Xintong Wang

PQ509, Mong Man Wai Building, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China

(+86) 155-2146-5116

https://ethanscuter.github.io/

m.e.xintong@gmail.com

EDUCATION

South China University of Technology	**School of Computer Science and Engineering**	**Master of Engineering**
GPA: 3.6/4.0	Computer Science	09.2016-06.2019
South China University of Technology	**School of Computer Science and Engineering**	**Bachelor of Engineering**
GPA: 3.4/4.0	Computer Science & Technology	09.2012-06.2016

WORKING EXPERIENCE

The Hong Kong Polytechnic University	**Research Assistant**	**Hong Kong, China**
Advisor: Prof. Wenjie Li		Oct. 2018 - Aug. 2019
The Chinese Academy of Sciences	**Research Assistant**	**Shenzhen, China**
Advisor: Prof. Min Yang		Mar. 2018 - Oct. 2018

RESEARCH INTERESTS

Deep Learning, Natural Language Processing, Computer Vision

Current focus on Natural Language Generation, Abstractive Summarization

PUBLICATIONS

[1] **Xintong Wang**, Wenjie Li. Fusing External Language Model in Abstractive Summarization. WWW-2020: The Web Conference, Under Review. 2019

[2] <u>Min Yang</u>, **Xintong Wang**, Yao Lu, Jianming Lv, Qiang Qu. Generative Adversarial Network for Abstractive Text Summarization with Multi-task Constraint. Information Science Under Review. 2019

[3] <u>Jianming Lv</u>, **Xintong Wang**. Cross-dataset Person Re-Identification Using Similarity Preserved Generative Adversarial Networks. KSEM-18: International Conference on Knowledge Science, Engineering and Management. 2018

[4] <u>Jianming Lv</u>, Qing Li, **Xintong Wang**. T-CONV: A Convolutional Neural Network For Multi-scale Taxi Trajectory Prediction. BigComp-18: IEEE International Conference on Big Data and Smart Computing. 2018

[5] HaitaoYang, Jianming Lv, Fei Xu, **Xintong Wang**, Yilin Huang, Lanting Xia, and Xuewu Zhu. Regression Approach for Optimal Purchase of Hosts Cluster in Fixed Fund for Hadoop Big-data Platform. 19th International Conference on Smart City, Transportation and Buildings. 2017

[6] <u>Jianming Lv</u>, **Xintong Wang**, Fengtao Huang, Junjie Yang. TREST: A Hadoop Based Distributed Mobile Trajectory Retrieval System. DSC-16: IEEE International Conference on Data Science in Cyberspace. 2016

[7] Haibiao Lin, Jianming Lv, Can Yang, Miaoyi Deng, Kaitao Wang, **Xintong Wang**. GPS Trajectory Mining : a Survey. In Journal of Computational Information Systems: Vol. 10 (16). 2014

Underline Means My Supervisor Who Is Both the First Author and Corresponding Author.

RESEARCH EXPERIENCE

Fusing External Language Model in Abstractive Summarization Comp@PolyU

Supervised by Prof. Wenjie Li, The Hong Kong Polytechnic University *Oct. 2018 - May. 2019*

Recent sequence-to-sequence neural network models provide a viable new solution to abstractive text summarization, which aims to rewrite a long text into a short and concise form while preserving the most crucial information. However, these models face significant challenges when generating both semantically and syntactically correct summaries. In this work, we explore the potential approaches to incorporate an external (pre-trained) language model to augment the linguistic quality of text generation. This allows the internal (decoder) language model to focus more on jointly learning

summary content selection and generation. Fused with the external language model, our abstractive summarization model achieves the results comparable to state-of-the-art models in terms of ROUGE scores, and meanwhile shows significant improvements in human evaluations.

The paper has been submitted to The Web Conference. WWW 2020.

Generative Adversarial Network for Abstractive Text Summarization SIAT@CAS

Supervised by Prof. Min Yang, Chinese Academy of Sciences *Mar. 2018 - Oct. 2018*

We proposed an adversarial process for abstractive text summarization, in which we simultaneously train a generative model and a discriminative model which attempts to distinguish the generated summary from the ground truth summary. Furthermore, we additionally propose extended regularizations for the generative model using the multi-task learning, sharing its LSTM encoder and LSTM decoder with the text categorization task and the syntax annotation task, respectively. The auxiliary tasks help to improve the quality of locating salient information of a document and generate high-quality summaries

The paper has been submitted to Information Science.

Cross-dataset Person Re-Identification Using Similarity Preserved Generative Adversarial Networks. Intelligent Information Fusion Lab@SCUT

Supervised by Prof. Jianming Lv, South China University of Technology *Sep. 2017 - Jan. 2018*

Due to the expensive cost of data labeling, most of the proposed Re-ID algorithms conduct supervised learning on small labeled datasets. Directly deploying these trained models to the real-world large-scale camera networks may lead to poor performance. We address this cross-dataset Re-ID challenge by transforming the unlabeled images in the target domain to fit the classifier using our proposed similarity preserved generative adversarial networks model. Comprehensive experiments based on real datasets indicate that our model is better than the state-of-the-art cross-dataset unsupervised transfer learning algorithm.

The paper has been accepted by KSEM. Long Paper (Acceptance rate of 23%).

A Convolutional Neural Network For Multi-scale Taxi Trajectory Prediction. Intelligent Information Fusion Lab@SCUT

Supervised by Prof. Jianming Lv, South China University of Technology *Sep. 2016 - Sep. 2017*

We propose TCONV which models trajectories as two-dimensional images and adopts multi-layer convolutional neural networks to combine multi-scale trajectory patterns to achieve precise prediction. Furthermore, we integrate multiple local enhancement convolutional fields to explore these important areas deeply for better prediction.

Comprehensive experiments based on real trajectory data show that T-CONV can achieve higher accuracy than the state-of-the-art methods.

The paper has been accepted by BigComp. Long Paper.

Hadoop Based Distributed Mobile Trajectory Retrieval System. New Media Lab@SCUT

Supervised by Prof. Jianming Lv, South China University of Technology *May. 2015 - Sep. 2016*

We develop a mobile trajectory retrieval system named TREST, which is based on the distributed Hadoop and HBase systems. TREST makes use of the horizontal expansion mechanism of Hadoop to store overwhelming spatio-temporal trajectories and supports frequent incremental insertion of data stream. Meanwhile, TREST maps the spatio-temporal features of trajectories into the simple key-value schema of HBase to support fast retrieval.

Experiments on this data set show that TREST can efficiently support both Single-track and All-track retrieval within milliseconds on average.

The paper has been accepted by DSC. Long Paper.

SCHOLARSHIPS / AWARDS

South China University of Technology Research Scholarships 2017, 2018, 2019
This scholarship is awarded to the outstanding students in the university.

Hongping And Changqing Scholarships 2016
This scholarship is awarded to the outstanding students with excellent performance in academic competitions.

Tencent Scholarship of the Science and Technology (Rank 1st) 2016

This scholarship is awarded to the outstanding students with excellent academic performance.

Anju Bao Scholarship of the Science and Technology (Rank 1st) 2016

This scholarship is awarded to the outstanding students with excellent academic performance.

South China University of Technology Scholarships 2015

This scholarship is awarded to the outstanding students in the university.

Merit Student of South China University of Technology 2015

This award is awarded to the outstanding students with excellent performance in academic and leadership.

Honorable Mention of Mathematical Contest In Modeling Certificate of Achievement 2015

Awarded by the Consortium for Mathematics and Its Application, USA.

Gold Prize in National COMAP's Computer Software Design Contest. 2015

Awarded by the China Computer Federation, China.

SERVICES

EMNLP-IJCNLP *Summarization and Generation Area* 2019

NLPCC *Conversational Bot/QA/IR/Dialogue Area* 2019

PROGRAMMING SKILLS

Computer Language:	**Python,** C++, Java
Development Software:	**Pycharm, Eclipse, Jupyter Notebook**
Database Software:	**MySQL, Oracle, HBase, Hive**
Toolkit for Data Analysis:	**Numpy, Pandas, Matplotlib, Seaborn, NLTK, Sklearn**
Framework for Deep Learning:	**PyTorch, MxNet, Tensorflow, Keras**
Framework for Distributed Computing:	**Hadoop, Spark**
Editing Software:	**Word, Excel, PowerPoint, Latex**