Ruotian Luo

CONTACT INFORMATION

E-mail: rluo@ttic.edu
Telephone: +1 312-608-0749
github: https://github.com/ruotianluo
Homepage: http://ttic.uchicago.edu/~rluo/

Address:
1607 E 50th Pl, APT 6C
Chicago, IL
60615

RESEARCH INTERESTS

Vision + Language, Computer Vision, Natural Language Processing, Deep Learning, Artificial Intelligence

EDUCATION

Toyota Technological Institute at Chicago, Chicago, IL. **September, 2015 - Present**

- Ph.D. in Computer Science (Expected to graduate in Summer 2021), GPA: 3.96/4
- Thesis title: Goal-driven text descriptions for images.
- Advisor: Greg Shakhnarovich

Shanghai Jiao Tong University, Shanghai, China. **September, 2011 - June, 2015**

- B.Eng. in Computer Science (IEEE Honor Class)
- Overall GPA: **89.43/100** (Rank 5 of 97)

HONORS AND AWARDS

ICCV'19, COCO+Mapillary Challenge, COCO Panoptic Segmentation Track: **Innovation Award**.

CVPR'19, Conceptual Captions Challenge: **Winning team** .

2014 The Interdisciplinary Contest in Modeling: **Finalist** (Top 11/1028)

XinDong Scholarship (Second-Class), XinDong Co. Ltd., 2013 (Top 5%)

Academic Excellence Scholarship (Second-Class), Shanghai Jiao Tong University, 2012-2014 (Top 10%)

National Scholarship, Ministry of Education, 2012 (Top 4%)

RESEARCH EXPERIENCE

Research Intern, Microsoft Multimodal AI Team **Summer, 2020**
Collaborators: Zhe Gan(mentor), Yen-Chun Chen, Luowei Zhou, Jingjing Liu

- Learning visual representation from image-text data in the wild with image-text matching as proxy task.

Research Intern, Snap Research **Summer, 2018**
Mentors: Linjie Yang, Ning Zhang, Bohyung Han

- Zero-shot object recognition: using knowledge graph and geometric context for better recognizing objects of unseen categories (published at AAAI 2020).

Research Intern, Adobe Research **Summer, 2017**
Mentors: Scott Cohen, Brian Price

- Discriminative Image Captioning: incorporated a discriminability objective into captioning training loss to generate more discriminative and more descriptive image captions (published at CVPR 2018).

Research Assistant, Toyota Technological Institute at Chicago **September, 2015 - Present**
Advisor: Greg Shakhnarovich

- Informative Image tagging: generate image tags that convey the most information about an image. (Collaboration with UChicago and Bar-Ilan University. In submission.)

- Controllable image captioning: designed several models that can generate image captions of designated lengths (presented at CVPR2020 VQA Workshop).
- Panoptic segmentation: a panoptic segmentation framework based on Generalized Hough Transform; each pixel votes for the center of the instance it belongs to (published at CVPR 2020).
- Diverse image captioning: analyzed the role of sampling methods and training objectives in generating diverse and accurate image captions; proposed a new metric AllSPICE (presented at ICCV2019 CLVL Workshop).
- Referring expressions generation: proposed two approaches to utilize models trained for referring expression comprehension task to generate better referring expressions (published at CVPR 2017).

Research Intern, University of Ottawa **Summer, 2014**
Advisor: Robert Laganière

- Project: time how long each customer waits in a fast food restaurant, from paying the money to getting the food.
- Accelerated an existing upper-body detection algorithm with CPU and GPU parallel methods.
- Applied SDALF, a re-id algorithm, to re-identify customers, and evaluated it on our fast food restaurant dataset.

Research Assistant, Shanghai Jiao Tong University **September, 2013 - June, 2015**
Advisor: Yuncai Liu

- Person re-identification: applied multi-view pictorial structures methods to jointly predict pedestrian poses and viewpoints to improve person re-id results.
- 3D pose reconstruction: used pictorial structures based 2D human pose estimation as a preprocessing for 3D pose reconstruction.

Publications

Luo, R. and Shakhnarovich, G., 2020. Controlling Length in Image Captioning. In CVPR 2020 VQA Workshop.

Wang, H., **Luo, R.**, Maire, M. and Shakhnarovich, G., 2020. Pixel Consensus Voting for Panoptic Segmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 9464-9473).

Luo, R., Zhang, N., Han, B. and Yang, L., 2020. Context-Aware Zero-Shot Recognition. In AAAI (pp. 11709-11716).

Luo, R. and Shakhnarovich, G., 2019. Analysis of diversity-accuracy tradeoff in image captioning. In ICCV19 CLVL Workshop.

Vasiljevic, I., Kolkin, N., Zhang, S., **Luo, R.**, Wang, H., Dai, F.Z., Daniele, A.F., Mostajabi, M., Basart, S., Walter, M.R. and Shakhnarovich, G., 2019. DIODE: A Dense Indoor and Outdoor DEpth Dataset. In CVPR19 3D Scene Understanding Workshop.

Luo, R., Price, B., Cohen, S. and Shakhnarovich, G., 2018, March. Discriminability objective for training descriptive captions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 6964-6974).

Zhao, W., Wang, B., Ye, J., Yang, M., Zhao, Z., **Luo, R.** and Qiao, Y., 2018. A Multi-task Learning Approach for Image Captioning. In IJCAI (pp. 1205-1211).

Luo, R. and Shakhnarovich, G., 2017. Comprehension-guided referring expressions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 7102-7111).

Zhao, Y., Zhao, X., **Luo, R.** and Liu, Y., 2016. Person Re-identification by encoding free energy

feature maps. Multimedia Tools and Applications, 75(8), pp.4795-4813.

Liang, S., **Luo, R.**, Chen, G., Ma, S., Wu, W., Song, L., Tian, X. and Wang, X., 2014, December. Are we still friends: Kernel multivariate survival analysis. In 2014 IEEE Global Communications Conference (pp. 405-410). IEEE.

Preprints

Gilton, D., **Luo, R.**, Willett, R. and Shakhnarovich, G., 2020. Detection and Description of Change in Visual Streams.

Luo, R., 2020. A Better Variant of Self-Critical Sequence Training.

Teachings

Teaching Assistant **September, 2016 - December, 2016**
TTIC 31020, Introduction to Statistical Machine Learning, Autumn 2016.

Services

Reviewer/Program Committee
TPAMI, ACL2021, ICML2021, CVPR2021, WACV2021, ACMMM2020, Neurips2020, EMNLP2020, ECCV2020, CVPR2020, ICCV2019, ICML-HIL2019, NAACL2019, CVPR 2019

Patents

Discriminative Caption Generation, US2019/0377987 A1 **June, 2018**
Inventors: Brian Lynn Price, **Ruotian Luo**, Scott David Cohen

Intelligent taxi scheduling system, CN103680128A **January, 2014**
Inventors: Tianyuan Liu, **Ruotian Luo**, Yang Zhang, Feng Yang, Xiaoying Gan, Xiaohua Tian, Xinbing Wang

Projects

Daily Arxiv Radiostation, *Core Member* **October, 2019 - Present**
A daily podcast on cutting-edge research papers of computer science, generated and published automatically.

Lane tracking with transfer learning in Duckietown, *Core Member* **May, 2018**
Trained a CNN which takes images from the front camera of a DuckieBot and predicts the angle and distance to centerline. The model is first trained on simulated data with domain randomization and finetuned on 72 real images data.

Faster R-CNN in PyTorch, *Individual developer* **August, 2017 - October, 2017**
A PyTorch implementation of Faster R-CNN: the first at the time that can achieve comparable performance (with training) compared to codebases of other frameworks. (3.5k stars on github)

Campus news recommendation system, *Core Member* **March, 2013 - July, 2013**
Developed an Android APP which crawls the news from campus website or BBS etc., and recommends news of users' interests.

IAmHungry GO AI, *Team leader* **December, 2013-January, 2014**
Developed a computer go AI based on UCT (Upper Confidence bounds applied to Trees) and Monte Carlo Tree Search.

Object tracking system on Smart Robot Car, *Core Member* **July, 2014**
Built a tracking system on Smart Robot Car based on TLD(Tracking-Learning-Detection) algorithm.

Skills

- Languages: Python, C++, C, Lua, MATLAB, Java, PASCAL, LaTeX, SQL, PHP.
- Deep Learning Frameworks: PyTorch, Tensorflow, Torch7.