Tianzhe Chu

Web : tianzhechu.com

Email : tianzhechu@gmail.com
Google Scholar: Tianzhe Chu

Education

- **ShanghaiTech University** — Shanghai, China
  B.Eng in Computer Science and Technology; GPA:3.75/4.0 — *Sep 2020 - Jun 2024*
  Selected Courses: Introduction to Machine Learning, Probability and Statistics, Computer Architecture I, Data Structure and Algorithms, Discrete Mathematics, Signals and Systems, Mathematical Analysis.
- **University of California, Berkeley** — Berkeley, CA, US
  Visiting undergraduate in EECS; GPA: 3.87/4.0 — *Aug 2022-May 2023*
  Selected Courses: Deep Learning, Deep Reinforcement Learning, Foundation of Graphics, Applications of Parallel Computing, Computer Vision, Optimization Model in Engineering, Introduction to Artificial Intelligence.

Research Interest

- I'm interested in unsupervised/self-supervised representation learning and interpretable deep learning architectures. My goal is to develop principled learning techniques that model the structures of real-world information at scale, with applications on visual recognition, 3D generation, multimodality, etc.

Publications

(* means equal contribution)

- Yaodong Yu*, **Tianzhe Chu***, Shengbang Tong, Ziyang Wu, Druv Pai, Sam Buchanan, Yi Ma, *Emergence of Segmentation with Minimalistic White-Box Transformers*, Under Review(CPAL 2024), https://arxiv.org/abs/2308.16271.
- **Tianzhe Chu***, Shengbang Tong*, Tianjiao Ding*, Xili Dai, Benjamin D. Haeffele, René Vidal, Yi Ma, *Image Clustering via the Principle of Rate Reduction in the Age of Pretrained Models*, Under Review(ICLR 2024), https://arxiv.org/abs/2306.05272.
- Yaodong Yu, Sam Buchanan, Druv Pai, **Tianzhe Chu**, Ziyang Wu, Shengbang Tong, Benjamin D. Haeffele, Yi Ma, *White-Box Transformers via Sparse Rate Reduction*, **NeurIPS 2023**, https://arxiv.org/abs/2306.01129.

Research Experience

- **Berkeley Artificial Intelligence Research (BAIR) in UC Berkeley** — Berkeley, CA, US
  Undergraduate research assistant advised by Prof. Yi Ma — *Nov 2022-Now*
  - In general, seeking low-dimensional structures of high-dimensional natural signals, i.e. images and languages
  - Empirically investigating the properties of white-box transformers
  - Exploring mathematically interpretable deep learning architectures driven by optimizing sparse rate reduction
  - Pushing the limits of image clustering in the age of pre-trained models

Activities and Awards

- **Outstanding Individual Award as Leader of Social Practice Group** — Enshi, Hubei, China
  Affiliation: ShanghaiTech University — *July 2021*
- **Outstanding Individual Award as Member of Industrial Practice Group** — Shanghai, China
  Affiliation: ShanghaiTech University — *July 2022*
- **Provincial First Prize for 35th National Physics Olympics Competition** — Nanjing, Jiangsu, China
  Affiliation: Suzhou High School of Jiangsu Province — *Sep 2018*

Skills

- **Languages:** Python, C/C++, Matlab, RISC-V, Mandarin(native), Cantonese, English(fluent)
- **Tools:** Pytorch, Jax, Markdown, git, LaTeX

Project Highlights

- **Emergence Properties in White-box Transformers** Berkeley, CA, US
  Mentor: Prof. Yi Ma *Summer 2023*
  - We discover that white-box transformer leads to the emergence of segmentation properties in the network's self-attention maps, solely through a minimalistic supervised training recipe.
  - Qualitatively, supervised white-box transformer(named CRATE) learns (i) explicit attention maps with semantic meanings; (ii) structured patch-wise representations with less spurious correlations.
  - Quantitatively, supervised CRATE, though not trained for segmentation, achieves a much higher segmentation score than supervised ViT.
- **Clustering via Principle of rate reduction and Pretrained model** Berkeley, CA, US
  Mentor: Prof. Yi Ma and Dr. Benjamin David Haeffele *Mar 2023 - May 2023*
  - We propose a novel image clustering pipeline (named CPP) that integrates pre-trained models and rate reduction, enhancing clustering accuracy and introducing an effective self-labeling algorithm for unlabeled datasets at scale.
  - Our pipeline learns a highly clusterable image representation that can be extended to CIFAR-100/ImageNet-1k/LAION-Aesthetic/WikiArt, where few previous methods succeeded in achieving a decent performance.
  - Our pipeline goes beyond deep clustering via proposing solutions for (i) measuring optimal number of clusters; (ii) better image-to-image search; (iii) labeling clusters with semantic meanings.
- **White-box Transformers** Berkeley, CA, US
  Mentor: Prof. Yi Ma *Feb 2023 - May 2023*
  - We develop white-box transformer-like deep network architectures which are mathematically interpretable and achieve performance very close to ViT.
  - The white-box transformer(named CRATE) is designed as an unrolled optimization of the sparse rate reduction objective over layers.