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 (UT) August 2019 - Present
PhD student in Statistics and Data Science.
GPA: 4.0/4.0

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

PUBLICATIONS

[1] Y. Qin, **H. Zheng**, J. Yao, M. Zhou and Y. Zhang. Class-Balancing Diffusion Models. CVPR 2023 submission.

[2] 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. CVPR 2023 submission.

[3] **H. Zheng**, P. He, W. Chen and M. Zhou. Truncated Diffusion Probabilistic Models and Diffusion-based Adversarial Autoencoders. ICLR submission, 2022. [PDF]

[4] Z. Wang **H. Zheng**, P. He, W. Chen and M. Zhou. Diffusion-GAN: Training GANs with Diffusion. ICLR submission, 2022. [PDF]

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

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

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

[8] 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]

[9] **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. Preprint, 2021. Under review of JMLR. [PDF]

[10] **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]

[11] 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).

[12] 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)

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

[14] 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]

[15] 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]

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

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

[18] **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]

[19] **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]

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

[21] **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]

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

PROJECTS

Truncated diffusion probabilistic models September 2021 - Present

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

- We propose a faster and cheaper 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.
- Experimental results show 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.
- The idea of truncated diffusion is able to be extend to several

Large-scale self-supervised pretraining May 2022 - August 2022

Internship at Microsoft

- We propose to investigate the relation between contrastive learning and the pretraining of masked image models (MIM). The encoders trained with contrastive learning are encouraged to minimize the representation distribution of different views, and show its merits in linear probing. MIMs are inspired by the success of language processing, where the mask language model successfully pretrain the model to provide powerful representations withe the masked tokens. The MIMs are known for being powerful in downstream fine-tuning, but the linear probing does not show satisfactory performance compared with contrastive learning.
- We further study if the pretraining techniques of these two methods could enhance each other. In our initial studies, we observe the contrastive loss and the mask language loss could help each other to provide better model performance.
- 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. Moreover, the pretraining does not require as many epochs as previous SOTA methods like MAE. Our results show the downstream fine-tuning is improved by 0.6% and the linear probing is improved by 5% compared with MAE with ViT-B/16 backbone.

Contrastive Conditional Transport for Representation Learning October 2020 - June 2022

University of Texas at Austin, advised by Prof. Mingyuan Zhou

- We propose to investigate the contrastive learning in a view of the conditional transport. An intuition is the feature extractor is encouraged to minimize the transport cost of the positive samples and maximize the transport cost of the negative samples.
- We propose a new objective function which learns two conditional distributions to determine the importance of positive and negative pairs for the cost computation. These two learnable conditional distributions encourage the learner to capture local connectivity of positive and 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. For example, on the ImageNet dataset, our model can be adapted on Moco/Moco v2 framework and improve the performance by around 2%.
- We recently updated our method in Moco v3 framework and achieve new SOTA results. Our method also show good generalization ability of the learned representations. Our method ranks at the 5th place on ELEVATER benchmark while requires only 10% parameter tuning compred to the other methods.
- Our method shows the robustness in learning representations on some imbalanced datasets. For example, we decrease the number of samples in some classes of the datasets for training and keep the same testing set for testing. On this class-imbalanced datasets, most of existing contrastive learning methods underperform than in the normal case to large extent, while our method preserves the effectiveness of learning and shows better performance.

Asymptotic Conditional Transport May 2020 - November 2021

University of Texas at Austin, advised by Prof. Mingyuan Zhou

- We propose asymptotic conditional transport (ACT) as a new divergence to measure the difference between two probability distributions. ACT 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 ACT 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 ACT divergence is shown to consistently improve the performance.

- Apart from generative models, we have already started working on applying ACTs 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 preliminary results on these tasks.

Generative models for the predictions in Markov process October 2020 - May 2021
University of Texas at Austin, advised by Prof. Stephen G. Walker

- We investigate to model natural Monte Carlo and fully nonparametric estimators of multivariate distributions with deep generative models as conditional distribution functions.
- Being able to model using conditional distribution functions we can study a number of problems, such as prediction for Markov processes, imputation of missing entries which depend on covariates, and general multivariate data.
- These aspects arise immediately from the integral theorem, being explicit Monte Carlo based and require no recursive or iterative algorithms.

MCMC interacted Variational Inference August 2019 - October 2020
University of Texas at Austin, advised by Prof. Mingyuan Zhou

- To improve MCMC strategies and Variational Inference, we propose MCMC-interactive variational inference (MIVI) to not only estimate the posterior in a time constrained manner, but also facilitate the design of MCMC transitions.
- We propose to utilize stochastic gradient Langevin dynamics (SGLD) and Gibbs sampling for more flexible posterior modeling and optimization.
- To prevent "over-pruning" and "posterior collapsing" in VAE models, we propose various encoding strategies and gradient estimators to make VAEs more robust and obtain tighter lower bound, as well as better latent code.
- We also propose to generalize the lower bound, and the novel encoding strategy can be applied to the generator in generative adverserial net (GAN), which provides a new way to combine VAEs and GANs.

Learning on Attribute-Missing Graphs July 2018 - July 2019
Shanghai Jiao Tong University, advised by Prof. Ya Zhang

- We study the graph whose nodes partially provide attribute information and existing graph learning methods including the popular GNN cannot provide satisfied learning performance since they are not specified for attribute-missing graphs. By making a shared-latent space assumption, we develop a novel distribution matching based GNN for structure-attribute transform and get promising results, which can benefit numerous real-world applications such as link prediction and node attribute completion tasks.

Unsupervised Image Anomaly Detection in the Light of Information Theory July 2018 - July 2019
Shanghai Jiao Tong University, advised by Prof. Ya Zhang

- We study Variational AutoEncoders (VAEs) and anomaly detection problems in the perspective of information theory. We propose insight of using an information theoretic lower bound for optimization in anomaly detection. With this bound, we can achieve better results than SOTA methods with only normal data.

Improving Variational Autoencoders with Fisher Information July 2015 - July 2018
Shanghai Jiao Tong University, advised by Prof. Ya Zhang

- We explore why the trade-off between representation learning quality and likelihood maximization exists in VAEs, and how to balance the trade-off effectively, through Modeling with Fisher information and Shannon entropy, and with the *Uncertainty Principle*.
- We study on two-stage combination of generative model and deep learning and find out how to learn hierarchical representation using topic model and low-level features extracted with deep learning based method.
- We propose an application in mining descriptive characteristics of a region from social images, using the learned representation.

Aligning Arterial and Venous Phases for CT images segmentation May 2018 - Dec 2018
Johns Hopkins University, advised by Prof. Alan L. Yuille

- We investigate the problem of organ segmentation on multi-phase medical images. In this specific work, we study the abdominal CT scans that often have arterial phase and venous phase, which provide complementary information for the task.
- From the perspective of generative models, we explore the intrinsic gap between the ideal setting and the real world scenario, which often makes it difficult to use multi-phase knowledge.
- We propose a framework that combines knowledge transfer and segmentation to incorporate the useful information from both phases. Our model improves the segmentation result on two-phase and mono-phase data.

ACADEMIC SERVICES

ICLR 2022 reviewer	2022
NeurIPS 2022 reviewer	2022
ICML 2022 reviewer	2022
IEEE Transactions on Neural Networks and Learning Systems reviewer	2022
ICLR 2022 reviewer	2021
ACML 2021 reviewer	2021
Machine Learning Journal reviewer	2021
Neurips 2021 reviewer	2021
ICML 2021 reviewer	2021
ICLR 2021 reviewer	2020

SELECTED HONORS AND AWARDS

Outstanding Graduates of Shanghai (Top 5% in Shanghai)	2019
National Scholarship for Graduate Students	2017
Excellent Teaching Assistant	2017
Outstanding Graduates 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