A/Prof. Qiang devoted lots of efforts to better my researches on computer vision. My first paper was accepted by a top-ranked journal (TIP, IF:9.34) under the patient guidance of A/Prof. Qiang. I learned a lot of skills about academic writing in my first paper from my supervisor. My second paper was accepted by one of the most challenging computer vision conference ICCV2019, and its extension version has been accepted by IJCV which is one of the top journal in computer vision field. I can still clearly remember how A/Prof. Qiang shared his valuable insight to this paper. His scientific and academic rigour impresses me much more than the outcomes. In addition, A/Prof. Qiang provided lots of support in data acquisition for my research. I am also very fortunate to have chance to work with other researchers and PhD students introduced by A/Prof. Qiang, which influenced my a lot to my research career. In the end, I appreciate A/Prof. Qiang's strong support to my scholarship (AGRTP and IRS) applications, making my lives much easier in Sydney.

Second, I would like to thank my co-supervisor A/Prof. Jian Zhang and Dr. Jingsong Xu. Jian organized many seminars that provided me opportunities to communicate with other researchers. Jian also organized the VCIP2019 conference at UTS. As a volunteer of the conference, I broaden my mind and learn something new from industries and researchers all over the world. Jingsong gave me a lot of constructive suggestions to my researches and always encourages me to focus on the

cutting-edge techniques of computer vision.

Third, I would like to thank all my dear co-authors and colleagues during my doctoral study. Thanks to Prof. Zhaoxiang Zhang (from NLPR), Asst/Prof. Yi Zhong (from BIT), Dr. Peng Zhang (from UTS), and Dr. Zhedong Zheng (from UTS) for their strong support to my publications. Thanks to A/Prof. Deyang Liu and Junyi Wu for their collaborative papers. Thanks to A/Prof. Yifan Zuo who introduced me to my supervisor. Thanks to Dr. Zongjian Zhang, Dr. Lu Zhang, Prof. Yazhou Yao, Dr. Xiaoshui Huang, Dr. Zhibin li, Dr. Huaxi Huang, Dr. Muming Zhao, Dr. Xunxiang Yao, Dr. Junjie Zhang, Dr. Lina Li, Dr. Yongshun Gong, Dr. Lingxiang Yao, Dr. Qian Li, Dr. Anan Du, Jialiang Shen and all other colleagues. With them, my lives in Sydney were rich and colourful.

In the end, I would like to thank my parents for their support, trust, encouragement, and love throughout my doctoral studies these years.

Yan Huang

Sydney, Australia, 2021.

List of Publications

Journal Papers

J-1. **Y. Huang**, Q. Wu, J. Xu, Y. Zhong and Z Zhang. "Unsupervised Domain Adaptation with Background Shift Mitigating for Person Re-Identification," *Springer International Journal of Computer Vision (IJCV)*, 2021.

J-2. **Y. Huang**, Q. Wu, J. Xu, Y. Zhong, P. Zhang and Z. Zhang. "Alleviating Modality Bias Training for Infrared-Visible Person Re-Identification," *IEEE Transactions on Multimedia (TMM)*, accepted, 2021.

J-3. **Y. Huang**, J. Xu, Q. Wu, Y. Zhong, P. Zhang and Z. Zhang, "Beyond scalar neuron: Adopting vector-neuron capsules for long-term person re-identification," *IEEE Transactions on Circuits and Systems for Video Technology (TCSVT)*, vol. 30, no. 10, pp. 3459-3471, 2020.

J-4. **Y. Huang**, J. Xu, Q. Wu, Z. Zheng, Z. Zhang and J. Zhang, "Multi-Pseudo Regularized Label for Generated Data in Person Re-Identification," *IEEE Transactions on Image Processing (TIP)*, vol. 28, no. 3, pp. 1391-1403, 2019.

J-5. **Y. Huang**, Y. Zhong, Q. Wu, E. Dutkiewicz and T. Jiang, "Cost-effective foliage penetration human detection under severe weather conditions based on auto-encoder/decoder neural network," *IEEE Internet of Things Journal (IoTJ)*, vol. 6, no. 4, pp. 6190-6200, 2019.

J-6. P. Zhang, J. Xu, Q. Wu, **Y. Huang** and X. Ben, "Learning Spatial-temporal Representations over Walking Tracklet for Long-term Person Re-Identification in The Wild," *IEEE Transactions on Multimedia (TMM)*, 2020.

J-7. L. Huang, Q. Yang, J. Wu, **Y. Huang**, Q. Wu and J. Xu, "Generated Data With Sparse Regularized Multi-Pseudo Label for Person Re-Identification,"

IEEE Signal Processing Letters (SPL), vol. 27, pp. 391-395, 2020.

J-8. P. Zhang, J. Xu, Q. Wu, **Y. Huang** and J. Zhang, "Top-Push Constrained Modality-Adaptive Dictionary Learning for Cross-Modality Person Re-Identification," *IEEE Transactions on Circuits and Systems for Video Technology (TCSVT)*, 2020.

J-9. Y. Zhong, Y. Yang, X. Zhu, **Y. Huang***, E. Dutkiewicz, Z. Zhou and T. Jiang, "Impact of Seasonal Variations on Foliage Penetration Experiment: A WSN-Based Device-Free Sensing Approach," *IEEE Transactions on Geoscience and Remote Sensing (TGRS)*, vol. 56, no. 9, pp. 5035-5045, 2018.

J-10. D. Liu, **Y. Huang**, Q. Wu, R. Ma and P. An. "Multi-Angular Epipolar Geometry based Light Field Angular Reconstruction Network," *IEEE Transactions on Computational Image (TCI)*, 2020.

J-11. Y. Zhong, S. Wu, J. Wang, T. Jiang, **Y. Huang*** and Q. Wu. "Multi-Location Human Activity Recognition via MIMO-OFDM Based Wireless Networks: An IoT-Inspired Device-Free Sensing Approach," *IEEE Internet of Things Journal (IoTJ)*, 2020.

J-12. J. Wu, **Y. Huang***, Q. Wu, Z. Gao, J. Zhao and L. Huang. "Dual-Stream Guided-Learning Via Priori Optimization for Person Re-identification," *Springer Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, accepted, 2020.

Conference Papers

C-1. **Y. Huang**, Q. Wu, J. Xu and Y. Zhong, "SBSGAN: Suppression of Inter-Domain Background Shift for Person Re-Identification," *Proc. IEEE International Conference on Computer Vision (ICCV)*, pp. 9526-9535, Oct, 2019.

C-2. **Y. Huang**, Q. Wu, J. Xu and Y. Zhong, "Celebrities-ReID: A Benchmark for Clothes Variation in Long-Term Person Re-Identification," *Proc. IEEE International Joint Conference on Neural Network (IJCNN)*, pp. 1-8, July, 2019.

C-3. J. Wu, L. Yao, **Y. Huang**, J. Xu, Q. Wu and L. Huang, "Improving Person Re-Identification Performance Using Body Mask Via Cross-Learning Strategy," *Proc. IEEE Visual Communications and Image Processing (VCIP)*, Dec, 2019.

Submitted Papers

J-1. **Y. Huang**, Q. Wu, J. Xu, Y. Zhong, P. Zhang and Z Zhang. "Learning Directly from Style Invariant Component for Infrared-Visible Person Re-Identification," *IEEE Transactions on Cybernetics*, 2020 (Under review).

J-2. X. Ding, T. Jiang, J. Yang, Y. Zhong, **Y. Huang** and S. Wu. "WiLISensing: Device-free Location-independent Human Activity Recognition via Wi-Fi," *IEEE Internet of Things Journal*, 2020 (Under review).

Contents

List of Figures

List of Tables

Abbreviation

Person re-ID - Person re-IDentification

ST-reID - Short-Term person re-IDentification

LT-reID - Long-Term person re-IDentification

GAN - Generative Adversarial Network

MBT - Modality Bias Training

DLS - Dual-level Learning Strategy

CNN - Convolutional Neural Network

SBSGAN - Suppression of Background Shift Generative Adversarial Network

MpRL - Multi-pseudo Regularized Label

DA-2S - Densely Associated 2-Stream

SN - Scalar Neuron

VN - Vector Neuron

IDE - ID-discriminative Embedding

HOF - Histograms of Optical Flow

MBH - Motion Boundary Histogram

LSRO - Label Smooth Regularization for Outliers

LSR - Label Smoothing Regularization

sMpRL - static Multi-pseudo Regularized Label

dMpRL-I - dynamic Multi-pseudo Regularized Label-I

dMpRL-II - dynamic Multi-pseudo Regularized Label-II

DPM - Deformable Part Model

ISDC - Inter-Stream Densely Connection

DCCV - Dynamic Clustering Confidence Value

GAP - Global Average Pooling

FC - Fully-Connected layer

CE - Cross-Entropy loss

BN - Batch Normalization

mAP - mean Average Precision

SEA - Soft Embedding Attention

FSR - Feature Sparse Representation

ReIDCaps - Re-IDentification Capsule network

P-Caps - Primary Capsules

C-Caps - Classification Capsules

R-by-A - Routing-by-Agreement

SE - Squeeze-and-Excitation

TS1 - Training Strategy 1

TS2 - Training Strategy 2

RS - Robustness Score

dIDeS - dynamic ID-exclusive Smooth

CMC - Cumulated Matching Characteristics

SGD - Stochastic Gradient Descent

WRT - Weighted Regularization Triplet