YIBO YANG

Research Scientist at Chan Zuckerberg Biohub

email: yyang@biohub.org ⋄ website: yiboyang.com

SUMMARY – *generative models, AI/ML for science, compression, representation learning*

I am a research scientist on the AI/ML team at Chan Zuckerberg Biohub, working on scalable and principled machine learning methods for accelerating scientific discovery. I am broadly interested in applications of probability, statistics, and information theory to solving challenging real-world problems. My PhD research pioneered machine-learning-based approaches to data compression and information theory.

EDUCATION

The University of California, Irvine   September 2019 - December 2024
M.S. and Ph.D. in Computer Science; GPA: 3.978
Research Advisor: Prof. Stephan Mandt
Supported by a fellowship from the Hasso Plattner Institute.

The University of Texas at Dallas, Computing Scholars Honors Program   August 2014 - May 2018
B.S. in Computer Science, magna cum laude.
Research Advisor: Prof. Nicholas Ruozzi
Supported by a full-tuition academic excellence scholarship.

PROFESSIONAL EXPERIENCE

Chan Zuckerberg Initiative   Redwood City, California
Research Scientist, AI/ML   *Dec. 2024 - Present*

· Supervisor: Theofanis Karaletsos
· Research area: deep generative modeling for cellular biology and solving inverse problems.

The Vector Institute for AI   Toronto, Canada
Research Intern.   *Fall 2023*

· Hosts: Alireza Makhzani, Roger Grosse.
· Research area: monitoring and steering Transformer language models for AI safety.

Google Research   Mountain View, CA
Research Intern on the neural compression team at Google.   *Summer 2022*

· Hosts: David Minnen, Lucas Theis.
· Research area: instance-adaptive / computationally-efficient neural image compression.

The University of Texas at Dallas   Dallas, TX
Research Assistant, computer science department.   *May 2017 - May 2019*

· Advisors: Nicholas Ruozzi, Vibhav Gogate.
· Research area: scalable inference and learning for probabilistic graphical models; neural network compression.

Johns Hopkins University   Baltimore, MD
Research Intern at the Center for Language and Speech Processing.   *Summer 2016*

· Advisor: Lukas Burget
· Research area: Bayesian nonparametric methods (sampling/VI) for under-resourced speech recognition.

PUBLICATIONS (* denotes equal contribution; † denotes alphabetical order.)

Book

"An Introduction to Neural Data Compression". **Yibo Yang**, Lucas Theis, Stephan Mandt. *Foundations and Trends in Computer Graphics and Vision, 15(2), 113-200*, 2023.

Journals

"Foundations of a Fast, Data-Driven, Machine-Learned Simulator". Jessica N. Howard, Stephan Mandt, Daniel Whiteson, **Yibo Yang**†. *Nature, Scientific Reports*, 2022.

"Insights from Generative Modeling for Neural Video Compression". Ruihan Yang, **Yibo Yang**, Joseph Marino, Stephan Mandt. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023.

Conference Proceedings

"Transformers for Mixed-type Event Sequences". Felix Draxler, Yang Meng, Kai Nelson, Lukas Laskowski, **Yibo Yang**, Theofanis Karaletsos, Stephan Mandt. *Conference on Neural Information Processing Systems (NeurIPS)*, 2025.

"Progressive Compression with Universally Quantized Diffusion Models". **Yibo Yang**, Justus C. Will, Stephan Mandt. *International Conference on Learning Representations (ICLR)*, 2025; awarded oral presentation.

"AstroCompress: A benchmark dataset for multi-purpose compression of astronomical data". Tuan Truong, Rithwik Sudharsan, **Yibo Yang**, Peter Xiangyuan Ma, Ruihan Yang, Stephan Mandt, Joshua S. Bloom. *International Conference on Learning Representations (ICLR)*, 2025.

"Estimating the Rate-Distortion Function by Wasserstein Gradient Descent". **Yibo Yang**, Stephan Eckstein, Marcel Nutz, Stephan Mandt. *Conference on Neural Information Processing Systems (NeurIPS)*, 2023.

"Computationally-Efficient Neural Image Compression with Shallow Decoders". **Yibo Yang**, Stephan Mandt. *International Conference on Computer Vision (ICCV)*, 2023.

"Towards Empirical Sandwich Bounds on the Rate-Distortion Function". **Yibo Yang**, Stephan Mandt. arXiv:2111.12166. *International Conference on Learning Representations (ICLR)*, 2022.

"Hierarchical Autoregressive Modeling for Neural Video Compression". Ruihan Yang, **Yibo Yang**, Joseph Marino, Stephan Mandt. *International Conference on Learning Representations (ICLR)*, 2021.

"Improving Inference for Neural Image Compression". **Yibo Yang**, Robert Bamler, and Stephan Mandt. *Conference on Neural Information Processing Systems (NeurIPS)*, 2020.

"Variational Bayesian Quantization". **Yibo Yang***, Robert Bamler*, Stephan Mandt. *International Conference on Machine Learning (ICML)*, 2020.

"Lifted Hybrid Variational Inference". Yuqiao Chen*, **Yibo Yang***, Sriraam Natarajan, Nicholas Ruozzi. *International Joint Conference on Artificial Intelligence (IJCAI)*, 2020.

"One-Shot Marginal MAP Inference in Markov Random Fields". Hao Xiong *, Yuanzhen Guo*, **Yibo Yang***, and Nicholas Ruozzi. *Uncertainty in Artificial Intelligence*, 2019.

Workshop Papers and Technical Reports

"Estimating the Rate-Distortion Function by Wasserstein Gradient Descent". **Yibo Yang**, Stephan Eckstein, Marcel Nutz, Stephan Mandt. *ICML 2023 Workshop on Neural Compression* (spotlight).

"Autoencoding Implicit Neural Representations for Image Compression". Tuan Pham, **Yibo Yang**, Stephan Mandt. *ICML 2023 Workshop on Neural Compression: From Information Theory to Applications.*

"Lower Bounding Rate-Distortion From Samples". **Yibo Yang**, Stephan Mandt. *ICLR 2021 Workshop on Neural Compression: From Information Theory to Applications* (spotlight).

"Efficient Neural Network Pruning and Quantization by Hard Clustering". **Yibo Yang**, Nicholas Ruozzi, Vibhav Gogate. *AAAI 2019 Workshop on Network Interpretability for Deep Learning.*

"Alternative Inference Methods for Acoustic Unit Discovery Model Training". **Yibo Yang**, Lucas Ondel, Lukas Burget. *Building Speech Recognition System from Untranscribed Data, Report from JHU Workshop 2016.*

TALKS

Recent Advances in Diffusion-Based Generative Compression. Invited talk at ML for Wireless Communications workshop at ICML 2025, June 26, 2025, and at Learn to Compress workshop at ISIT 2025, July 18, 2025.

Estimating the Rate-Distortion Function with Statistical Learning and Optimal Transport, ITA Workshop 2024 (best student presentation award). Feb 21, 2024.

"Data Compression with Machine Learning", guest lecture for CS 274 at UCI. Nov 21, 2023.

Estimating the Rate-Distortion Function with Statistical Learning and Optimal Transport, Stanford University. Nov 3, 2023.

"Data Compression with Machine Learning", tutorial at UAI 2023. July 31, 2023.

"Data Compression with Machine Learning", tutorial at AAAI 2023. February 8, 2023.

"Data Compression with Machine Learning", tutorial at NeurIPS 2022. December 5, 2022.

"Data Compression with Machine Learning", guest lecture for CS 274 at UCI. Nov 22, 2022.

"Towards Empirical Sandwich Bounds on the Rate-Distortion Function", Qualcomm Research. March 1, 2022.

"Towards Empirical Sandwich Bounds on the Rate-Distortion Function", Intel. February 28, 2022.

"Hierarchical Autoregressive Modeling for Neural Video Compression", reading group at Google Research. July 2022.

"Exploring the limits of lossy data compression with deep learning", UCI Center for Machine Learning Seminar. April 26, 2021.

"Discrete Variational Optimization with Stochastic Gumbel Annealing", SoCal ML Symposium. March 22, 2021.

MISCELLANEOUS

Professional Service
· Co-organizer for the ICML 2023 and NeurIPS 2024 workshops on machine learning and compression.
· Speaker/co-organizer for the tutorial on Data Compression with Machine Learning at NeurIPS 2022, AAAI 2023, UAI 2023. See `https://neuralcompression.github.io/`.
· Regular reviewer for ICML, NeurIPS, and ICLR since 2020. *Outstanding Reviewer Awards* at NeurIPS 2020, 2021, and ICLR 2021.

Computing
Python (Numpy, Tensorflow, Jax, PyTorch, Cython), C/C++, MATLAB.

References
Dr. Stephan Mandt, associate professor at UC Irvine, mandt@uci.edu
Dr. Daniel Whiteson, professor at UC Irvine, daniel@uci.edu
Dr. Alireza Makhzani, faculty member at the Vector Institute, makhzani@vectorinstitute.ai
Dr. Robert Bamler, professor at University of Tübingen, robert.bamler@uni-tuebingen.de