HUANGJIE ZHENG

Statistics & Data Science Department, University of Texas at Austin
email: huangjie.zheng@utexas.edu homepage: https://huangjiezheng.com/

RESEARCH INTERESTS

General Interests: Machine Learning and Deep Bayesian methods for probabilistic modeling with application to several domains, such as Deep Generative Models, Representation Learning, etc. My recent research topics involves:

- Deep generative model, *e.g.* Generative Adversarial Networks, Variational Autoencoders, Diffusion-based methods, *etc.*
- Self-supervised Representation Learning from high-dimensional data.
- Corresponding problems in relevant fields like computer vision, language modeling, *etc.*

EDUCATION

University of Texas at Austin August 2019 - Present
PhD student in Statistics and Data Science.
GPA: 4.0/4.0

Shanghai Jiao Tong University March 2017 - March 2019
M.S. in Information Engineering (Sino-French Cooperative Education Program).
GPA: 3.87/4.0

Telecom ParisTech ENST September 2016 - March 2017
Grande Ecole Engineer Cycle in Data Science and Computer Networking.
Average Course Grade: A

Shanghai Jiao Tong University September 2012 - August 2016
B.S. in Information Engineering (Sino-French Cooperative Education Program).
GPA: 3.66/4.3

WORK EXPERIENCE

Research intern at ByteDance	2023 Summer
Research intern at Microsoft Research	2022 Summer
Research intern at Microsoft Research	2021 Summer
Research Assistant at the University of Texas at Austin	Academic year 2021-2024
Teaching Assistant at the University of Texas at Austin	Academic year 2019-2021

SELECTED PUBLICATIONS

[1] **H. Zheng**, Z. Wang, J. Yuan, G. Ning, P. He, Q. You, H. Yang, and M. Zhou. Learning Stackable and Skippable LEGO Bricks for Efficient, Reconfigurable, and Variable-Resolution Diffusion Modeling. Preprint, 2023 [PDF]

[2] **H. Zheng**, X. Chen, J. Yao, H. Yang, C. Li, Y. Zhang H. Zhang, I. W. Tsang, J. Zhou and M. Zhou. Contrastive Attraction and Contrastive Repulsion for Representation Learning. Transactions on Machine Learning Research (TMLR). [PDF]

[3] M. Armandpour, **H. Zheng***, A. Sadeghian, A. Sadeghian, M. Zhou. Re-imagine the Negative Prompt Algorithm: Transform 2D Diffusion into 3D, alleviate Janus problem and Beyond. Preprint, 2023 [PDF]

[4] Y. Qin, **H. Zheng**, J. Yao, M. Zhou and Y. Zhang. Class-Balancing Diffusion Models. The IEEE/CVF Conference on Computer Vision and Pattern Recognition 2023 (CVPR 2023). [PDF]

*Equal contribution to the first author

[5] **H. Zheng**, P. He, W. Chen and M. Zhou. Truncated Diffusion Probabilistic Models and Diffusion-based Adversarial Autoencoders. International Conference on Learning Representations 2022. (ICLR 2023). [PDF]

[6] X. Han, **H. Zheng*** and M. Zhou. CARD: Classification and Regression Diffusion Models. Advances in Neural Information Processing Systems 36. (NeurIPS 2022) [PDF]

[7] **H. Zheng**, P. He, C. Li, J. Gao, W. Chen and M. Zhou. Good Image Generators Improve Discriminative Visual Representation Learning. 2022.

[8] **H. Zheng**, P. He, W. Chen and M. Zhou. Mixing and Shifting: Exploiting Global and Local Dependencies in Vision MLPs. Preprint, 2022. [PDF]

[9] **H. Zheng** and M. Zhou. Exploiting Chain Rule and Bayes' Theorem to Compare Probability Distributions. Advances in Neural Information Processing Systems 35. (NeurIPS 2021) [PDF]

[10] K. Tanwisuth, X. FAN, **H. Zheng***, S Zhang, H. Zhang, B. Chen, and M. Zhou. A Prototype-Oriented Framework for Unsupervised Domain Adaptation. Advances in Neural Information Processing Systems 35. (NeurIPS 2021, first three authors have equal contribution). [PDF]

[11] **H. Zheng**, L. Xie, T. Ni, Y. Zhang, Y. Wang, Q. Tian, E. K. Fishman, A. L. Yuille. Incorporating Multi-Phase Information for Medical Imaging Segmentation. Preprint, 2019. [PDF]

[12] **H. Zheng**, J. Yao, Y. Zhang, I. W. Tsang and J. Wang. Understanding VAEs in Fisher-Shannon Plane. AAAI Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence, 2019. [PDF]

[13] **H. Zheng**, J. Yao, Y. Zhang and I. W. Tsang. Degeneration in VAE: in the Light of Fisher Information Loss. Preprint, 2018. [PDF]

[14] **H. Zheng**, Y. Wang, C. Han, F. Le, R. He and J. Lu. Learning and Utilizing Ontology with Machine Learning in Attack Detection. 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/ 12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE) 2018. (Oral Presentation) [Link]

[15] **H. Zheng**, J. Yao, and Y. Zhang. Describing Geographical Characteristics with Social Images. MultiMedia Modeling. Springer International Publishing, 2017. (Oral Presentation) [Link]

OTHER PUBLICATIONS

[16] M. Zhou, T. Chen, Z. Wang and **H. Zheng**. Beta diffusion. Advances in Neural Information Processing Systems 37. (NeurIPS 2023). [PDF]

[17] Z. Wang, Y. Jiang **H. Zheng**, P. Wang, P. He, Z. Wang, W. Chen and M. Zhou, Patch Diffusion: Faster and More Data-Efficient Training of Diffusion Models. Advances in Neural Information Processing Systems 37. (NeurIPS 2023). [PDF]

[18] K. Tanwisuth, S Zhang, **H. Zheng**, P. He, and M. Zhou. POUF: Prompt-oriented unsupervised fine-tuning for large pre-trained models. International Conference of Machine Learning. (ICML 2023). [PDF]

[19] Z. Wang, X. Zhang, Z. Zhang, **H. Zheng**, J. Yao, and Y. Zhang, Y. Wang and M. Zhou. DR2: Diffusion-based Robust Degradation Remover for Blind Face Restoration. The IEEE/CVF Conference on Computer Vision and Pattern Recognition 2023 (CVPR 2023). [PDF]

[20] Z. Wang, **H. Zheng**, P. He, W. Chen and M. Zhou. Diffusion-GAN: Training GANs with Diffusion. International Conference on Learning Representations 2022. (ICLR 2023). [PDF]

[21] D. Wang, D. Guo, H. Zhao, **H. Zheng**, K. Tanwisuth, B. Chen and, M. Zhou, Representing Mixtures of Word Embeddings with Mixtures of Topic Embeddings. International Conference on Learning Representations 2022. (ICLR 2022) [PDF]

[22] S. Zhang, X. Fan, **H. Zheng**, K. Tanwisuth and M. Zhou. Alignment Attention by Matching Key and Query Distributions. Advances in Neural Information Processing Systems 35. (NeurIPS 2021) [PDF]

[23] H. Zhang, Z. Wang, J. Xu, **H. Zheng**, M. Zhou, J. Bian and F. Wang. Semantic Transport for Unsupervised Domain Adaptation.

[24] X. Chen, S. Chen, J. Yao, **H. Zheng**, Y. Zhang, and I. W. Tsang. Learning on Attribute-Missing Graphs. IEEE transactions on pattern analysis and machine intelligence, 2020. [PDF]

[25] F. Ye, **H. Zheng**, C. Huang and Y. Zhang. Deep Unsupervised Image Anomaly Detection: An Information Theoretic Framework. IEEE International Conference on Image Processing 2021 (ICIP 2021). [PDF]

[26] Q. Zhang, **H. Zheng**, M. Zhou. MCMC-Interactive Variational Inference. Preprint, 2020. [PDF]

[27] T. Ni, L. Xie, **H. Zheng**, E. K. Fishman, A. L. Yuille. Elastic Boundary Projection for 3D Medical Imaging Segmentation. CVPR 2019 [PDF]

SELECTED PROJECTS

LEGO: Stackable and Skippable Bricks for Efficient, Reconfigurable, and Variable-Resolution Diffusion Modeling.

University of Texas at Austin, advised by Prof. Mingyuan Zhou and are done in the internship at Bytedance

- Currently diffusion models with architecture like U-Net and Vision Transformer often rely on resource-intensive deep networks and lack the flexibility needed for generating images at variable resolutions.
- We introduces LEGO bricks, which seamlessly integrate *Local-feature Enrichment and Global-content Orchestration*. Each brick models learns to generate images at variable resolutions and can be stacked to create a giant diffusion model backbone.
- The training of LEGO bricks is efficient, as each brick can be trained with sampled patches, which largely reduces the computation FLOPs. Even for the brick that handles the entire image, it is less massive in terms of model size and is more computational efficient.
- We show LEGO is flexible in sampling, as we may select to skip bricks to reduce sampling costs. Moreover, LEGO is able to generate higher-resolution images than the training data by repeating the process of *Local-feature Enrichment and Global-content Orchestration*. In our experiments, we successfully generate images with resolution as large as 4k, using LEGO model trained on image with only 256×256 resolution.

Truncated diffusion probabilistic models and GANs with diffusion

University of Texas at Austin, advised by Prof. Mingyuan Zhou and some parts are done in the internship at Microsoft

- We propose a more efficient approach that adds noise not until the data become pure random noise, but until they reach a hidden noisy-data distribution that we can confidently learn with implicit models such as GANs.
- We reveal that the proposed model can be cast as an adversarial auto-encoder empowered by both the diffusion process and a learnable implicit prior.
- Even with a significantly smaller number of reverse diffusion steps, the proposed truncated diffusion probabilistic models can provide consistent improvements over the non-truncated ones in terms of performance image generations. With a well-learned implicit model, the truncated diffusion model only requires a few reverse steps, e.g., 5, 10 steps to generate satisfactory images in both unconditional or text-guided generation.

Large-scale self-supervised pretraining

Internship at Microsoft Research

- We investigate contrastive learning in view of conditional transport. Intuition is the feature extractor that is encouraged to align the distributions of the positive samples and distinguish those of the negative samples. Better interpretability and robustness are very nice properties of our method. Our method shows better performance than SOTA methods on various datasets. Moreover, our method is compatible with most of the existing methods.
- We further investigate the relation between contrastive learning and the pretraining of masked image models (MIM). MIMs are inspired by the success of language processing, where the mask language model successfully pretrain the model to provide powerful representations with the masked tokens. We propose an adversarial self-supervised pretraining framework that consists of both a generator and a discriminator. The generator is a masked image generator, aiming to fool the discriminator, which is trained with a patch-level and an image-level

discrimination loss and a contrastive loss. Different than previous masked image models, we take the discriminator backbone for the downstream tasks and find such design could handle both fine-tuning and linear probing.

Conditional Transport for Unsupervised Domain Adaptation, and Unsupervised fine-tuning
University of Texas at Austin, advised by Prof. Mingyuan Zhou

- To avoid the sampling variability, class imbalance, and data-privacy concerns, we propose a prototype method that leverages conditional transport to connect data from different domain so as to benefit downstream tasks.
- Furthermore, when working with large pre-trained models, we leverage the prototype to connect the knowledge from the large foundation models with the incoming data for fast adaptation.

Conditional Transport for distribution matching May 2020 - November 2021
University of Texas at Austin, advised by Prof. Mingyuan Zhou

- We propose Conditional Transport (CT) as a new divergence to measure the difference between two probability distributions. CT consists of the expected cost of a forward transport from a data point of one distribution to the other distribution, and that of a backward CT which reverses the transport direction.
- Equipped with two navigators that amortize the computation of conditional transport plans, the CT divergence comes with unbiased sample gradients that are straightforward to compute, making it amenable to mini-batch stochastic gradient descent based optimization.
- On a wide variety of benchmark datasets for generative modeling, substituting the default statistical distance of an existing GAN with the CT divergence is shown to consistently improve the performance.
- Apart from generative models, we have already started working on applying CTs conditional transport plans to a wide variety of tasks, including contrastive representation learning, image-to-image translation, and imitation learning (inverse reinforcement learning) tasks, etc. and get good results on these tasks.

ACADEMIC SERVICES

Transaction of Machine Learning Research reviewer	active
IEEE Transactions on Pattern Analysis and Machine Intelligence reviewer	active
IEEE Transactions on Neural Networks and Learning Systems reviewer	active
Machine Learning Journal reviewer	active
ICML reviewer	2020-2023
ICLR reviewer	2020-2023
NeurIPS reviewer	2020-2023
IEEE ICDM reviewer	2022
ACML reviewer	2021
Teaching Assistant at SDS department, UT Austin	
Research Assistant at the research group of Dr. Mingyuan Zhou, UT Austin	

SELECTED HONORS AND AWARDS

Outstanding Graduates of Shanghai (Top 5% in Shanghai)	2019
National Scholarship for Graduate Students	2017
Excellent Teaching Assistant	2017
Outstanding Undergraduates of Shanghai (Top 5% in Shanghai)	2016
Excellent Undergraduate Thesis in SJTU (Top 1% in University)	2016
Meritorious Winner in Mathematical Contest in Modeling (MCM)	2015
Undergraduate-Entrance Bursary (Top 1% in Chinese University Entrance Exam)	2012