

# Hao Li

4431 A. V. Williams Building  
College Park, MD 20742

☎ +1 (240) 485-7596  
✉ [haoli@cs.umd.edu](mailto:haoli@cs.umd.edu)  
🌐 <http://cs.umd.edu/~haoli>  
🔗 <http://github.com/ljk628>

## Education

- University of Maryland, College Park** College Park, MD  
*Ph.D. in Computer Science* 2012 – 2018
- Advisors: Prof. Hanan Samet and Prof. Tom Goldstein
  - Thesis: Towards Fast and Efficient Representation Learning
- Institute of Computing Technology, Chinese Academy of Sciences** Beijing, China  
*M.S. in Computer Science* 2009 – 2012
- Advisor: Prof. Shiguang Shan
- Shandong University** Jinan, China  
*B.Eng. in Software Engineering* 2005 – 2009

## Research Interests

Machine Learning, Deep Learning, Computer Vision, HPC, Distributed System

## Experience

- University of Maryland Institute for Advanced Computer Studies** College Park, MD  
*Research Assistant with Prof. Hanan Samet and Prof. Tom Goldstein* May 2013 - present
- Visualizing the loss surface of neural networks [NIPS'18]
  - Optimization methods for training quantized neural networks [NIPS'17].
  - Accelerating Convolutional Neural Networks for fast inference [ICLR'17].
  - Distributed machine learning with HPC [EuroSys'15, ICML'17].
- Machine Learning Team, Nervana Systems** San Diego, CA  
*Algorithms Engineer Intern with Scott Cyphers and Arjun Bansal* May - Aug 2016
- Design and implement the importer module for Neon's graph-backend (Nervana Graph), enabling executing the computation graph of TensorFlow with nGraph.
- Machine Learning Department, NEC Labs America** Princeton, NJ  
*Research Intern/Visiting Student* Jan - Aug'15/Jan - May'16
- Research on the architecture design of Convolutional Neural Networks.
  - Research on filter pruning for acceleration and compression of CNNs [ICLR'17].
- Storage System Department, NEC Labs America** Princeton, NJ  
*Research Intern with Asim Kadav, Erik Kruus and Cristian Ungureanu* May – Aug 2014
- Built MALT, a distributed data-parallel machine learning system with flexible peer-to-peer communication, utilizing one-side RDMA write over InfiniBand [EuroSys'15].
- Institute of Computing Technology, Chinese Academy of Sciences** Beijing, China  
*Research Assistant with Prof. Shiguan Shan* Jan 2011 - July 2012
- Multi-person tracking with re-identification in non-overlapping camera networks [ACCV'14].
- Internet Graphics Group, Microsoft Research Asia** Beijing, China  
*Research Intern with Ying-Qing Xu* May – Nov 2010

- Bing image search result beautification.

**Media Computing Group, Microsoft Research Asia**

Research Intern with Meng Wang

Beijing, China

Mar – July 2009

- Research on graph-based reranking for web image search [T-IP'12, SIGIR'11].
- Built a million-scale image/video dataset for web image/video search reranking [ICDMW'09].

## Publications

1. *Visualizing the Loss Landscape of Neural Nets*  
**Hao Li**, Zheng Xu, Gavin Taylor, Tom Goldstein  
Annual Conference on Neural Information Processing Systems (**NIPS**), 2018
2. *Training Quantized Nets: A Deeper Understanding*  
**Hao Li\***, Soham De\*, Zheng Xu, Christoph Studer, Hanan Samet, Tom Goldstein  
Annual Conference on Neural Information Processing Systems (**NIPS**), 2017
3. *Towards a Deeper Understanding of Training Quantized Neural Networks*  
**Hao Li\***, Soham De\*, Zheng Xu, Christoph Studer, Hanan Samet, Tom Goldstein  
ICML Workshop on Principled Approaches to Deep Learning, 2017 (**Google Best Student Paper**)
4. *Pruning Filters for Efficient ConvNets*  
**Hao Li**, Asim Kadav, Igor Durdanovic, Hanan Samet, Hans Peter Graf  
International Conference on Learning Representations (**ICLR**), 2017  
NIPS Workshop on Efficient Methods for Deep Neural Networks (**EMDNN**), 2016
5. *Adaptive Consensus ADMM for Distributed Optimization*  
Zheng Xu, Gavin Taylor, **Hao Li**, Mario Figueiredo, Xiaoming Yuan, and Tom Goldstein  
International Conference on Machine Learning (**ICML**), 2017
6. *MALT: Distributed Data-Parallelism for Existing ML Applications*  
**Hao Li**, Asim Kadav, Erik Kruus, Cristian Ungureanu  
ACM European Conference on Computer Systems (**EuroSys**), 2015  
NIPS Workshop on Machine Learning Systems (**LearningSys**), 2015
7. *Streaming News Image Summarization*  
**Hao Li**, Shangfu Peng, Hanan Samet  
International Conference on Pattern Recognition (**ICPR**), 2016 (Oral)
8. *Simplification and Refinement for Speedy Spatio-temporal Hot Spot Detection Using Spark (GIS Cup)*.  
Shangfu Peng, Hong Wei, **Hao Li**, and Hanan Samet  
ACM Int. Conf. on Advances in Geographic Information Systems (**SIGSPATIAL**), 2016 (Short)
9. *View Adaptive Metric Learning for Multi-view Person Re-identification*  
Canxiang Yang, Shiguang Shan, Dan Wang, **Hao Li**, Xilin Chen  
Asian Conference on Computer Vision (**ACCV**), 2014
10. *Multimodal Graph-Based Reranking for Web Image Search*  
Meng Wang, **Hao Li**, Dacheng Tao, Ke Lu, Xindong Wu  
IEEE Transactions on Image Processing (**T-IP**), 2012
11. *Optimizing Multimodal Reranking for Web Image Search*  
**Hao Li**, Meng Wang, Zhisheng Li, Zheng-Jun Zha, Jialie Shen  
ACM SIGIR Conf. on Research and Development in Information Retrieval (**SIGIR**), 2011 (Short)
12. *MSRA-MM 2.0: A Large-Scale Web Multimedia Dataset*  
**Hao Li**, Meng Wang, Xian-Sheng Hua  
IEEE ICDM Workshop on Internet Multimedia Mining, 2009

## Awards

Google Best Student Paper Award, ICML PADL Workshop	2017
Conference Travel Awards: NIPS'17, ICML'17, ICLR'17, EuroSys'15, SIGIR'11	
Jacob K. Goldhaber Travel Grant, University of Maryland	2016
John D. Gannon Student Travel Grant, University of Maryland	2013
Dean's Fellowship, Department of Computer Science, University of Maryland	2012, 2013
Stars of Tomorrow Excellent Intern Award, Microsoft Research Asia	2010
Outstanding College Graduate, Shandong Province and Shandong University	2009
Meritorious Winner in Mathematical Contest in Modeling, USA	2008
National Second Prize in China Undergraduate Mathematical Contest in Modeling	2007

## Academic Services

Reviewer for IEEE Transactions on Pattern Analysis and Machine Intelligence	2018
Reviewer for IEEE Transactions on Neural Network and Learning Systems	2018
Reviewer for IEEE Transaction on Industrial Informatics	2015
Reviewer for Neurocomputing	2015
Reviewer for Multimedia Systems	2014
Reviewer for International Conference on Learning Representation (ICLR)	2019
Reviewer for Neural Information Processing Systems (NIPS)	2016, 2018
Reviewer for AAAI Conference on Artificial Intelligence (AAAI)	2019
Shadow PC for European Conference on Computer Systems (EuroSys)	2018
Reviewer for NIPS Workshop on Efficient Methods for Deep Neural Networks (EMDNN)	2016
Reviewer for IEEE International Conference on Multimedia and Expo (ICME)	2013, 2014
External Reviewer for ACM Conf. on Information and Knowledge Management (CIKM)	2014

## Talks

### *Towards Fast and Efficient Representation Learning*

Facebook Research, Menlo Park, CA, May, 2018; IBM Research, Cambridge, MA, May, 2018; Amazon AWS AI, Seattle, WA, April, 2018; Apple AI Research, Cupertino, CA, April, 2018; Vicarious, Union City, CA, April, 2018; Comcast Labs, Washington D.C., Mar, 2018; Microsoft Research Asia, Beijing, Jan. 2018; NVIDIA Research, Santa Clara, CA, Nov. 2017

### *Towards a Deeper Understanding of Training Quantized Neural Networks*

ICML'17 Workshop on Principled Approaches on Deep Learning, Sydney, Aug. 2017

### *Towards Efficient ConvNets*

NEC Labs America, Princeton, NJ, Aug. 2015

### *MALT: Efficient Data Parallelism for Existing Machine Learning Software*

NEC Labs America, Princeton, NJ, Aug. 2014

## Programming Skills

Languages: Python, C/C++, Lua, Java, C#, Matlab, Go

Tools: PyTorch, Torch7, TensorFlow, Neon, Theano, MPI, GASPI, ZMQ, Hadoop

## Teaching Experience

CMSC 764 Advanced Numerical Optimization, with Prof. Tom Goldstein	2017
CMSC 828Z Reinforcement Learning, with Prof. Hal Daumé III	2016
CMSC 216 Introduction to Computer Systems, with Dr. Larry Herman	2015
CMSC 426 Programming Handheld Systems, with Prof. Atif Memon	2013
CMSC 330 The Organization of Programming Languages, with Prof. Chau-Wen Tseng	2012

## Patents

Parallelized machine learning with distributed lockless training. US Patent 9,984,337  
Passive Pruning of Filters in a Convolutional Neural Network. US Patent App. 15/590,620  
Security System Using a Conv. Neural Network with Pruned Filters. US Patent App. 15/590,666  
Parallelized Machine Learning With Distributed Lockless Training. US Patent App. 14/875,773  
MALT: Distributed Data-Parallelism for Existing ML Applications. US Patent App. 14/872,521

## References

Prof. Hanan Samet  
Dept. of Computer Science  
University of Maryland  
hjs@cs.umd.edu

Prof. Tom Goldstein  
Dept. of Computer Science  
University of Maryland  
tomg@cs.umd.edu

Dr. Hans Peter Graf  
Machine Learning Dept.  
NEC Labs America  
hpg@nec-labs.com