Zhe Gan

Microsoft Advanta-B, 6227, Phone: Provided upon request 4200 150th Ave NE, Email: zhe.gan@microsoft.com Redmond, WA 98052 Homepage: http://zhegan27.github.io/

Research Interests

I am a Senior Researcher at Microsoft Dynamics 365 AI Research, primarily working on Vision-and-Language Representation Learning, Generative Pre-training, and Adversarial Machine Learning. I also have broad interests on other machine learning topics.

Education

• Duke University, Durham, NC

Ph.D., Electrical and Computer Engineering

09/2013 - 03/2018

Peking University, Beijing, China

M.S., Electrical Engineering B.S., Electrical Engineering 09/2010 - 07/2013

09/2006 - 07/2010

Experience

• Microsoft Cloud and AI

04/2018 - present

Senior Researcher. Manager: Jingjing Liu, MS D365 AI

Vision-and-Language Representation Learning, Generative Pre-training, Self-supervised Learning, and Adversarial Machine Learning

• Information Initiative at Duke (iiD)

09/2013 - 03/2018

Research Assistant. Advisor: Prof. Lawrence Carin

- (i) Deep Bayesian Learning: developing deep generative models for computer vision and natural language processing applications, including VAE and GAN
- (ii) Bayesan Deep Learning: designing stochastic gradient variational inference algorithms and stochastic gradient MCMC methods for scalable Bayesian inference

• Microsoft Research Redmond

05/2017 - 08/2017

Research Intern. Advisor: Xiaodong He, Lihong Li, Ph.D

Deep reinforcement learning for vision and language intelligence, with focus on the visual storytelling task.

• Microsoft Research Redmond

05/2016 - 08/2016

Research Intern. Advisor: Xiaodong He, Jianfeng Gao, Li Deng, Ph.D

- (i) image captioning: using deep learning techniques to improve the state-of-the-art of image and video captioning.
- (ii) deep conflation: using deep learning techniques to implement conflation for business data analytics.

• Adobe Research

06/2015 - 09/2015

Data Scientist Intern. Advisor: Hung Bui, Ph.D

Recurrent neural networks (RNN) for NLP applications, including sentence classification, sentence retrieval and sentence generation

Publications

arXiv preprints

- 1. **Z. Gan**, Y.-C. Chen, L. Li, C. Zhu, Y. Cheng and J. Liu "Large-Scale Adversarial Training for Vision-and-Language Representation Learning", arXiv preprint arXiv:2006.06195
- 2. Y. Fang*, S. Wang*, **Z. Gan**, S. Sun and J. Liu "FILTER: An Enhanced Fusion Method for Crosslingual Language Understanding", arXiv preprint arXiv:2009.05166 Leaderboard #1 on XTREME and XGLUE as of Sep. 8, 2020
- 3. S. Wang, L. Zhou, **Z. Gan**, Y.-C. Chen, Y. Fang, S. Sun, Y. Cheng and J. Liu "Cluster-Former: Clustering-based Sparse Transformer for Long-Range Dependency Encoding", arXiv preprint arXiv:2009.06097
- 4. Y. Fang, S. Wang, **Z. Gan**, S. Sun and J. Liu "Accelerating Real-Time Question Answering via Question Generation", *arXiv preprint arXiv:2009.05167*
- 5. C. Zhu, Y. Cheng, **Z. Gan**, F. Huang, J. Liu and T. Goldstein "Adaptive Learning Rates with Maximum Variation Averaging", arXiv preprint arXiv:2006.11918
- 6. S. Dai, **Z. Gan**, Y. Cheng, C. Tao, L. Carin and J. Liu "APo-VAE: Text Generation in Hyperbolic Space", *arXiv* preprint arXiv:2005.00054
- 7. S. Dai, Y. Cheng, Y. Zhang, **Z. Gan**, J. Liu and L. Carin "Contrastively Smoothed Class Alignment for Unsupervised Domain Adaptation", *arXiv* preprint arXiv:1909.05288

2021

1. W. Chen, **Z. Gan**, L. Li, Y. Cheng, W. Wang and J. Liu "Meta Module Network for Compositional Visual Reasoning", *Winter Conf. on Applications of Computer Vision* (WACV), 2021 Oral

2020

- 1. S. Sun, **Z. Gan**, Y. Fang, Y. Cheng, S. Wang and J. Liu "Contrastive Distillation on Intermediate Representations for Language Model Compression", *Conf. on Empirical Methods in Natural Language Processing* (EMNLP), 2020
- 2. S. Wang, Y. Fang, S. Sun, **Z. Gan**, Y. Cheng, J. Liu and J. Jiang "Cross-Thought for Sentence Encoder Pre-training", Conf. on Empirical Methods in Natural Language Processing (EMNLP), 2020
- 3. Y. Dong, S. Wang, **Z. Gan**, Y. Cheng, J. Cheung and J. Liu "Multi-Fact Correction in Abstractive Text Summarization", Conf. on Empirical Methods in Natural Language Processing (EMNLP), 2020
- 4. L. Li*, Y.-C. Chen*, Y. Cheng, **Z. Gan**, L. Yu and J. Liu "HERO: Hierarchical Encoder for Video+Language Omni-representation Pre-training", *Conf. on Empirical Methods in Natural Language Processing* (EMNLP), 2020 Leaderboard #1 on TVR and TVC as of Sep. 15, 2020
- 5. Y. Zhang*, G. Wang*, C. Li, **Z. Gan**, C. Brockett and B. Dolan "POINTER: Constrained Text Generation via Insertion-based Generative Pre-training", *Conf. on Empirical Methods in Natural Language Processing* (EMNLP), 2020
- 6. Y. Fang, S. Sun, **Z. Gan**, R. Pillai, S. Wang and J. Liu "Hierarchical Graph Network for Multihop Question Answering", arXiv preprint arXiv:1911.03631 Conf. on Empirical Methods in Natural Language Processing (EMNLP), 2020 Leaderboard #1 on HotpotQA as of Dec. 1st, 2019
- 7. Y. Cheng, **Z. Gan**, Y. Zhang, O. Elachqar, D. Li and J. Liu "Contextual Text Style Transfer", Findings of Empirical Methods in Natural Language Processing (Findings of EMNLP), 2020
- 8. J. Cao, **Z. Gan**, Y. Cheng, L. Yu, Y.-C. Chen and J. Liu "Behind the Scene: Revealing the Secrets of Pre-trained Vision-and-Language Models", *European Conf. on Computer Vision* (ECCV), 2020 Spotlight (Top 5% among all submissions)
- Y.-C. Chen*, L. Li*, L. Yu*, A. Kholy, F. Ahmed, Z. Gan, Y. Cheng and J. Liu "UNITER: UNiversal Image-TExt Representation Learning", European Conf. on Computer Vision (ECCV), 2020 SOTA on 13 Vision+Language Datasets/Tasks, No. 1 on VCR and NLVR2 leaderboards as of Sep. 2019
- 10. Y. Cheng, **Z. Gan**, Y. Li, J. Liu and J. Gao "Sequential Attention GAN for Interactive Image Editing", ACM International Conference on Multimedia (ACMMM), 2020

- 11. P. Cheng, W. Hao, S. Dai, J. Liu, **Z. Gan** and L. Carin "CLUB: A Contrastive Log-ratio Upper Bound of Mutual Information", *Int. Conf. Machine Learning* (ICML), 2020
- 12. L. Chen, **Z. Gan**, Y. Cheng, L. Li, L. Carin and J. Liu "Graph Optimal Transport for Cross-Domain Alignment", *Int. Conf. Machine Learning* (**ICML**), 2020
- 13. J. Xu, **Z. Gan**, Y. Cheng and J. Liu "Discourse-Aware Neural Extractive Text Summarization", *Association for Computational Linguistics* (ACL), 2020
- 14. Y. Chen, **Z. Gan**, Y. Cheng, J. Liu and J. Liu "Distilling Knowledge Learned in BERT for Text Generation", Association for Computational Linguistics (ACL), 2020
- 15. R. Zhang, C. Chen, **Z. Gan**, W. Wang, D. Shen, G. Wang, Z. Wen and L. Carin "Improving Adversarial Text Generation by Modeling the Distant Future", *Association for Computational Linguistics* (ACL), 2020
- 16. Y. Li, Y. Cheng, **Z. Gan**, L. Yu, L. Wang, J. Liu "BachGAN: High-Resolution Image Synthesis from Salient Object Layout", *Computer Vision and Pattern Recognition* (CVPR), 2020
- 17. J. Liu, W. Chen, Y. Cheng, **Z. Gan**, L. Yu, Y. Yang, J. Liu "VIOLIN: A Large-Scale Dataset for Video-and-Language Inference", Computer Vision and Pattern Recognition (CVPR), 2020
- 18. R. Zhang, C. Chen, **Z. Gan**, Z. Wen, W. Wang, L. Carin "Nested-Wasserstein Self-Imitation Learning for Sequence Generation", *Artificial Intelligence and Statistics* (AISTATS), 2020
- 19. C. Zhu, Y. Cheng, **Z. Gan**, S. Sun, T. Goldstein and J. Liu "FreeLB: Enhanced Adversarial Training for Natural Language Understanding", *Int. Conf. Learning Representations* (ICLR), 2020 Spotlight (Leaderboard #1 on GLUE, ARC Easy/Challenge and Commonsense QA as of Sep. 2019)
- 20. W. Wang, H. Xu, **Z. Gan**, B. Li, G. Wang, L. Chen, Q. Yang, W. Wang and L. Carin "Graph-Driven Generative Models for Heterogeneous Multi-Task Learning", *Proc. American Association of Artificial Intelligence* (AAAI), 2020 Spotlight
- 21. J. Hu, Y. Cheng, **Z. Gan**, J. Liu, J. Gao and G. Neubig "What Makes A Good Story? Designing Composite Rewards for Visual Storytelling", *Proc. American Association of Artificial Intelligence* (AAAI), 2020 Spotlight

2019

- 1. W. Wang, C. Tao, **Z. Gan**, G. Wang, L. Chen, X. Zhang, R. Zhang, Q. Yang, R. Henao and L. Carin "Improving Textual Network Learning with Variational Homophilic Embeddings", *Neural Information Processing Systems* (NeurIPS), 2019
- 2. R. Zhang, C. Chen, **Z. Gan**, Z. Wen, W. Wang, L. Carin "Nested-Wasserstein Distance for Sequence Generation", Workshop on Bayesian Deep Learning, NeurIPS 2019
- 3. S. Sun, Y. Cheng, **Z. Gan**, and J. Liu "Patient Knowledge Distillation for BERT Model Compression", Conf. on Empirical Methods in Natural Language Processing (EMNLP), 2019
- 4. H. Wang, **Z. Gan**, X. Liu, J. Liu, J. Gao and H. Wang "Adversarial Domain Adaptation for Machine Reading Comprehension", *Conf. on Empirical Methods in Natural Language Processing* (EMNLP), 2019
- 5. D. Li, Y. Zhang, **Z. Gan**, Y. Cheng, C. Brockett, M. Sun and B. Dolan "Domain Adaptive Text Style Transfer", Conf. on Empirical Methods in Natural Language Processing (EMNLP), 2019
- 6. M. Jiang, Q. Huang, L. Zhang, X. Wang, P. Zhang, **Z. Gan**, J. Diesner and J. Gao "TIGEr: Text-to-Image Grounding for Image Caption Evaluation", *Conf. on Empirical Methods in Natural Language Processing* (EMNLP), 2019
- 7. L. Li, **Z. Gan**, Y. Cheng and J. Liu "Relation-Aware Graph Attention Network for Visual Question Answering", *Int. Conf. on Computer Vision* (ICCV), 2019
- 8. **Z. Gan**, Y. Cheng, A. Kholy, L. Li, J. Liu and J. Gao "Multi-step Reasoning via Recurrent Dual Attention for Visual Dialog", *Association for Computational Linguistics* (**ACL**), 2019
- 9. L. Ke, X. Li, Y. Bisk, A. Holtzman, Z. Gan, J. Liu, J. Gao, Y. Choi, and S. Srinivasa "Tactical Rewind: Self-Correction via Backtracking in Vision-and-Language Navigation", Computer Vision and Pattern Recognition (CVPR), 2019 Oral

- 10. Y. Li, **Z. Gan**, Y. Shen, J. Liu, Y. Cheng, Y. Wu, L. Carin, D. Carlson and J. Gao "StoryGAN: A Sequential Conditional GAN for Story Visualization", *Computer Vision and Pattern Recognition* (CVPR), 2019
- 11. W. Wang, **Z. Gan**, H. Xu, R. Zhang, G. Wang, D. Shen, C. Chen and L. Carin "Topic-Guided Variational Autoencoders for Text Generation", North American Chapter of the Association for Computational Linguistics (NAACL), 2019 Oral
- 12. L. Chen, Y. Zhang, R. Zhang, C. Tao, **Z. Gan**, H. Zhang, B. Li, D. Shen, C. Chen and L. Carin "Improving Sequence-to-Sequence Learning via Optimal Transport", *Int. Conf. Learning Representations* (ICLR), 2019
- 13. Q. Huang*, **Z. Gan***, A. Celikyilmaz, D. Wu, J. Wang and X. He "Hierarchically Structured Reinforcement Learning for Topically Coherent Visual Story Generation", *Proc. American Association of Artificial Intelligence* (**AAAI**), 2019 **Spotlight**

2018

- 1. Y. Zhang, M. Galley, J. Gao, Z. Gan, X. Li, C. Brockett and B. Dolan "Generating Informative and Diverse Conversational Responses via Adversarial Information Maximization", *Neural Information Processing Systems* (NeurIPS), 2018
- 2. L. Chen, S. Dai, C. Tao, D. Shen, **Z. Gan**, H. Zhang, Y. Zhang and L. Carin "Adversarial Text Generation via Feature-Mover's Distance", *Neural Information Processing Systems* (**NeurIPS**), 2018
- 3. X. Zhang, R. Henao, **Z. Gan**, Y. Li and L. Carin "Multi-Label Learning from Medical Plain Text with Convolutional Residual Models", *Machine Learning for Healthcare* (MLHC), 2018 Spotlight
- 4. Y. Pu, S. Dai, **Z. Gan**, W. Wang, G. Wang, Y. Zhang, R. Henao and L. Carin "JointGAN: Multi-Domain Joint Distribution Learning with Generative Adversarial Nets", *Int. Conf. Machine Learning* (ICML), 2018
- 5. T. Xu, P. Zhang, Q. Huang, H. Zhang, Z. Gan, X. Huang and X. He "AttnGAN: Fine-Grained Text to Image Generation with Attentional Generative Adversarial Networks", Computer Vision and Pattern Recognition (CVPR), 2018
- 6. W. Wang, Z. Gan, W. Wang, D. Shen, J. Huang, W. Ping, S. Satheesh and L. Carin "Topic Compositional Neural Language Model", Artificial Intelligence and Statistics (AISTATS), 2018
- 7. Y. Pu, M. R. Min, **Z. Gan** and L. Carin "Adaptive Feature Abstraction for Translating Video to Text", *Proc. American Association of Artificial Intelligence* (**AAAI**), 2018

2017

- 1. **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* (NeurIPS), 2017
- 2. Y. Pu, W. Wang, R. Henao, L. Chen, **Z. Gan**, C. Li, and L. Carin "Adversarial Symmetric Variational Autoencoder", *Neural Information Processing Systems* (NeurIPS), 2017
- 3. Y. Pu, **Z. Gan**, R. Henao, C. Li, S. Han and L. Carin "VAE Learning via Stein Variational Gradient Descent", *Neural Information Processing Systems* (**NeurIPS**), 2017
- 4. Y. Zhang, D. Shen, G. Wang, **Z. Gan**, R. Henao and L. Carin "Deconvolutional Paragraph Representation Learning", *Neural Information Processing Systems* (**NeurIPS**), 2017
- 5. **Z. Gan**, Y. Pu, R. Henao, C. Li, X. He and L. Carin "Learning Generic Sentence Representations Using Convolutional Neural Networks", *Conf. on Empirical Methods in Natural Language Processing* (EMNLP), 2017 Oral
- 6. Y. Zhang, **Z. Gan**, K. Fan, Z. Chen, R. Henao, D. Shen and L. Carin "Adversarial Feature Matching for Text Generation", *Int. Conf. Machine Learning* (ICML), 2017
- 7. Y. Zhang, C. Chen, **Z. Gan**, R. Henao and L. Carin "Stochastic Gradient Monomial Gamma Sampler", *Int. Conf. Machine Learning* (ICML), 2017
- 8. **Z. Gan***, C. Li*, C. Chen, Y. Pu, Q. Su and L. Carin "Scalable Bayesian Learning of Recurrent Neural Networks for Language Modeling", Association for Computational Linguistics (ACL), 2017 Oral

- 9. **Z. Gan**, C. Gan, X. He, Y. Pu, K. Tran, J. Gao, L. Carin and L. Deng "Semantic Compositional Networks for Visual Captioning", *Computer Vision and Pattern Recognition* (CVPR), 2017 Spotlight
- 10. C. Gan, **Z. Gan**, X. He, J. Gao and L. Deng "StyleNet: Generating Attractive Visual Captions with Styles", Computer Vision and Pattern Recognition (CVPR), 2017
- 11. **Z. Gan**, P. D. Singh, A. Joshi, X. He, J. Chen, J. Gao and L. Deng "Character-level Deep Conflation for Business Data Analytics", *Int. Conf. Acoustics, Speech and Signal Processing* (ICASSP), 2017
- 12. Y. Xian, Y. Pu, **Z. Gan**, L. Lu and A. Thompson "Adaptive DCTNet for Audio Signal Classification", *Int. Conf. Acoustics, Speech and Signal Processing* (ICASSP), 2017
- 13. Q. Su, X. Liao, C. Li, **Z. Gan** and L. Carin "Unsupervised Learning with Truncated Gaussian Graphical Models", *Proc. American Association of Artificial Intelligence* (**AAAI**), 2017 Oral

2016

- Y. Zhang, Z. Gan and L. Carin "Generating Text via Adversarial Training", NeurIPS Workshop, 2016
- 2. Y. Xian, Y. Pu, **Z. Gan**, L. Lu and A. Thompson "Modified DCTNet for Audio Signals Classification", *Journal of the Acoustical Society of America*, 2016
- 3. Y. Pu, **Z. Gan**, R. Henao, X. Yuan, C. Li, A. Stevens and L. Carin "Variational Autoencoder for Deep Learning of Images, Labels and Captions", *Neural Information Processing Systems* (**NeurIPS**), 2016
- 4. J. Song, **Z. Gan** and L. Carin "Factored Temporal Sigmoid Belief Networks for Sequence Learning", *Int. Conf. Machine Learning* (ICML), 2016
- 5. 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
- 6. 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

2015

- 1. **Z. Gan**, C. Li, R. Henao, D. Carlson and L. Carin "Deep Temporal Sigmoid Belief Networks for Sequence Modeling", *Neural Information Processing Systems* (**NeurIPS**), 2015
- 2. R. Henao, **Z. Gan**, J. Lu and L. Carin "Deep Poisson Factor Modeling", *Neural Information Processing Systems* (NeurIPS), 2015
- 3. **Z. Gan**, C. Chen, R. Henao, D. Carlson and L. Carin "Scalable Deep Poisson Factor Analysis for Topic Modeling", *Int. Conf. Machine Learning* (**ICML**), 2015
- 4. **Z. Gan**, R. Henao, D. Carlson and L. Carin "Learning Deep Sigmoid Belief Networks with Data Augmentation", *Artificial Intelligence and Statistics* (AISTATS), 2015

Book Chapter

 Z. Gan, X. Yuan, R. Henao, E. Tsalik and L. Carin "Inference of Gene Networks Associated with the Host Response to Infectious Disease", Chapter 13 of Book Big Data Over Networks. Cambridge University Press. In Press.

PhD Dissertation

1. **Z. Gan** "Deep Generative Models for Vision and Language Intelligence", Duke University.

Tutorial and Workshop

Tutorial

1. **Z. Gan**, L. Yu, Y. Cheng, L. Zhou, L. Li, Y.-C. Chen, J. Liu and X. He "Recent Advances in Vision-and-Language Research", *Computer Vision and Pattern Recognition* (CVPR), 2020

Professional Activities

Area Chair: ICLR 2021, NeurIPS 2020, NeurIPS 2019

Senior Program Committee (SPC) Member: AAAI 2021, AAAI 2020

Interactive Arts Chair: ACMMM 2020

Awarded as Top Reviewer: ICML 2020, NeurIPS 2018

Conference Reviewer/PC Member:

- 2020: ICML, ICLR, IJCAI; CVPR, ECCV, ACMMM; ACL, EMNLP, COLING, AACL, CoNLL
- 2019: ICML, ICLR, AAAI; CVPR, ICCV, ACMMM; EMNLP, CoNLL
- 2018: NeurIPS, EMNLP, CVPR, ACCV
- 2016: NIPS

Journal Reviewer:

- Transactions on Pattern Analysis and Machine Intelligence (PAMI)
- Transactions on Image Processing (TIP)
- Transactions on Knowledge and Data Engineering (KDE)
- Journal of Selected Topics in Signal Processing (STSP)
- Transactions on Multimedia Computing Communications and Applications (TOMM)
- Transactions on Audio, Speech and Language Processing (ASL)
- Science China
- Transactions on Cybernetics, IET Computer Vision, Entropy

Workshop Reviewer/PC Member:

- 2020: ACL Workshop on Advances in Language and Vision Research
- 2019: ICCV Workshop on Closing the loop between Vision and Language
- 2019: ICLR Workshop on Deep Generative Models for Highly Structured Data
- 2018: ICML Workshop on Theoretical Foundations and Applications of Deep Generative Models

Talks

- "Deep Generative Models for Vision and Language Intelligence", *Ph.D. Final Defense*, Durham, NC, February 2018
- "Deep Generative Models for Vision and Language Intelligence", IBM Thomas J. Watson Research Center, Yorktown, NY, October 2017
- "Deep Generative Models for Vision and Language Intelligence", NVIDIA, Santa Clara, CA, September 2017
- "Deep Generative Models for Vision and Language Intelligence", Apple, Cupertino, CA, September 2017
- "Learning Generic Sentence Representations Using Convolutional Neural Networks", EMNLP, Copenhagen, Denmark, September 2017
- "Semantic Compositional Networks for Visual Captioning", CVPR, Hawaii, July 2017
- "Semantic Compositional Networks for Visual Captioning", *Ph.D. Preliminary Exam*, Durham, NC, April 2017
- "Deep Generative Models for Sequence Learning", Ph.D. Qualifying Exam, Durham, NC, December 2015

Competitions

- 2020/09: Rank 1st on TVR and TVC leaderboards
- 2020/09: Rank 1st on XTREME and XGLUE leaderboards
- 2020/05: Rank 4th on VQA Challenge 2020
- 2019/12: Rank 1st on HotpotQA leaderboard
- 2019/10: Rank 1st on VCR and NLVR2 leaderboards
- 2019/09: Rank 1st on GLUE, ARC Easy/Challenge and Commonsense QA leaderboards
- 2019/06: Rank 2nd in Visual Dialog Challenge 2019
- 2018/09: Rank 3rd in Visual Dialog Challenge 2018

Software Skills

Python (Theano, Tensorflow, PyTorch), Matlab, R and C

Awards

AAAI-2020 Outstanding SPC Award. ECE Fellowship, Duke University, 2013. National Scholarship, Department of Minister of Education of China, 2010-2013.

Graduate Coursework

Bayesian and Modern Statistics, Probabilistic Machine Learning, Advanced Machine Learning, Statistical Inference, Statistical Computation, Information Theory, Graphical Models & Inference, Optimization For Engineers

Teaching Experience

Teaching Assistant 09/2014-12/2014

STA 601 - Bayesian and Modern Statistics Instructor: David Dunson, Ph.D

Teaching Assistant 01/2015-05/2015

ECE 587 - Information Theory Instructor: Ahmad Beirami, Ph.D