

Yubei Chen

Postdoctoral Associate, Center for Data Science, NYU

yubeichen@nyu.edu
510-332-0359
<https://yubeichen.com>

Research Interest

My work is at the intersection of computational neuroscience and deep unsupervised learning—enhancing our understanding of the computational principles governing unsupervised representation learning in both brains and machines, and reshaping our insights into natural signal statistics.

Keywords: Unsupervised (self-supervised) representation learning, theoretical neuroscience, natural signal statistics, energy-based models, sensorimotor representation learning, efficient processing

Education

- Ph.D. in Electrical Engineering and Computer Science, UC Berkeley** Dec. 2019
with Designated Emphasis in Communication, Computation, and Statistics
- Major: artificial intelligence; Minor: signal processing & mathematics
 - Advisor: Professor Bruno Olshausen
 - Thesis: *The Sparse Manifold Transform and Unsupervised Learning for Signal Representation*
 - Committee: Professor Bruno Olshausen, Professor Michael Lustig, Professor Pieter Abbeel
- M.S. in Electrical Engineering and Computer Science, UC Berkeley** Dec. 2018
- M.A. in Mathematics, UC Berkeley** May 2015
- Committee: Professor Marc A. Rieffel, Professor John Lott
- B.E. in Electrical Engineering, Tsinghua University, Beijing** Jul. 2012
- Advisors: Professor Yangdong Deng, Professor Wenguang Chen

Professional and Research Experience

- Center for Data Science, NYU** Oct. 2022 – Present
- Postdoctoral Associate, Advisor: Professor Yann LeCun
 - Topic: Hierarchical white-box unsupervised learning, geometric self-supervised learning for world models, self-supervised learning for science
- Meta - Fundamental AI Research; Visiting Center for Data Science, NYU** Oct. 2020 – Oct. 2022
- Postdoctoral Researcher, Advisor: Professor Yann LeCun
 - To construct minimalistic unsupervised learning model from the neural and statistical principles
 - To build theory for state-of-the-art deep self-supervised learning by simplification, unification and visualization
- Aizip, Inc. Cupertino, CA** Dec. 2019 – Oct. 2020
- Co-founder; Participated in building a core team and key IP portfolio for company to grow
 - To build robust, efficient, and scalable real-world AI-IoT (AIoT) solutions
 - Under semi-stealth model, achieved volume shipping and break-even
 - Full-stack machine learning experience
- Redwood Center for Theoretical Neuroscience & Berkeley AI Research (BAIR) Lab, UC Berkeley** Jan. 2020 – Oct. 2020

- Postdoctoral Associate, Advisor: Professor Bruno Olshausen
- Constructed the simplest known Bayesian mathematical model for building unsupervised factorized representations based on sparse coding and Lie group transformations
- Designed the state-of-the-art neural MCMC sampler

Redwood Center for Theoretical Neuroscience & Berkeley AI Research (BAIR) Lab , UC Berkeley

- Graduate Research Assistant, Advisor: Professor Bruno Olshausen Jan. 2013 – Dec. 2019
- Proposed the sparse manifold transform, which unifies the key ideas from sparse coding, manifold learning, and slow feature analysis
- Proposed a geometric theory of complex cells
- Discovered word factors in different word embedding methods
- Constructed the first energy-based model capable of high-quality natural image sampling
- Proposed orthogonal convolution neural networks

Computer Graphics Lab, Stanford University Jul. 2011 – Sep. 2011

- Stanford UGVR Fellowship, Advisor(s): Professor Ronald Fedkiw
- Implemented tetrahedron-based deformable mesh with multi-threading in PhysBAM.

HPC Embedded Computing Lab, Tsinghua University Mar. 2010 – Jun. 2012

- Undergraduate Research Assistant, Advisor: Professor Yangdong Deng
- Designed Hermes, an integrated CPU/GPU microarchitecture for IP routing
- Optimized task-level pipeline GPGPU architecture

PACMAN Group, Institute of HPC, Tsinghua University Jun. 2010 – Mar. 2012

- Undergraduate Research Assistant, Advisors: Professor Wenguang Chen
- Participated in the development of MapCG, the first MapReduce framework on GPU.
- Implemented a set of highly tuned benchmark algorithms with CUDA

Selected Honors and Awards

Outstanding Paper Honorable Mentions Award 2023

- The International Conference on Learning Representations (ICLR)

First Place Winner in Object Detection Track 2020

- IEEE Low Power Computer Vision Challenge (LPCVC)

NSF GRFP Fellowship 2012 – 2015

- National Science Foundation, USA

Chair's Excellence Scholarship Dec. 2012

- Department of Electrical Engineering and Computer Science, UC Berkeley

Undergraduate Visiting Research Fellowship (UGVR) 2011

- School of Engineering, Stanford University

Comprehensive Scholarship for Academic Excellence 2011 – 2012

- Tsinghua University, Beijing, China

Scholarship for Science, Technology, and Innovation 2009 – 2011

- Tsinghua University, Beijing, China

Second Prize Winner (the Youngest Winner)

2010

- NVIDIA China University GPU Programming Contest

Finalist in Chinese National Physics Olympiad

2007

- The Chinese Physical Society

Grant Writing

This section lists all the proposals I have been involved with, where I participate in regular discussions.

MURI: Neuro-Inspired Distributed Deep Learning (Under Review)

2022

- Title: Emergence of Structured Memory via Compression and Self-Consistency

- PI: Yi Ma, Doris Tsao, Stefano Fusi, John Wright, Rob Fergus, Yann LeCun

- My Contribution: drafting, editing, and reviewing every section of the proposal, with a particular emphasis on Task 1 and Task 2. Total percentage of my contribution: 15% - 20%.

- Budget: \$1,250,000/year

NSF: the Mathematical and Scientific Foundations of Deep Learning (Not Awarded)

2021

- Title: A Lie Theoretic Framework for Deep Learning

- PI: Bruno Olshausen, Stella Yu, Nina Miolane

- My Contribution: drafting, editing, and reviewing every section of the proposal, which a particular emphasis on Aim 3. Total percentage of my contribution: 15%.

- Budget: \$1,200,000 for 3 years

ONR: Science of Artificial Intelligence (Not Awarded)

2019

- Title: Bidirectional computational models of vision

- PI: Christopher Rozell and Bruno Olshausen

- My Contribution: drafting, editing, and reviewing every section of the proposal, which a particular emphasis on Thrust 1. Total percentage of my contribution: 33%.

- Budget: \$1,700,000 for 4 years

NSF: Information & Intelligent Systems, (Award IIS-1718991)

2017

- Title: Extracting and understanding sparse structure in spatiotemporal data

- PI: Friedrich Sommer, Bruno Olshausen, and Saeed Saremi

- My Contribution: drafting, editing, and reviewing every section of the proposal, with a particular emphasis on Aim 2 and Aim 3. Total percentage of my contribution: 40%.

- Award: Granted \$500,000, 9/1/2017 – 8/31/2021

I believe the scope of my research will be diversely supported by agencies such as NSF, ONR, DoD, and NIH. In particular, my work fits particularly well to the current emphasis of NSF-IIS/RI and NSF-CCF/CIF.

Teaching Experience

Vision Science 265 Neural Computation, UC Berkeley Fall 2016
Teaching Assistant

- Instructor: Professor Bruno Olshausen
- I gave guest lectures on manifold learning and unsupervised learning; graded and designed the projects; led discussion sessions and hold regular office hours.
- I co-organized the restructuring of the course to include new advancements in this field, designing the new curricula, partition and distribution of different tasks, and creating the GitHub repository of the refreshed course.

Computer Science 188 Introduction to Artificial Intelligence, UC Berkeley Fall 2015
Teaching Assistant

- Instructors: Professor Stuart Russell & Dr. Pat Virtue
- I gave regular discussion lectures on a variety of topics including Bayesian inference, sampling, MDP & optimal control etc; designed new projects and graded exams; hold regular office hours. I was also invited as a final exam grader for CS188 in the Spring 2016 semester.

CRCNS Course on Mining and Modeling of Neuroscience Data, UC Berkeley Jul.06-17 2015
Teaching Assistant

- Instructors: Prof.Frederic Theunissen, Prof. Rob Kass, Prof. Maneesh Sahani, Prof. Odelia Schwartz, Prof. Sonja Gruen, Prof. Eric Shea-Brown, Prof. Christos Papadimitriou
- Teaching Assistant: I helped with the daily problem and coding sessions, coordinated logistics, the invited talks, and events with the organizers.

[K12] Advising Wevin AI Club at Saratoga High School and Los Altos High School, CA 2018 - 2020
Research Mentor; I supervised 4 high schoolers on Animal Voice Recognition

- I gave regular lectures on speech recognition, deep neural network training, data collection and cleaning, etc. I also coach my students with hands-on AI coding and debugging sessions.
- My students built the SOTA cat voice recognition system and presented a poster at International Conference on Statistical Language and Speech Processing (SLSP), 2020

Student Advising and Mentorship

- [12] Shengbang Tong, NYU (CDS undergraduate summer research, co-advised with Prof. Yann LeCun)
May 2022 - Sep 2022
- [11] Martin Yuqi Zhai, UC Berkeley (Undergraduate research assistant, co-advised with Jiayun Wang)
May 2022 - Present
- [10] Vijay Veerabadran, Facebook AI Research (Ph.Dc., Research Internship, co-advised with Prof. Yann LeCun and Dr. Stephane Deny)
Jun. 2021 - Aug. 2021
- [9] Juexiao Zhang, NYU (M.Sc. graduate research assistant, co-advised with Prof. Chen Feng)
Dec. 2021 - Present
- [8] Chun-Hsiao Yeh, Academia Sinica Taiwan & UC Berkeley (Ph.D student, role: research mentor)
Sep. 2020 - Dec. 2021
- [7] Tejasvi Kothapalli, UC Berkeley (Undergraduate research assistant, co-advised with Prof. Stella Yu)
Sep. 2020 - May 2021

- [6] Zeyu Yun, UC Berkeley (Undergraduate research assistant, co-advised with Prof. Bruno Olshausen) May 2020 - Present
- [5] Kaylo Littlejohn, UC Berkeley (Ph.D. student, EECS research lab rotation) May 2020 - Aug 2020
- [4] Ho Yin Chau, UC Berkeley (Undergraduate research assistant, co-advised with Prof. Bruno Olshausen) Sep. 2019 - Dec. 2020
- [3] Jiayun Wang, UC Berkeley (Ph.Dc. in Vision Science, role: research mentor) Sep. 2019 - Present
- [2] Juexiao Zhang, UC Berkeley (Redwood undergraduate summer research, co-advised with Prof. Bruno Olshausen) May 2019 - Aug 2019
- [1] Noah Gundotra, UC Berkeley (BAIR undergraduate mentoring program) Sep.2018 - Nov. 2019

Professional Service

Reviewer for:

- Transactions on Machine Learning Research 2022
- IEEE Transactions on Signal Processing 2022
- Journal of Applied and Computational Topology 2022
- Neural Computation 2021 - 2022
- European Conference on Computer Vision (ECCV) 2022
- Committee member, NeurIPS NeurReps Workshop 2022
- IEEE/CVF Winter Conference on Applications of Computer Vision (WACV) 2022
- IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) 2021 - 2022
- NeurIPS SVRHM Workshop 2020, 2022
- AAAI Conference on Artificial Intelligence (AAAI) 2020 - 2021
- International Conference on Learning Representations (ICLR) 2019 - 2021
- International Conference on Machine Learning (ICML) 2019 - 2021
- ICML Human-in-the-Loop Learning (HILL) 2019
- Advances in Neural Information Processing Systems (NeurIPS) 2018 - 2022

TinyML Datasets & Benchmarking Technical Program Committee 2022 - Present

NSF GRFP Fellowship Ambassador, UC Berkeley 2012 - 2014

- Resource person to help senior undergraduate and junior graduate students prepare for the NSF GRFP Fellowship
- Outreaching to different student groups to share the funding opportunities and my experience

Publications

(P): Preprint; (C): Peer-Reviewed Conference; (J): Journal; (A): Abstract; (Pre): In preparation;
 Authorship convention in the field: lead senior authors is listed last; lead junior author is listed first.
 *: Equal contribution as the first author; †: Equal contribution as the lead senior author

- [32] (Pre). Zeyu Yun, Juexiao Zhang, Bruno Olshausen, Yann LeCun, **Yubei Chen**, *Unsupervised representation learning without topology and stationarity*. (In preparation for Nature Machine Intelligence)

- [31] (Pre). Jiayun Wang*, **Yubei Chen***, Yann LeCun, Stella Yu, *Geometry-aware self-supervised learning*. (In preparation)
- [30] (Pre). Yunhui Gao, **Yubei Chen**, Stella Yu, *The emergence of prototypicality: unsupervised feature learning in hyperbolic space*. (Submitted to the International Conference on Computer Vision (ICCV), 2023)
- [29] (P). Shengbang Tong*, **Yubei Chen***, Yi Ma, Yann LeCun, *EMP-SSL: Towards Self-Supervised Learning in One Training Epoch*. (Submitted to the International Conference on Computer Vision (ICCV), 2023)
- [28] (P). Bobak Kiani, Randall Balestriero, **Yubei Chen**, Seth Lloyd, Yann LeCun, *Joint-embedding self-supervised learning in the kernel regime*. arXiv preprint: 2209.14884 (Submitted to The International Conference on Machine Learning (ICML), 2023)
- [27] (P). Shengbang Tong*, Xili Dai*, **Yubei Chen**, Mingyang Li, Zengyi Li, Brent Yi, Yann LeCun, Yi Ma, *Unsupervised learning of structured representations via closed-loop transcription*. arXiv preprint: 2210.16782
- [26] (P). **Yubei Chen***, Adrien Bardes*, Zengyi Li, Yann LeCun, *Intra-instance VICReg: bag of self-supervised image patch embedding explains the performance*. arXiv preprint: 2206.08954
- [25] (P). Zengyi Li, **Yubei Chen**, Yann LeCun, Friedrich T. Sommer, *Neural manifold clustering and embedding*. (In revision for Journal of Machine Learning Research, (JMLR))
- [24] (C). **Yubei Chen**, Zeyu Yun, Yi Ma, Bruno Olshausen, Yann LeCun, *Minimalistic unsupervised learning with the sparse manifold transform*, International Conference on Learning Representations, (**ICLR**), 2023 (**Spotlight, notable-top-25%**)
- [23] (C). Quentin Garrido, **Yubei Chen**, Adrien Bardes, Laurent Najman, Yann Lecun, *On the duality between contrastive and non-contrastive self-supervised learning*, International Conference on Learning Representations, (**ICLR**), 2023 (**Oral, notable-top-5%, Outstanding Paper Honorable Mentions Award**)
- [22] (C). Pu Hua, **Yubei Chen**[†], Huazhe Xu[†], *Simple Emergent Action Representations from Multi-task Policy Training*, International Conference on Learning Representations, (**ICLR**), 2023
- [21] (C). Ho Yin Chau, Frank Qiu, **Yubei Chen**, Bruno Olshausen, *Disentangling images with lie group transformations and sparse coding*. Proceeding of Machine Learning Research (**PMLR**) Volume on Symmetry and Geometry in Neural Representations (NeurReps Workshop), Advances in Neural Information Processing Systems (**NeurIPS-W**), 2022
- [20] (C). Jiayun Wang*, **Yubei Chen***, Stella X. Yu, Brian Cheung, Yann LeCun, *Compact and optimal deep learning with recurrent parameter generators*. IEEE/CVF Winter Conference on Applications of Computer Vision (**WACV**), 2023
- [19] (C). Jiayun Wang, Jierui Lin, Qian Yu, Runtao Liu, **Yubei Chen**, Stella X. Yu, *3D shape reconstruction from free-hand sketches*. The 2nd Workshop on Drawings and abstract Imagery: Representation and Analysis (DIRA), European Conference on Computer Vision (**ECCV-W**), 2022 (Full-paper Spotlights)
- [18] (J). Hong-Y Hu, Dian Wu, Yi-zhuang You, Bruno Olshausen, **Yubei Chen**, *RG-Flow: A hierarchical and explainable flow model based on renormalization group and sparse prior*. Machine Learning: Science and Technology, (**MLST**), 2022.
- [17] (C). Yunhui Guo, Xudong Wang, Yubei Chen, Stella X. Yu, *Clipped hyperbolic classifiers are super-hyperbolic classifiers*. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (**CVPR**), 2022.
- [16] (C). Chun-Hsiao Yeh, Cheng-Yao Hong, Yen-Chi Hsu, Tyng-Luh Liu, **Yubei Chen**, Yann LeCun, *Decoupled contrastive learning*. European Conference on Computer Vision (**ECCV**), 2022.

- [15] (C). Zeyu Yun*, **Yubei Chen***, Bruno Olshausen, Yann LeCun, *Transformer visualization via dictionary learning: contextualized embedding as a linear superposition of transformer factors*. Deep Learning Inside and Out (DeeLIO) Workshop (**NAACL-W**), 2021.
- [14] (J). Zengyi Li, **Yubei Chen**, Friedrich T. Sommer, *A neural network MCMC sampler that maximizes proposal entropy*. (**Entropy**), 2021.
- [13] (C). Vincent Lu, Weilin Sun, Aaron Truong, Hermione Bossolina and **Yubei Chen**, *End-to-End Domestic Cat Sound and Emotion Recognition with Deep Neural Networks*. International Conference on Statistical Language and Speech Processing (**SLSP**), 2020.
- [12] (C). Jiayun Wang, **Yubei Chen**, Rudrasis Chakraborty, Stella X. Yu. *Orthogonal convolutional neural networks*. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (**CVPR**), 2020.
- [11] (P). Zengyi Li, **Yubei Chen**, Friedrich T Sommer, *Learning energy-based models in high-dimensional spaces with multi-scale denoising score matching*. arXiv preprint: 1910.07762, 2019
- [10] (P). Juexiao Zhang*, **Yubei Chen***, Brian Cheung, Bruno Olshausen, *Word embedding visualization via dictionary learning*. arXiv preprint: 1910.03833, 2019
- [9] (A). **Yubei Chen**, Dylan M. Paiton, Bruno Olshausen, Friedrich T. Sommer. *A geometric theory for complex cells*. Annual Meeting of Society for Neuroscience (**SFN**), 2019.
- [8] (C). Brian Cheung, Alex Terekhov, **Yubei Chen**, Pulkit Agrawal, Bruno Olshausen, *Superposition of many models into one*. Advances in Neural Information Processing Systems (**NeurIPS**), 2019.
- [7] (C). **Yubei Chen**, Dylan Paiton, Bruno Olshausen, *The sparse manifold transform*. Advances in Neural Information Processing Systems (**NeurIPS**), 2018.
- [6] (J). Cong Wang, Yu Zhao, **Yubei Chen**, Masayoshi Tomizuka, *Nonparametric statistical learning control of robot manipulators for trajectory or contour tracking*. Robotics and Computer-Integrated Manufacturing (**RCIM**), 2015.
- [5] (J). Shuai Mu, Yandong Deng, **Yubei Chen**, Huaiming Li, Jianming Pan, Wenjun Zhang, Zhihua Wang, *Orchestrating cache management and memory scheduling for GPGPU applications*. IEEE Transactions on Very Large Scale Integration (**VLSI**) Systems, 2013.
- [4] (J). Shuai Mu, Dongdong Li, **Yubei Chen**, Yangdong Deng, Zhihua Wang, *Exploiting the task-pipelined parallelism of stream programs on many-core GPUs*. IEICE Transactions on Information and Systems, 2013.
- [3] (J). Chuntao Hong, Dehao Chen, **Yubei Chen**, Wenguang Chen, Weimin Zheng, Haibo Lin, *Providing source code level portability between CPU and GPU with mapCG*. Journal of Computer Science and Technology (**JCST**), 2012.
- [2] (C). Yuhao Zhu, Yangdong Deng, **Yubei Chen**, *Hermes: an integrated CPU/GPU microarchitecture for IP routing*. ACM/EDAC/IEEE Design Automation Conference (**DAC**), 2011.
- [1] (Ph.D. Thesis). **Yubei Chen**, *The Sparse Manifold Transform and Unsupervised Learning for Signal Representation*. UC Berkeley, 2020.

Invited Talks

Seeking the principles of unsupervised representation learning

- Research Seminar at Cognitive Science Department, UCSD Mar. 2023
- Wu Tsai Institute Symposium on Computation and Cognition, Yale University Mar. 2023
- Research Seminar at ECE Department, UC Davis Feb. 2023
- Redwood Center for Theoretical Neuroscience Seminar, UC Berkeley Jan. 2023
- Center for Computational Neuroscience (CCN) Seminar, Flatiron Institute Jan. 2023
- The System Level Design Group (SLD) Seminar, UT Austin Dec. 2022
- University of Washington NeuroAI Seminar Nov. 2022
- UC Berkeley's 34th Annual Bay Area Vision Research Day (BAVRD) Sep. 2022

RG-Flow: hierarchical and explainable flow model based on renormalization group and sparse prior

- Facebook AI Research Annual Conference Dec. 2020

The sparse manifold transform and unsupervised learning for signal representation

- Research seminar at Prof. Yanzhi Wang's group at Northeastern U., ECE Jan. 2021
- Research seminar at Facebook AI Research Mar. 2020
- Research seminar at Prof. Eero Simoncelli's group at NYU, Center for Neural Science Mar. 2020

The sparse manifold transform

- Research seminar at Prof. Bin Yu's group at UC Berkeley, Stats Dept. Mar. 2019
- Vision group research seminar at International Computer Science Institute (ICSI) Feb. 2019
- AI Yanxishe – Online AI Open Lectures Dec. 2018
- Cresta AI research seminar Nov. 2018
- Cerebras systems research seminar Oct. 2018

Will the three basic ideas in signal modeling give us new directions in control theory?

- Mechanical Systems Control (MSC) Lab lunch seminar, UC Berkeley May. 2016