# Xindi (Cindy) Wu

+1 (412)-726-0470 | xindiw@princeton.edu | Website | Scholar | Github | Linkedin

## Education

**Princeton University** — *Princeton, NJ*
Ph.D. student, Computer Science Department, School of Engineering and Applied Science — *Aug. 2022 - Now*
Advisor: Olga Russakovsky

**Carnegie Mellon University** — *Pittsburgh, PA*
Master of Science in Computer Vision, Robotics Institute, School of Computer Science — *Aug. 2020 - Dec. 2021*
Advisor: Deva Ramanan

**Xi'an Jiaotong University** — *Xi'an, China*
Bachelor of Science in Computer Science, Honors Youth Program — *Sept. 2016- July 2020*
Advisors: Jinjun Wang & Pengju Ren

## Publications & Preprints

[1] Corgi: Cached Memory-Guided Video Generation
**X. Wu**, U. Singer, Z. Lin, A. Madotto, X. Xia, PA. Crook, YE. Xu, XL. Dong, S. Moon — *WACV 2025*

[2] ConceptMix: A Compositional Image Generation Benchmark with Controllable Difficulty
**X. Wu***, D. Yu*, Y. Huang*, O. Russakovsky, S. Arora — *NeurIPS D&B 2024*

[3] Vision-Language Dataset Distillation
**X. Wu**, B. Zhang, Z. Deng, O. Russakovsky — *TMLR 2024*

[4] SWE-bench Multimodal: Do AI Systems Generalize to Visual Software Domains?
J. Yang*, C. Jimenez*, ..., **X. Wu**, ..., O. Press — *Arxiv 2024*

[5] CharXiv: Charting Gaps in Realistic Chart Understanding in Multimodal LLMs
Z. Wang, ..., **X. Wu**, ..., D. Chen — *NeurIPS D&B 2024*

[6] Language Models as Science Tutors
A. Chevalier,..., **X. Wu**,..., D. Chen — *ICML 2024*

[7] Pix2Map: Cross-modal Retrieval for Inferring Street Maps from Images
**X. Wu**, K. Lau, F. Ferroni, A. Osep, D. Ramanan — *CVPR 2023*

[8] Ego4D: Around the World in 3,000 Hours of Egocentric Video
K. Grauman,..., **X. Wu**,..., Jitendra Malik — *CVPR 2022*

[9] Toward Learning Robust and Invariant Representations with Alignment Regularization and Data Augmentation
H. Wang, Z. Huang, **X. Wu** and EP. Xing — *KDD 2022*

[10] CryoETGAN: Cryo-electron Tomography Image Synthesis Using Unpaired Image Translation
**X. Wu**, C. Li, H. Wei, H. Deng, J. Zhang and M. Xu — *Frontiers in Physiology Computational Physiology and Medicine, 2022*

[11] Squared l2 Norm as Consistency Loss for Leveraging Augmented Data to Learn Robust and Invariant Representations
H. Wang, Z. Huang, **X. Wu** and EP. Xing — *Arxiv 2021*

[12] Marrying Motion Forecasting and Offline Model-Based Reinforcement Learning for Self-Driving Cars
S. Pande and **X. Wu** — *Preprint 2021*

[13] High Frequency Component Helps Explain the Generalization of Convolutional Neural Networks.
H. Wang, **X. Wu**, Z. Huang, EP. Xing — *CVPR 2020*

[14] Transferable Adversarial Attacks on Deep Reinforcement Learning
X. Pan, Y. Cao, **X. Wu**, E. Zelikman, C. Xiao, Y. Sui, R. Chakraborty, RS. Fearing — *Workshop on Adversarial ML at CVPR 2020*

[15] Reducing Exploitation of Data Idiosyncrasy Helps Robustify Trained Models
**X. Wu**, H. Wang, E. Zelikman, M. Xu and EP. Xing — *Preprint 2020*

[16] Regularized Adversarial Training (RAT) for Robust Cellular Electron Cryo Tomograms Classification
**X. Wu**, Y. Mao, H. Wang, X. Zeng, X. Gao, EP. Xing, M. Xu — *BIBM 2019*

[17] Template-based and Template-free Approaches in Cellular Cryo-electron Tomography Structural Pattern Mining.
**X. Wu**, X. Zeng, Z. Zhu, X. Gao and M. Xu — *Computational Biology, Codon Publications, Brisbane, Australia, 2019*

[18] Deep Self-Paced Learning for Semi-supervised Person Re-identification Using Multi-View Self-Paced Clustering
X. Xin, **X. Wu**, Y. Wang, J. Wang — *ICIP 2019*

[19] Multitask Learning With Enhanced Modules
Z. Zheng, Y. Wei, Z. Zhao, **X. Wu**, Z. Li and P. Ren — *DSP 2018*

## Experience

**Meta Reality Lab** — *Redmond, WA*
*Research Scientist Intern w/ Dr. Shane Moon* — *May. 2023 - Aug. 2023*
- We introduce compositional video generation, a new paradigm for multi-scene T2V generation guided by subject finetuning.

**Robotics Institute - CMU Argo AI Center for Autonomous Vehicle Research** — *Pittsburgh, PA*
*CMU Sponsered Capstone | Research Assistant w/ Prof. Deva Ramanan* — *Jan. 2021 - Jan. 2022*

- Proposed a contrastive cross-modal approach to dynamic street map construction from camera data. Trained the graph encoder and image encoder with a shared latent space building on recent advances in multimodal representation learning.
- Defined a new task and benchmark for map maintenance, evaluating both fidelity and generalization. Demonstrated that this approach has the ability to generalize both to novel observations within a city as well as to unseen cities.

**Snap Inc. Perception Team** *New York, NY*
*Research Intern w/Dr. Alireza Zareian and Dr. Chen Wang* *May 2021 - Aug. 2021*

- Developed a sample-efficient method to generate self-supervised vision and language representations incorporating ideas from CLIP, supporting a variety of downstream zero-shot tasks including classification, object detection, and segmentation.
- Achieved a 24% relative improvement on top-1 ImageNet accuracy over CLIP trained with the Conceptual Captions 3M dataset.

**Facebook AI Research & Carnegie Mellon University, Robotics Institute** *Pittsburgh, PA*
*Research Assistant w/ Prof. Kris Kitani* *Sept. 2020 - Dec. 2020*

- Developed de-identification tool based on object tracking to efficiently de-identify arbitrary objects including faces, license plates, etc., in egocentric video at near real time, allowing 3x faster de-identification than other SOTA methods.

**Megvii Research (Face++)** *Beijing, China*
*Computer Vision Research Intern w/ Banghuai Li* *June 2020 - Sept. 2020*

- Researched & designed few shot learning models built on Detectron2 with metric learning based methods for object detection.
- Implemented mixup data augmentation and contrastive loss to improve the post-Region Proposal Network relation graph.

**Carnegie Mellon University, Language Technology Institute** Pittsburgh, PA
*Research Assistant w/ Haohan Wang* *Apr. 2019 - June 2020*

- Demonstrated a relationship between the frequency spectrum of image data and generalization behavior of CNNs.
- Designed a regularization scheme that penalizes large differences between adjacent components within kernels.

**Carnegie Mellon University, Computational Biology Department** *Pittsburgh, PA*
*Research Assistant w/ Prof. Min Xu* *Mar. 2019 - June 2020*

- Proposed Regularized Adversarial Training to push the decision boundary away from training data while maximizing accuracy on unperturbed examples to improved the robustness of subtomogram SoTA classification models.
- Designed a model to achieve unsupervised image-to-image translation for Cryo-ET images which is stable to train and capable of generating plausibly diverse image samples [10].

**Xi'an Jiaotong University, Institute of Artificial Intelligence and Robotics** *Xi'an, China*
*Research Assistant w/ Prof. Jinjun Wang & Prof. Pengju Ren* *Dec. 2017 - Feb. 2019*

- Introduced a self-paced regularizer to select reliable samples for fine-tuning each CNNs and implemented self-paced clustering.
- Designed an inverse adversarial learning regime that take classifiers to supervise each generator extract discriminate features and take discriminators for regularizing generators to learn complementary features.

## Talks and Poster Presentations

- Corgi: Cached Memory Guided Video Generation *ECCV AI for Visual Arts Workshop, Milan, Oct. 2024*
- Vision-Language Dataset Distillation *ECCV Dataset Distillation Workshop, Milan, Oct. 2024*
- ConceptMix *ECCV Knowledge in Generative Models Workshop, Milan, Oct. 2024*
- Compositional Generation Evaluation *Google Research, New York, July 2024*
- Scaling Down before Scaling Up: Recent Progress on Dataset Distillation *CVPRW Dataset Distillation, Seattle, June 2024*
- Corgi: Compositional Memory-Guided Video Generation *Cornell Tech, New York, Nov. 2023*
- Pix2Map: Cross-modal Retrieval for Inferring Street Maps from Images *CVPR, Vancouver, June 2023*
- Regularized Adversarial Training for Robust Cellular Electron Cryo Tomograms Classification *BIBM, San Diego, Nov. 2019*

## Professional Service

- **Reviewer** - ICLR 25'/24', ICML 24', CVPR 24'/23'/22', ECCV 24'/22', ICRA 24', ACCV 24', Neurips 23', ICCV 23', ICLR 23' Workshop ME-FoMo, Neurips Interpolate Workshop 22', BMVC 20', IJCAI 20'
- **Committee Member** - Diversity, Equity and Inclusion Committee in Robotics Institute, CMU
- **Volunteer** - vGHC(Grace Hopper Celebration of Women in Computing) 2021 Volunteer
- **Panelist** - Robotics Institute MS Student Panel, 2021, Robotics Institute Summer Scholars (RISS) program 2021
- **Co-Host** - Weekly RI Meets! 2021
- **Mentor** - CMU Society of Women Engineeers (SWE) mentoring program 2021

## Teaching

- TA: COS 429 Computer Vision by Vikram V. Ramaswamy and Felix Heide *Princeton, Spring 2024*
- TA: COS 597O Advanced Topics in Computer Science: Deep Generative Models by Adji Bousso Dieng *Princeton, Fall 2023*

## Awards

ECCV 2024 Dataset Distillation Workshop Best Paper Award: Vision-Language Dataset Distillation. *Sept. 2024*