Xiaolong Wang

412-2655517 • xiaolonw@cs.cmu.edu • www.cs.cmu.edu/~xiaolonw

Education

Carnegie Mellon University — Pittsburgh, PA, USA
Ph.D. Student in Robotics — *2014 - present*
Advisor: Prof. Abhinav Gupta
Awarded Facebook Fellowship, Baidu Fellowship and Nvidia Fellowship

Sun Yat-Sen University — Guangzhou, China
M.S. in Computer Science — *2011 - 2014*
Advisor: Prof. Liang Lin

South China Agricultural University — Guangzhou, China
B.S. in Computer Science — *2007 - 2011*
Advisors in ACM-ICPC: Prof. Caixing Liu and Prof. Xiangji Chen

Academic Experience

Berkeley AI Research (BAIR)
Visitor (Supervisor: Prof. Alexei A. Efros) — *Jun 2018 - Nov 2018*

Facebook AI Research (FAIR)
Intern (Supervisor: Dr. Kaiming He and Dr. Ross Girshick) — *May 2017 - Nov 2017*

Allen Institute for Artificial Intelligence (AI2)
Intern (Supervisor: Prof. Ali Farhadi) — *May 2015 - Aug 2015*

Institute of Deep Learning (IDL) in Baidu, Inc.
Intern (Supervisor: Dr. Kai Yu and Dr. Yinan Yu) — *Jul 2013 - Jan 2014*

Selected Awards

- Facebook Fellowship — *2018*
- Baidu Fellowship — *2018*
- Nvidia Fellowship — *2017*
- Best Student Paper Award of ICME — *2014*
- Google Scholarship — *2013*
- ACM-ICPC World Finals, Honorable Mention — *2010*
- ACM-ICPC Asia Regional Programming Contest (1 Gold, 2 Silver, 1 Bronze Medals) — *2008 - 2010*

Publications

Peer-Reviewed Conference Publications

[1] **Xiaolong Wang***, Allan Jabri* and Alexei A. Efros
Learning Correspondence from the Cycle-consistency of Time
in Computer Vision and Pattern Recognition (CVPR), 2019.
(**Oral Presentation**)

[2] Xueting Li, Sifei Liu, Kihwan Kim, **Xiaolong Wang**, Ming-Hsuan Yang, and Jan Kautz
Putting Humans in a Scene: Learning Affordance in 3D Indoor Environments
in Computer Vision and Pattern Recognition (CVPR), 2019.

[3] Wei Yang, **Xiaolong Wang**, Ali Farhadi, Abhinav Gupta, and Roozbeh Mottaghi
Visual Semantic Navigation using Scene Priors
in International Conference on Learning Representations (ICLR), 2019.

[4] **Xiaolong Wang** and Abhinav Gupta
Videos as Space-Time Region Graphs
in European Conference on Computer Vision (ECCV), 2018.

[5] Tian Ye, **Xiaolong Wang**, James Davidson, and Abhinav Gupta
Interpretable Intuitive Physics Model
in European Conference on Computer Vision (ECCV), 2018.

[6] **Xiaolong Wang**, Ross Girshick, Abhinav Gupta, and Kaiming He
Non-local Neural Networks
in Computer Vision and Pattern Recognition (CVPR), 2018.

[7] **Xiaolong Wang***, Yufei Ye*, and Abhinav Gupta
Zero-shot Recognition via Semantic Embeddings and Knowledge Graphs
in Computer Vision and Pattern Recognition (CVPR), 2018.

[8] Wei Yang , Wanli Ouyang, **Xiaolong Wang**, Jimmy Ren, Hongsheng Li, and Xiaogang Wang
3D Human Pose Estimation in the Wild by Adversarial Learning
in Computer Vision and Pattern Recognition (CVPR), 2018.

[9] **Xiaolong Wang**, Kaiming He, and Abhinav Gupta
Transitive Invariance for Self-supervised Visual Representation Learning
in International Conference on Computer Vision (ICCV), 2017.

[10] Yuan Yuan, Xiaodan Liang, **Xiaolong Wang**, Dit-Yan Yeung, and Abhinav Gupta
Temporal Dynamic Graph LSTM for Action-driven Video Object Detection
in International Conference on Computer Vision (ICCV), 2017.

[11] **Xiaolong Wang***, Rohit Girdhar*, and Abhinav Gupta
Binge Watching: Scaling Affordance Learning from Sitcoms
in Computer Vision and Pattern Recognition (CVPR), 2017.
(**Spotlight Oral Presentation**)

[12] **Xiaolong Wang**, Abhinav Shrivastava, and Abhinav Gupta
A-Fast-RCNN: Hard Positive Generation via Adversary for Object Detection
in Computer Vision and Pattern Recognition (CVPR), 2017.

[13] **Xiaolong Wang** and Abhinav Gupta
Generative Image Modeling using Style and Structure Adversarial Networks
in European Conference on Computer Vision (ECCV), 2016.

[14] Gunnar A. Sigurdsson, Gül Varol, **Xiaolong Wang**, Ivan Laptev, Ali Farhadi, and Abhinav Gupta
Hollywood in Homes: Crowdsourcing Data Collection for Activity Understanding
in European Conference on Computer Vision (ECCV), 2016.

[15] **Xiaolong Wang**, Ali Farhadi, and Abhinav Gupta
Actions $\sim$ *Transformations*
in Computer Vision and Pattern Recognition (CVPR), 2016.

[16] **Xiaolong Wang** and Abhinav Gupta
Unsupervised Learning of Visual Representations using Videos
in International Conference on Computer Vision (ICCV), 2015.

[17] **Xiaolong Wang**, David F. Fouhey, and Abhinav Gupta
Designing Deep Networks for Surface Normal Estimation
in Computer Vision and Pattern Recognition (CVPR), 2015.

[18] **Xiaolong Wang**, Liliang Zhang, Liang Lin, Zhujin Liang, and Wangmeng Zuo
Deep Joint Task Learning for Generic Object Extraction
in Neural Information Processing Systems (NIPS), 2014.

[19] Keze Wang, **Xiaolong Wang**, and Liang Lin
Deep Structured Models for 3D Human Activity Recognition
in ACM International Conference on Multimedia (MM), 2014.
(**Full Paper, Oral Presentation**)

[20] Zhujin Liang, **Xiaolong Wang**, Rui Huang, and Liang Lin
An Expressive Deep Model for Parsing Human Action from a Single Image
in IEEE International Conference on Multimedia and Expo (ICME), 2014.
(**Oral Presentation, Best Student Paper Award**)

[21] **Xiaolong Wang**, Liang Lin, Lichao Huang, and Shuicheng Yan
Incorporating Structural Alternatives and Sharing into Hierarchy for Multiclass Object Recognition and Detection
in Computer Vision and Pattern Recognition (CVPR), 2013.

[22] **Xiaolong Wang** and Liang Lin
Dynamical And-Or Graph Learning for Object Shape Modeling and Detection
in Neural Information Processing Systems (NIPS), 2012.

[23] Liang Lin, **Xiaolong Wang**, Wei Yang, and Jian-Huang Lai
Learning Contour-Fragment-based Shape Model with And-Or Tree Representation
in Computer Vision and Pattern Recognition (CVPR), 2012.

[24] Wei Yang, **Xiaolong Wang**, Liang Lin, Chengying Gao
Interactive CT image segmentation with online discriminative learning
in International Conference on Image Processing (ICIP), 2011.

Peer-Reviewed Journal Publications

[25] Liang Lin, **Xiaolong Wang**, Wei Yang, and JianHuang Lai
Discriminatively Trained And-Or Graph Models for Object Shape Detection
in Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2015.

Technical Reports

[26] David F. Fouhey, **Xiaolong Wang**, and Abhinav Gupta
In Defense of the Direct Perception of Affordances
in arXiv, 2015.

Talks

- *Learning and Reasoning with Visual Correspondence in Time*
  University of California San Diego — *Mar 2019*
  University of Illinois at Urbana-Champaign — *Mar 2019*
  University of Massachusetts Amherst — *Mar 2019*
- *Exploiting Redundancy for Learning Visual Representations*
  Invited Talk in University of California, Los Angeles — *Oct 2018*
- *Looking into Recognition in the Deep Era*
  Computer Vision Seminar, University of California, Berkeley — *Oct 2018*
- *Adversaries for Detection and Action*
  CVPR 2018 Tutorial on GANs — *Jun 2018*
- *Videos as Space-Time Region Graphs*
  CVPR 2018 Workshop on Fine-grained Instructional Video undERstanding (FIVER) — *Jun 2018*
- *Exploiting Redundancy for Learning Visual Representations*
  Invited Talk in University of California, Berkeley — *May 2018*
  Invited Talk in The Jiangmen — *May 2018*
  PhD Thesis Proposal, Carnegie Mellon University — *Feb 2018*

- *Non-local Neural Networks*
  Grad Fellow FastForward, NVIDIA's GPU Technology Conference (GTC) *Mar 2018*
- *Learning Visual Representations for Object Detection*
  AI Seminar sponsored by Apple *Oct 2017*
- *Binge Watching: Scaling Affordance Learning from Sitcoms*
  Spotlight Presentation, CVPR *Jul 2017*
- *Actions ∼ Transformations*
  Allen Institute for Artificial Intelligence (AI2) *Aug 2016*
- *Unsupervised Learning of Visual Representations using Videos*
  PhD Speaking Qualifier, Carnegie Mellon University *Apr 2016*
- *Designing Deep Networks for Surface Normal Estimation*
  Mid-Atlantic Computer Vision (MACV) Workshop *Mar 2015*

Academic Services

Teaching Assistant

Visual Learning and Recognition (CMU 16-824) *2016*

Journal Reviewer

IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)
International Journal of Computer Vision (IJCV)
IEEE Transactions on Circuits and Systems for Video Technology (TCSVT)
IEEE Transactions on Neural Networks and Learning Systems (TNNLS)

Conference Reviewer

Computer Vision and Pattern Recognition (CVPR)
European Conference on Computer Vision (ECCV)
International Conference on Computer Vision (ICCV)
International Conference on Robotics and Automation (ICRA)
Conference on Robot Learning (CoRL)

Open Source

Codes and models on GitHub: https://github.com/xiaolonw