
Chunyuan Li

Senior Researcher
Microsoft Research, Redmond

Office: Room 3840, Microsoft Building 99
Email: chunyuan.li@hotmail.com
Homepage: <http://chunyuan.li>

Research Interests

My recent research focuses on deep learning and probabilistic modeling with large-scale datasets/training. These tools have been applied to various domains, including natural language processing, computer vision and deep reinforcement learning etc.

Education

- **Duke University**, Durham, NC 2014 - 2018
Ph.D., Electrical and Computer Engineering, GPA: 3.9/4.0.
Bayesian and Modern Statistics Statistical Computation
Probabilistic & Advanced Machine Learning Information Theory
Discrete Optimization Graphical Models & Inference
- **Concordia University**, Montreal, Canada 09/2011 - 2013
M.S., Quality System Engineering, University Merit Award (top 0.5%)
- **Huazhong University of Science and Technology**, Wuhan, China 09/2007 - 06/2011
B.E., Electronics and Information Engineering, Excellent Undergraduate Thesis (top 1%)

Research Experiences

- **Microsoft Research**, Redmond, WA 2018 - present
Senior Researcher, deep learning team at MSR AI
Self-supervised Representation Learning & Generative Pre-training Models via Large-scale Data & Training, with demonstrated project experience/papers in natural language modeling, image generation, vision-and-language, dialog tasks.
- **Information Initiative at Duke (iiD)** 2014 - 2018
Research assistant. Advisor: Prof. Lawrence Carin
(i) **Bayesian Deep Learning**: Scalable Bayesian learning methods for the weight uncertainty of deep neural networks, e.g., SG-MCMCs
(ii) **Deep Bayesian Learning**: Deep neural networks as flexible representation methods in Bayesian models, e.g., GANs and VAEs.
- **Uber AI Labs**, San Francisco, CA Summer, 2017
Research Scientist Intern. Mentor: Jason Yosinski
Subspace training of neural networks; one paper & patent
- **Adobe Research**, San Jose, CA Summer, 2016
Data Scientist Intern. Mentors: H. Bui, M. Ghavamzadeh and G. Theocharous
(i) Product: Recurrent neural networks for digital market forecasting; one patent
(ii) Science: Investigation of Bayesian deep reinforcement learning
- **National Institute of Standards and Technology**, MD 09/2013 - 08/2014
Organized and participated shape retrieval contests in Eurographics 2014
- **Geometrica Group of INRIA Saclay**, France Summer, 2013
Research intern. Mentors: Maks Ovsjanikov and Frederic Chazal
Developed algorithms for object recognition via topological persistence

- Concordia University, Montreal, Canada 09/2011 - 04/2013
Deformable 3D shape analysis via spectral geometry
- Ankon International Summer, 2011
Developed novel online redundant image elimination algorithms for wireless capsule endoscopy
- Huazhong University of Science and Technology 2009 - 2011
Worked on algorithms for 2D shape analysis and classification

Publications [Citations=3041, h-index=31, i10-index=45]

Preprint

1. B. Peng, C. Li, J. Li, S. Shayandeh, L. Liden, J. Gao
Soloist: Few-shot Task-Oriented Dialog with A Single Pre-trained Auto-regressive Model

Selected Journal and Conference

1. C. Li, X. Gao, Y. Li, X. Li, B. Peng, Y. Zhang, J. Gao
Optimus: Organizing Sentences via Pre-trained Modeling of a Latent Space
Empirical Methods in Natural Language Processing (EMNLP) 2020
2. B. Peng, C. Zhu, C. Li, X. Li, J. Li, M. Zeng, and J. Gao
Few-shot Natural Language Generation for Task-Oriented Dialogue
Empirical Methods in Natural Language Processing (EMNLP) 2020
3. Y. Zhang*, G. Wang*, C. Li, Z. Gan, C. Brockett, B. Dolan
Pointer: Constrained Text Generation via Insertion-based Generative Pre-training
Empirical Methods in Natural Language Processing (EMNLP) 2020
4. B. An, J. Lyu, Z. Wang, C. Li, C. Hu, F. Tan, R. Zhang, Y. Hu and C. Chen
Repulsive Attention: Rethinking Multi-head Attention as Bayesian Inference
Empirical Methods in Natural Language Processing (EMNLP) 2020
5. J. Li*, C. Li*, G. Wang*, H. Fu, Y. Lin, L. Chen, Y. Zhang et al
Improving Text Generation with Student-Forcing Optimal Transport
Empirical Methods in Natural Language Processing (EMNLP) 2020
6. S. Yuan, K. Bai, L. Chen, Y. Zhang, C. Tao, C. Li, G. Wang, R. Henao, L. Carin
Weakly Supervised Cross-domain Alignment with Optimal Transport
British Machine Vision Conference (BMVC) 2020 **Oral Presentation**
7. X. Li, X. Yin, C. Li, P. Zhang, X. Hu, L. Zhang, H. Hu, L. Dong, F. Wei, Y. Choi and J. Gao
Oscar: Object-Semantics Aligned Pre-training for Vision-Language Tasks
European Conference on Computer Vision (ECCV) 2020
8. P. Yu, Y. Zhao, C. Li, J. Yuan and C. Chen
Structure-Aware Human-Action Generation
European Conference on Computer Vision (ECCV) 2020
9. Y. Zhao*, C. Li*, J. Gao, and C. Chen (* Equal contribution)
Feature Quantization Improves GAN Training
International Conference on Machine Learning (ICML) 2020
10. W. Hao*, C. Li*, X. Li, L. Carin and J. Gao (* Equal contribution)
Towards Learning a Generic Agent for Vision-and-Language Navigation via Pre-training
Conference on Computer Vision and Pattern Recognition (CVPR) 2020
11. R. Zhang, C. Li, J. Zhang, C. Chen, and A. G. Wilson
Cyclical Stochastic Gradient MCMC for Bayesian Deep Learning
International Conference on Learning Representations (ICLR) 2020
Oral Presentation, acceptance rate 1.8%
12. S. Lobel*, C. Li*, J. Gao, and L. Carin (* Equal contribution)
Towards Amortized Ranking-Critical Training for Collaborative Filtering
International Conference on Learning Representations (ICLR) 2020

13. Y. Li, C. Li[†], Y. Zhang, X. Li, G. Zheng, L. Carin and J. Gao ([†] Corresponding author)
Complementary Auxiliary Classifiers for Label-Conditional Text Generation
AAAI Conference on Artificial Intelligence (AAAI) 2020
14. P. Chapfuwa, C. Li, N. Mehta, L. Carin, and R. Henao
Survival Cluster Analysis
ACM Conference on Health, Inference, and Learning (CHIL) 2020
15. M Gong, Y Xu, C. Li, K Zhang, K Batmanghelich
Twin Auxiliary Classifiers GAN
Neural Information Processing Systems (NeurIPS) 2019
Spotlight Presentation, acceptance rate 2.4%
16. Robust Navigation with Language Pre-training and Stochastic Sampling
X Li, C. Li, Q Xia, Y Bisk, A Celikyilmaz, J Gao, NA Smith, Y Choi
Empirical Methods in Natural Language Processing (EMNLP) 2019
17. L Fang, C. Li, J Gao, W Dong, C Chen
Implicit Deep Latent Variable Models for Text Generation
Empirical Methods in Natural Language Processing (EMNLP) 2019
18. H, Fu*, C. Li*, X. Liu, J. Gao, A. Celikyilmaz, and L. Carin (* Equal contribution)
"Cyclical Annealing Schedule: A Simple Approach to Mitigating KL Vanishing"
North American Association for Computational Linguistics (NAACL) 2019, **Oral Presentation**
19. C. Li, K. Bai, J. Li, G. Wang, C. Chen, and L. Carin
"Adversarial Learning of a Sampler Based on an Unnormalized Distribution"
Artificial Intelligence and Statistics (AISTATS) 2019
20. C. Li, C. Chen, Y. Pu, R. Henao and L. Carin
"Communication-efficient Stochastic Gradient MCMC for Neural Networks"
AAAI Conference on Artificial Intelligence (AAAI) 2019
21. C. Li, H. Farkhoor, R. Liu and J. Yosinski
"Measuring the Intrinsic Dimension of Objective Landscapes"
International Conference on Learning Representations (ICLR) 2018
22. C. Chen, C. Li, L. Chen, W. Wang, Y. Pu and L. Carin
"Continuous-Time Flows for Efficient Inference and Density Estimation"
International Conference on Machine Learning (ICML) 2018
23. R. Zhang, C. Chen, C. Li, and L. Carin
"Policy Optimization as Wasserstein Gradient Flows"
International Conference on Machine Learning (ICML) 2018
24. P. Chapfuwa, C. Tao, C. Li, C. Page, B. Goldstein, L. Carin, R. Henao
"Adversarial Time-to-Event Modeling"
International Conference on Machine Learning (ICML) 2018
25. G. Wang, C. Li[†], W. Wang, Y. Zhang, D. Shen, and L. Carin ([†] Corresponding author)
"Joint Word and Label Embeddings for Text Classification"
Annual Meeting of the Association for Computational Linguistics (ACL) 2018
26. D. Shen, G. Wang, W. Wang, M. Min, Q. Su, Y. Zhang, C. Li, R. Henao and L. Carin
"On Simple Word-Embedding-Based Models and Associated Pooling Mechanisms"
Annual Meeting of the Association for Computational Linguistics (ACL) 2018
27. R. Zhang, C. Li, C. Chen, and L. Carin
"Learning Structural Weight Uncertainty for Sequential Decision-Making"
Artificial Intelligence and Statistics (AISTATS) 2018
28. L. Chen, S. Dai, Y. Pu, C. Li, Q. Su, and L. Carin
"Symmetric Variational Autoencoder and Connections to Adversarial Learning"
Artificial Intelligence and Statistics (AISTATS) 2018
29. J. Lu, C. Li, J. Singh-Alvarado, Z. Zhou, F. Frohlich, R. Mooney and F. Wang
"MIN1PIPE: A Miniscope 1-photon-based Calcium Imaging Signal Extraction Pipeline"
Cell Report 2018 (Impact factor: 8.282)

30. C. Li, H. Liu, C. Chen, Y. Pu, L. Chen, R. Henao and L. Carin
"ALICE: Towards Understanding Adversarial Training for Joint Distribution Matching"
Neural Information Processing Systems (NIPS) 2017
31. Y. Pu, Z. Gan, R. Henao, C. Li, S. Han and L. Carin
"VAE Learning via Stein Variational Gradient Descent"
Neural Information Processing Systems (NIPS) 2017
32. Y. Pu, W. Wang, R. Henao, L. Chen, Z. Gan, C. Li, and L. Carin
"Adversarial Symmetric Variational Autoencoder",
Neural Information Processing Systems (NIPS) 2017
33. Z. Gan, L. Chen, W. Wang, Y. Pu, Y. Zhang, H. Liu, C. Li, and L. Carin
"Triangle Generative Adversarial Networks",
Neural Information Processing Systems (NIPS) 2017
34. Z. Gan*, C. Li*, C. Chen, Q. Su, Y. Pu, and L. Carin (* Equal contribution)
"Scalable Bayesian Learning of Recurrent Neural Networks for Language Modeling"
Annual Meeting of the Association for Computational Linguistics (ACL) 2017, **Oral Presentation**
35. Z. Gan, Y. Pu, R. Henao, C. Li, X. He and L. Carin
"Learning Generic Sentence Representations using Convolutional Neural Networks"
Empirical Methods on Natural Language Processing (EMNLP) 2017, **Oral Presentation**
36. Q. Su, X. Liao, C. Li, and Z. Gan, L. Carin
"Restricted Truncated Gaussian Graphical Models"
AAAI Conference on Artificial Intelligence (AAAI) 2017, **Oral Presentation**
37. C. Li, A. Stevens, C. Chen, Y. Pu, Z. Gan and L. Carin
"Learning Weight Uncertainty with Stochastic Gradient MCMC for Shape Classification"
Computer Vision and Pattern Recognition (CVPR) 2016,
Spotlight Presentation, acceptance rate 9.7%
38. C. Li, C. Chen, D. Carlson and L. Carin
"Preconditioned Stochastic Gradient Langevin Dynamics for Deep Neural Networks"
AAAI Conference on Artificial Intelligence (AAAI) 2016, **Oral Presentation**
39. C. Li, C. Chen, K. Fan and L. Carin
"High-Order Stochastic Gradient Thermostats for Bayesian Learning of Deep Models"
AAAI Conference on Artificial Intelligence (AAAI) 2016
40. C. Chen, N. Ding, C. Li, Y. Zhang and L. Carin
"Stochastic Gradient MCMC with Stale Gradients"
Neural Information Processing Systems (NIPS) 2016
41. Y. Pu, Z. Gan, R. Henao, Y. Xin, C. Li, A Stevens, and L. Carin
"Variational Autoencoder for Deep Learning of Images, Labels and Captions"
Neural Information Processing Systems (NIPS) 2016
42. K. Fan, C. Li, and K. Heller
"Hierarchical Graph-Coupled HMM with an Application to Influenza Infection"
AAAI Conference on Artificial Intelligence (AAAI) 2016
43. Y. Zhang, R. Henao, C. Li and L. Carin
"Bayesian Dictionary Learning with Gaussian Processes and Sigmoid Belief Networks"
Int. Joint Conference on Artificial Intelligence (IJCAI) 2016
44. C. Chen, D. Carlson, Z. Gan, C. Li and L. Carin
"Bridging the Gap Between Stochastic Gradient MCMC and Stochastic Optimization"
Artificial Intelligence and Statistics (AISTATS) 2016,
Oral Presentation, acceptance rate 6.5%
45. Y. Pu, X. Yuan, A. Stevens, C. Li and L. Carin
"A Deep Generative Deconvolutional Image Model"
Artificial Intelligence and Statistics (AISTATS) 2016
46. D. Pickup, X. Sun, P. L. Rosin, R. R. Martin, C. Li *et al.*
"Shape Retrieval of Non-Rigid 3D Human Models",
Int. Journal of Computer Vision (IJCV) 2016

47. Z. Gan, **C. Li**, R. Henao, D. Carlson and L. Carin
"Deep Temporal Sigmoid Belief Networks for Sequence Modeling",
Neural Information Processing Systems (NIPS) 2015
48. B. Li, Y. Lu, **C. Li**, A. Godil, T. Schreck, *et al.*
"A Comparison of 3D Shape Retrieval Methods: A Benchmark with Multimodal Queries",
Computer Vision and Image Understanding (CVIU) 2015
49. **C. Li**, M. Ovsjanikov and F. Chazal
"Persistence-based Structural Recognition"
Computer Vision and Pattern Recognition (CVPR) 2014
50. Z. Ren, J. Yuan, **C. Li** and W. Liu
"Minimum Near-Convex Decomposition for Shape Representation"
International Conference on Computer Vision (ICCV) 2011
51. **C. Li** and A. Ben Hamza
"Spatially Aggregating Spectral Descriptors for Non-Rigid 3D Shape Retrieval: A Comprehensive Comparison", *Multimedia Systems*, 2014
52. **C. Li** and A. Ben Hamza
"Symmetry Discovery and Retrieval of Nonrigid 3D Shapes using Geodesic Skeleton Paths",
Multimedia Tools and Applications, 2014
53. **C. Li** and A. Ben Hamza
"A Multi-Resolution Descriptor for Deformable 3D Shape Retrieval",
Visual Computer (Computer Graphics International, acceptance rate 18%), 2013

Patents

- "Metric Forecasting Employing a Similarity Determination in a Digital Medium Environment"
C. Li, H. Bui, M. Ghavamzadeh and G. Theoharous, *US20180276691A1*
- "Generating Compressed Representation Neural Networks Having High Degree of Accuracy"
Jason. Yosinski, **C. Li**, and Ruoqian Liu, *US20190130272A1*

Teaching Experiences

Teaching assistant. Besides grading and office hours, I gave the following lectures.

STA571 Machine Learning: Design discussion material and lead the discussion lecture every week

ECE681 Pattern Classification: Guest Lecture on *Introduction to Deep Neural Networks*

Students/Interns Mentored

- Yang Zhao, ICML 2020 paper, PhD student at University of Buffalo
- Paidamoyo Chapfuwa, CHIL 2020 paper, PhD student at Duke University
- Christy Li, AAAI 2020 paper, PhD student at Duke University
- Ruqi Zhang, ICLR 2020 paper, PhD student at Cornell University
- Le Fang, EMNLP 2019 paper, PhD student at University of Buffalo
- Hao Fu, NAACL 2019 paper, PhD student at Duke University
- Ke Bai, AISTATS 2019 paper, PhD student at Duke University
- Guoyin wang, ACL 2018 paper, PhD student at Duke University
- Ruiyi Zhang, AISTATS 2018 paper, PhD student at Duke University
- Sam Lobel, ICLR 2020 paper. Visiting student at Duke, now PhD student at Brown University
- Hao Liu, NIPS 2017 paper. Visiting student at Duke, now PhD student at Caltech

Academic Activities

Senior PC : AAAI 2020/2021

Reviewer / PC :

- Natural Sciences and Engineering Research Council of Canada (NSERC)
- NIPS 2020/2019/2018/2016, ICML 2019/2018, ICLR 2020/2019/2018, AISTATS 2019/2018
- ICCV 2019, CVPR 2019/2018, ECCV 2020, ACCV 2018
- ACL 2020/2018, NAACL 2020, NAACL 2019
- IJCAI 2020/2019, AAAI 2019/2018
- IEEE Transactions on Pattern Analysis and Machine Intelligence
- International Journal of Computer Vision
- Computer Vision and Image Understanding
- Pattern Recognition
- IEEE Transactions on Neural Networks and Learning Systems

Organizer:

- Weakly-supervised and Unsupervised Learning Workshop, *SIAM SDM* 2020
- Two SHREC 3D shape retrieval contests in *Eurographics workshop on 3DOR* 2014

Talks

- “Towards Better Representations with Deep/Bayesian Learning”
Salesforce Research, Palo Alto, January 2018
IBM Watson Research Center, Boston & Yorktown, April 2018
Microsoft Research, Redmond, June 2018
Google, Mountain View, August 2018
- “Scalable Bayesian Learning of Recurrent Neural Networks for Language Modeling”
ACL, Vancouver, Canada, August 2017
- “Scalable Bayesian Methods for Deep Learning”, OpenAI, San Francisco, Feb. 2017
- “Learning Weight Uncertainty with Stochastic Gradient MCMC for Shape Classification”
Computer Vision and Pattern Recognition, Las Vegas, NV, June 2016
- “Preconditioned Stochastic Gradient Langevin Dynamics for Deep Neural Networks”
AAAI Conference on Artificial Intelligence, Phoenix, AZ, Feb. 2016
- “Large-scale Comprehensive 3D Shape Retrieval”
Eurographics workshop on 3DOR, Strasbourg, France, April 2014

Software Skills

Python (Pytorch, Tensorflow and Keras), Matlab, R and C