Pengcheng (Ryan) Xu

+1(217)550-1337 | pengcheng.xu@uci.edu | https://github.com/explcre | https://linkedin.com/in/pengcheng-xu-ryan
Google Scholar: https://scholar.google.com/citations?user=LXzbDZEAAAAJ&hl=en
Personal Website: https://pengcheng-xu.com | ChatGPT version of me website: https://pengcheng-xu.com/mychat/
Open to: student researcher (Winter/Spring 2027) | research intern (Summer 2027, extendable) | full-time (after Summer 2027)

EDUCATION

University of California, Irvine Sep 2024 – Summer 2027
Ph.D. Candidate in Computer Science (passed candidacy exam Nov 25, 2026; graduating Summer 2027 or earlier). GPA: 4.0/4.0
Dissertation: *Efficient and Reliable Foundation Models: Knowledge Distillation from Genomic Language Models and On-Policy Distillation for LLM Agents*.
Research Interest: Large Language Model, Reinforcement Learning, On-Policy Distillation, Multi-Turn LLM Agents, Knowledge Distillation, Conditional & Controllable Generative Models, Multi-modal Language Model, Bioinformatics, AI for Science, Large Language Model for Science, DNA Language Model, Single Cell, Spatial Transcriptomics Foundation Model.
Courses: AI in Biology and Medicine(A), Innovation and Entrepreneurship(A), Light and Geometry in Computer Vision(A+), Deep Generative Models(A).

University of Illinois Urbana-Champaign Aug 2022 - Dec 2023
Master's Degree in Computer Engineering, GPA: 3.8/4.0. Research Scientist at AI@UIUC.
Core Courses: Distributed Systems, Communication Network(A+), Computer Vision(A), Advanced Distributed Systems, Transfer Learning(A), Engineering Entrepreneurship(A+), Machine Learning(A), Applied Parallel Programming.

University of Michigan - Shanghai Jiao Tong University Joint Institute Sep 2018 - Aug 2022
Bachelor's Degree in Electrical and Computer Engineering. Minor in Data Science. Outstanding Graduate (School Level).
Core Courses: Data structures and Algorithms, Intro to Computer Organization (A, Teaching Assistant), Computational Methods for Statistics and Data Science (A+), Intro to Operating Systems, Intro to Data Science (A), Undergraduate Research (A+).

PUBLICATIONS

(Accepted to Findings of EACL 2027) Pengcheng Xu, Xinyu Guan. SILICA: Certified Counterfactual Evaluation of Identifiability in Unseen-Language Induction, 2027. [Paper] [Code] [Dataset] [Project page] [Interactive demo]
(arXiv:2609.34036; under review at ICLR 2027; reviewer ratings 3 / 3 / 2 on a 1–4 scale) Wenbo Zhang*, Pengcheng Xu*, Weizhi Du, Jing Zhang, Hengrui Cai. UOPD: Uncertainty-Aware Intervention for On-Policy Distillation of Multi-Turn Agents, 2026. [Paper] [Code] [Interactive demo] (* equal contribution; co-first authors.)
(In preparation for ICML 2027) Pengcheng Xu, Junhao Liu, Peter Tan Dat Le, Haoyi Cai, Bruhati Aarushi Kuchi, Arnav Abhijit Dhariya, Chaoyang Wang, Hokin Deng, Jing Zhang. "DNAThinker: A Reasoning-Grounded Multimodal Foundation Model for Cell-Type-Conditioned DNA Generation, Editing, and Pair Prediction with Mask-Diffusion Reinforcement Learning".
(*Science Advances*: minor revision, accepted in principle) Pengcheng Xu*, Junhao Liu*, Yi Dai*, Kainoa Andrew Nishida, Dongbo Sun, Yutong Lei, Yaqi Hu, Chaoyang Wang, Jing Zhang. OmegaGenome: Toward Better Sub-Million Scale Expert Models from Large Genomic Language Models via Knowledge Distillation. (* equal contribution.)
(Published in the KDD 2026 proceedings) Simon Dongbo Sun, Junhao Liu, Pengcheng Xu, Yaqi Hu, Martin Jinye Zhang, Jing Zhang. MUGO: Differentiable Combinatorial Optimization for Causal Variant Discovery in the Non-coding Genome, pp. 12042–12053.
(Accepted to ICML 2026) Maijunxian Wang, ..., Pengcheng Xu, Hokin Deng. A Very Big Video Reasoning Suite. [Paper]
(arXiv:2610.04795) Xinyu Guan, Zhirong Zhang, Hongyuan Liu, Pengcheng Xu, Yu Sun, Chen Song, Qianyang Zhao. TIMBRE: Teaching Time Series Forecasters to Read, Remember, and Reconcile, 2026.
(Under review at a top-tier AI conference) Pengcheng Xu, et al., Jing Zhang. "Prediction Instability, Not Confidence, Routes Test-Time Compute in Recursive Reasoners and Language Models".
(Under review at a top-tier AI conference) Pengcheng Xu, et al., Jing Zhang. From Wasserstein Distances to Wasserstein Barycenters: Prediction-Space Ensembling in Set-Valued Knowledge Distillation, 2026. [Paper]
(To be submitted) Pengcheng Xu, et al. TexGemma-D1: An LLM Training Pipeline with a Verifiable Symbolic Reward for Scientific Diagram Generation. Pandita AI / diagram.ai.
(Under review at NAACL 2027; ARR overall assessment 4 / 3.5 / 2.5, soundness 4 / 4 / 4) Pengcheng Xu, et al. PACT: Harness-Side, Cache-Friendly Context Management for Long-Horizon Agents, 2026. [Paper]
(Under review at a top-tier AI/CV conference) Pengcheng Xu, et al. RateGS: Learning Motion Rates for 4D Gaussian Forecasting, 2026.
(Under review at a top-tier AI/CV conference; second author) Pengcheng Xu, et al. Counterfactual Tool-Use Distillation for Multimodal Question Answering, 2026. [Paper]
(To be submitted) Pengcheng Xu, et al. "Routed Co-Evolution of a Coding Agent's Weights and Code Harness under One Compute Budget".
(Under review) Yutong Lei, Yaqi Hu, Chaoyang Wang, Pengcheng Xu, Junhao Liu, Henry Cheng, Jing Zhang. SPATREG: A Hierarchical Framework for Histology Image Registration in Spatial Transcriptomics.
(Manuscript in preparation) Xi Li, Ziheng Duan, Diptanshu Sikdar, Pengcheng Xu, Jing Zhang. BRIDGE: Spatially Aware Bimodal Representation Integration with Hierarchical Graphs for Spatial Transcriptomics Resolution Enhancement.
(ACL BioNLP Workshop 2025) Pengcheng Xu, Sicheng Shen, Jieli Zhou, Hongyi Xin. Team XSZ at BioLaySumm2025: Section-Wise

Summarization, Retrieval-Augmented LLM, and Reinforcement Learning Fine-Tuning for Lay Summaries, pp. 275–280.
(ACM BCB 2024, SIGBIO Paper Award at BCB '24) Ziqi Rong, Jinpu Cai, Jiahao Qiu, Pengcheng Xu, Lana X. Garmire, Qiuyu Lian, Hongyi Xin. L2 Normalization and Geodesic Distance for Enhanced Information Preservation in Visualizing High-dimensional Single-cell Sequencing Data, pp. 1–11.
(Accepted to KDD-AIDSH Workshop 2024, Oral) Pengcheng Xu, Tianfan Fu, Wenhao Gao, Jimeng Sun. REINVENT-Transformer: Molecular De Novo Design through Transformer-based Reinforcement Learning, 2024.
(ACL BioNLP Workshop 2024) Jieli Zhou*, Cheng Ye*, Pengcheng Xu, Hongyi Xin. Team YXZ at BioLaySumm: Adapting Large Language Models for Biomedical Lay Summarization, pp. 818–825. (* equal contribution by the first two authors.)
(arXiv:2306.13866v2) Pengcheng Xu*, Jinpu Cai*, Yulin Gao, Ziqi Rong. MIRACLE: Multi-task Learning based Interpretable Regulation of Autoimmune Diseases through Common Latent Epigenetics, 2023. (* equal contribution by the first two authors, as defined in the public v2 source.)
(Unpublished technical report) Pengcheng Xu, Kaiyang Chen, Yuanrui Zhang, Indranil Gupta. Pipe-Dejavu: Hardware-aware Latency Predictable, Differentiable Search for Faster Config and Convergence of Distributed ML Pipeline Parallelism, 2023.
Zhikai Yang, Pengcheng Xu, Dekun Yang, Yufeng Chen, YanCong Ma. Vascular Intervention Training System Based on Electromagnetic Tracking Technology, ICVRV 2020, pp. 269–270.

Research Experience

SILICA: Certified Counterfactual Evaluation of Identifiability in Unseen-Language Induction 2026
Lead author (with Xinyu Guan). Accepted to Findings of EACL 2027; code (GitHub), data (Hugging Face), project page and interactive demo.

- ▷ Introduced SILICA, a benchmark of *certified identifiability* for unseen-language induction: each record publishes an executable Boolean program over a finite lexicon–grammar class, and exhaustive filtering proves whether every evidence-compatible lexicon and grammar agrees on the query (answer) or some disagree (the answer is underdetermined, so the model should abstain).
- ▷ Built a shortcut-audited artifact of 12,288 records over 192 semantically packed templates: each of 6,144 program/evidence states is queried two complementary ways with opposite correct actions at constant declared class size (128) and posterior size (16), so frozen shortcut controls and learned probes score exactly 50%; an independent clean-room enumerator reproduces all labels, payloads, and certificates.
- ▷ Under a silent single-pass protocol, a sealed panel of eight conditions from 0.5B to 32B answers almost independently of whether the answer is determined and completes none of the 192 eight-row route groups; scale only shifts models between the always-answer and always-abstain constants. Matched controls show a visible reasoning trace lifts Qwen3-235B to .671 balanced strict joint accuracy and 11/192 groups, while the same permission does not help Qwen3-1.7B and a scratchpad does not help Qwen2.5-14B/32B.

UOPD: Uncertainty-Aware Intervention for On-Policy Distillation of Multi-Turn Agents 2026 - Present
Co-first author (equal contribution with Wenbo Zhang); carried out most of the experiments. arXiv:2609.34036, under review at ICLR 2027 (reviewer ratings 3 / 3 / 2 on a 1–4 scale, avg. 2.67); code: github.com/onepounchman/UOPD; interactive demo.

- ▷ Co-developed UOPD, an on-policy distillation (OPD) method for multi-turn LLM agents that decides *when* the teacher should intervene: the frozen teacher scores its confidence in each student-proposed action; low-uncertainty turns execute the student action under the standard reverse-KL OPD loss, while high-uncertainty turns execute a sampled teacher action and train the student to imitate it via SFT (forward KL in expectation). An adaptive quantile threshold over recent uncertainty scores tracks a scheduled intervention rate (e.g., 30%→5% for the 3B student), so the schedule sets *how often* the teacher intervenes and uncertainty sets *where*.
- ▷ Carried out most of the paper's experiments, in particular the full ALFWorld and Search-R1 multi-hop QA pipelines, the complete ablation study, and the cross-scale study distilling RL-trained GiGPO-Qwen2.5-7B teachers into Qwen2.5-3B and Qwen2.5-1.5B students against OPD, TCOD-F2B/B2F, and FTB-OPD baselines over three evaluation seeds.
- ▷ UOPD attains the best ALFWorld success rate at both student sizes (3B: 90.0/88.8% seen/unseen vs. OPD 88.1/85.3%; 1.5B: 89.8/86.6% vs. 86.2/84.8%) and the best WebShop score (1.5B: 76.8 vs. 66.3, +15.8% relative; success 58.6% vs. 52.3%), with the fewest interaction turns in 5 of 6 settings; on four Search-R1 multi-hop QA datasets, a 1.5B student distilled from SearchR1-Qwen2.5-7B gains 2.7–8.0 exact-match points over OPD (e.g. Bamboogle 12.8→20.8) while average turns fall from 5.19 to 3.90.
- ▷ Ablations isolate each component: uncertainty-selected intervention beats random intervention at the same scheduled rate (unseen ALFWorld 88.8% vs. 83.8%; WebShop score 82.4 vs. 77.3 with 6.5 vs. 7.4 turns); removing the SFT term costs 8.0 points on unseen ALFWorld; executing teacher corrections lifts the WebShop score from 77.6 to 82.4; and UOPD cuts training time by 13.8%/36.5% relative to FTB-OPD on ALFWorld/WebShop at an 11.5%/14.7% overhead over OPD. A controlled ALFWorld study motivates the design: a single teacher correction at a low-confidence step raises student task success from 14.0% to 25.8%, versus 16.1% at a random step.
- ▷ Engineered the multi-GPU SLURM training and evaluation infrastructure on Trinity-RFT (verl / vLLM / Ray); reproduced the TCOD baseline stack (published teacher results matched to within one episode) and fixed a trajectory history-buffer bug that misaligned the distillation window in the curriculum baselines.

DNAThinker: A Reasoning-Grounded Multimodal Foundation Model for Cell-Type-Conditioned DNA Generation, Editing, and Pair Prediction Feb 2026 - Present
Lead researcher & primary author. Manuscript in preparation; planned submission to ICML 2027.

- ▷ Built DNAThinker, a unified regulatory-DNA foundation model coupling a Qwen3.5-0.8B + LoRA reasoning LLM with an NTv3-

100M-post + MDLM DNA generator through an SD3/FLUX-style cross-attention bridge and delta-residual conditioning path over a 1.5M-row fine-tuned regulatory-DNA corpus.

- ▷ Developed tool-grounded reasoning prompts with FIMO motif scans and per-cell expression context, so rationales cite verifiable transcription factors and JASPAR motifs across T1 enhancer generation, T2 pair prediction, and T3 enhancer editing.
- ▷ Applied strict-holdout grounded-RL post-training whose reward pairs a TF-grounding term with an explicit hallucination penalty, and measured it as a head-to-head against the supervised (SFT) checkpoint: on a scaled 7-cell holdout (n=700) under the refined v2 detector, RL lifts TF-grounding from the SFT baseline of 0.702/0.714/0.799 to 0.919/0.956/0.920 on T1/T2/T3 (e.g. ∼0.80→0.92 on T3 enhancer editing) while cutting hallucinated TFs by 74%/67%/67% – isolating the hallucination-reward RL step, not SFT, as the source of the grounding gain.
- ▷ Introduced vector-level TF-knockdown intervention on the 17,415-d cell-expression slot plus knockdown-aware policy-gradient training, achieving 63% conditional motif depletion on the Mic T3 test slice and 56% depletion / 71% depleting rows across the 7-cell stratified aggregate.
- ▷ Developed MDLM-RL, a policy-gradient framework for masked-diffusion DNA generators with GRPO, DAPO, GSPO, and SVGSPO variants; audited reward, algorithm, capacity, and learning-rate regimes, reporting T1 single-shot FID 8.78 (4.4x better than LEONINE), a 14x edit-tight T3 trunk-only motif lift, and clear saturated-trunk/no-RL regimes.

Prediction Instability, Not Confidence, Routes Test-Time Compute in Recursive Reasoners and Language Models 2026 - Present

Lead researcher. Under review at a top-tier AI conference.

- ▷ Established that *prediction instability* – how much a prediction moves under a label-preserving perturbation – not the model's *confidence*, is the right label-free, training-free signal for routing test-time compute across recursive reasoners (Tiny Recursion Model, HRM) and LLMs (Qwen2.5/3.5, Phi-3.5, Llama-3.1): confident-but-wrong predictions are stable fixed points, so confidence misroutes – entropy-based halting collapses to depth 1.8 / accuracy 0.199 where instability holds 0.859, and rescue-prediction AUROC is 0.88 (instability) vs. 0.48 (confidence) on GSM8k.
- ▷ Introduced a *free* latent-velocity signal that gates the depth axis – certifying on 3/3 tasks whether extra recursion pays (Sudoku alive, ARC/Maze converged) – and showed across 17 test-time per-step interventions that nothing beats taking the deepest step on a converged recursion; composing extended depth with augmentation-consistency gating (ACG) lifts Sudoku-Extreme exact-grid accuracy from 0.82 to 0.992 at depth-96 (+0.12 to +0.18 over the trained model, zero retraining).
- ▷ Built VR-MoR, a training-free Mixture-of-Recursion that routes per-token recursion depth by velocity – 3.8x less compute at near-full accuracy, dominating the uniform-depth frontier by up to +0.108 and matching the model's *trained* ACT halt head with no training, labels, or auxiliary network.
- ▷ Extended the same signal to a generalist→specialist model-capacity cascade (Qwen2.5-3B→7B; always-specialist accuracy at 73% compute, rescue-AUROC 0.83 vs. confidence 0.50) that transfers down the price ladder (3B→7B 0.74, 0.5B→7B 0.62), a unified cross-axis allocator that Pareto-dominates any single axis at every budget, and a routed self-refinement lever that exceeds refine-all at 38% of its cost.
- ▷ Built a realizable online controller (one calibrated threshold, no labels or reward model) that reaches the oracle accuracy–compute frontier (oracle accuracy at ∼14x vs. 66x uniform on a mixed Sudoku+Maze workload; 75% of self-consistency's gain at 44% of its samples on GSM8k); independently reproduced HRM-Text-1B's published benchmark table (5 matches), found and ablated a PrefixLM-masking pitfall inflating naive HF evaluation by +0.36, and beat the 2025 SOTA confidence method DeepConf head-to-head (routing AUROC 0.666 vs. chance).

From Wasserstein Distances to Wasserstein Barycenters: Prediction-Space Ensembling in Set-Valued Knowledge Distillation 2026 - Present

Lead researcher. Under review at a top-tier AI conference.

- ▷ Developed prediction-space ensembling for set-valued knowledge distillation from a frozen 108.7M-parameter single-cell perturbation teacher into independently seeded 3.21M-parameter students; a five-student pool improved the primary Pearson-Δ metric by +0.0179 against an exactly output-size-matched control, improving all 70/70 held-out perturbations (t=13.6).
- ▷ Derived metric-geometry results explaining why affine-invariant, first-moment evaluation limits the direct headroom of reweighting and higher-moment objectives, and when pooled predictions and a W_2 barycenter score equivalently; tested ten objective variants to convergence with same-seed pairing and found seven ties, three losses, and no positive arm in the fixed benchmark.
- ▷ Measured 26% of the converged student's residual as seed-averageable functional variance (R^2=0.999); matched ten independent members with five runs times two stored snapshots at half the training compute, while explicitly separating prediction pooling from cross-basin weight averaging.
- ▷ Validated the method across domains on CIFAR-100: a six-student prediction pool gained +2.42 top-1 and improved calibration, and distilling the pool into one student recovered +1.20 top-1 across 6/6 seeds at single-model inference; also identified seed-derived evaluation and emitted-output-count confounds worth up to 0.0177 and 0.34, respectively.

Routed Co-Evolution of a Coding Agent's Weights and Code Harness under One Compute Budget 2026 - Present

Lead researcher. To be submitted. Public AReaL research branch.

- ▷ Designed routed weight–harness co-evolution, which spends one GPU-hour budget on both a coding agent's weights and its code harness: LoRA GRPO on tasks chosen by their own rollout history, failed attempts routed to a channel where the policy edits the harness under runtime check-and-repair and an exact permutation gate, and an evolved test-time selector that picks one of several sampled attempts.

- ▷ Measured that 46% of rollout compute in an agentic GRPO run goes to groups with no outcome advantage, and that a task's own history predicts informative groups: history-replayed tasks yield them over twice as often as fresh ones.
- ▷ With Qwen3.5-9B on SWE-bench Verified, the selector, fixed in advance, lifts three attempts of the base model, the shared starting adapter and the full method by 3.2, 3.7 and 2.5 points over a random pick; runtime repair made 3/3 harness proposals work, against 0/3 in an earlier study without it.
- ▷ Ran pre-registered end-to-end comparisons on AReaL (multi-GPU SLURM, sandboxed SWE episodes) against a weights-only control at the same RL step, and against the MGM and HGM self-improvement baselines on 110 pre-registered held-out instances.

RateGS: Learning Motion Rates for 4D Gaussian Forecasting 2026 - Present

First author. Under review at a top-tier AI/CV conference; preprint.

- ▷ Identified and formalized a representational limitation of accumulated-pose decoding for long-horizon 4D Gaussian forecasting: sustained rigid motion with unbounded translation drives the pose decoder's inputs out of distribution, and observed-frame supervision does not constrain how motion errors propagate into the future.
- ▷ Proposed RateGS, which separates instantaneous motion prediction from pose accumulation: an autonomous Neural-ODE latent evolver with no explicit time input, a twist-rate decoder that predicts per-Gaussian rigid velocities, and a parameter-free SE(3) integrator that accumulates them into rigid transformations, so sustained constant-twist motion is represented at a latent equilibrium; a training-only Flow/Pose-Constrained (FC) correction propagates observation-driven latent perturbations into twist space.
- ▷ Reaches 35.90 / 39.92 / 28.20 dB extrapolation PSNR on Indoor / Multipart / GoPro (+4.80 / +3.78 / +1.41 dB over ParticleGS) and, under a unified re-training protocol, the highest PSNR on all four benchmarks (+0.15–3.48 dB over the retrained ParticleGS); autonomous twist-rate dynamics with SE(3) integration lift development-set PSNR from 35.95 to 38.20 dB over accumulated-pose decoding.

PACT: Harness-Side, Cache-Friendly Context Management for Long-Horizon Agents 2026 - Present

Lead researcher. Under review at NAACL 2027 (October 2026 ACL Rolling Review cycle); preprint. ARR reviews (three reviewers): overall assessment 4 / 3.5 / 2.5 (avg. 3.33); soundness 4 / 4 / 4; excitement 4 / 3.5 / 2.5 (avg. 3.33); reproducibility 4 / 4 / 4; reviewer confidence 4 / 4 / 4.

- ▷ Designed PACT, a context policy that runs in the agent harness and spends no agent turns on context management: past a threshold, tool results move verbatim into workspace files behind short stubs that are never edited again, so the prompt prefix stays cache-valid; a commit-with-recall variant answers from the archive near the end of the token budget.
- ▷ On held-out BrowseComp-Plus queries, pre-registered tests find PACT non-inferior to the self-managed state of the art (VISTA, run from its authors' code) with 51% fewer uncached prompt tokens, more accurate at VISTA's largest budget (0.56 vs. 0.46) at a third of the cost, and at the default budget with commit-with-recall (0.54 vs. 0.43).
- ▷ Non-inferiority and the saving hold on three more backbones up to 550B parameters and on GAIA's hidden test set; on LOCA-bench, PACT is not detectably less accurate than VISTA at three data scales while the server recomputes 28–36% fewer prompt tokens.

Counterfactual Tool-Use Distillation for Multimodal Question Answering 2026

Second author. Under review at a top-tier AI/CV conference; preprint.

- ▷ Counterfactual Tool-Use Distillation (CTD) trains a lightweight predictor on the quality difference between paired answers that a frozen multimodal model gives with and without each candidate search call's evidence; using only information available before execution, it estimates each call's added value and selects a call or stops under an adjustable cost threshold.
- ▷ On FVQA, InfoSeek, MMSearch, SimpleVQA and LiveVQA, CTD improves macro-averaged accuracy by 2.93–5.75 percentage points across three backbones while using 50.0–52.6% fewer calls than the strongest evaluated baseline for each backbone under a shared two-call limit.

A Very Big Video Reasoning Suite: Public Task Generators and Dataset Construction 2025 - 2026

Contributing researcher. Accepted to ICML 2026; paper PDF: `https://arxiv.org/abs/2602.20159`.

- ▷ Implemented eight public Python prototype data-generator repositories for gear-chain direction, arrow maze, scale-balance tilt, dot quantity comparison, layer-depth ordering, water-level reasoning, binary counter, and odd-one-out.
- ▷ Built prototype generation and validation logic mapping symbolic task states into controllable video prompts/assets and answer labels.
- ▷ Assisted with Wan video-generation-model fine-tuning experiments for dataset generation.

OmegaGenome: Toward Better Sub-Million Scale Expert Models from Large Genomic Language Models via Knowledge Distillation. Mar 2025 - Present

Lead researcher & primary author (I wrote most of the code + paper). *Science Advances*: minor revision after the second round of review; accepted in principle.

- ▷ Designed OmegaGenome, a two-stage framework that fine-tunes a large genomic language model as a teacher and distills it into a compact, task-specific expert; evaluated on 18 classification tasks (epigenetic marks, regulatory elements, splice sites) and a base-resolution 34-track regression benchmark. Across five teachers (7.7M–3.45B parameters) under three-seed evaluation, distilled experts are up to 28,750x smaller, retain up to 98% of teacher performance (95% on average), and consistently outperform size-matched baselines.
- ▷ **CNN experts for classification:** fine-tuned five teachers spanning *k*-mer, byte-pair, and single-nucleotide tokenization – NT-2.5B and the autoregressive Carbon-3B with LoRA (rank 48 / 16), DNABERT-2, Enformer, and Caduceus with full-parameter tuning – and distilled each into a 0.12M-parameter dilated-convolution CNN expert (BPNet architecture, one-hot input, 9 residual dilated

blocks, ∼2 kb receptive field) with a three-term objective: cross-entropy on labels, temperature-softened KL on teacher logits, and MSE feature alignment through a learned linear projection.

- ▷ The strongest expert beats the identical-architecture BPNet baseline on every task (6.8–10.8% overall gain; 3.6–14.2% on enhancer/promoter and splice tasks), and the same 0.12M student recovers 93% of the 3.45B Carbon-3B teacher. Inference is up to 658x faster on GPU (NT-2.5B: 3,652 s vs. 5.6 s over 38,822 test sequences), uses 29–680x less GPU peak memory (11.3 GB vs. 17 MB), and is several-thousand-fold faster on CPU.
- ▷ Benchmarked against KD baselines DKD, DIST, and Logit Standardization with per-method grid search (splice sites, NT teacher: MCC 0.884 vs. 0.880 / 0.864 / 0.863; BPNet 0.760), and ablated loss terms, loss weights, and student size (2K–3.6M parameters). Quantified teacher–student alignment with silhouette, cosine similarity, linear CKA, and logit correlation: distilled embeddings keep the teacher's class separation (silhouette 0.398 vs. teacher 0.420 and baseline 0.199) and logits track the teacher more closely (Pearson 0.777 vs. 0.674), consistently across all five teachers; DKD matches outputs but not features (CKA 0.412 vs. 0.539).
- ▷ **Per-base-pair multi-track regression:** extended OmegaGenome to sequence-to-track prediction – 34 coverage tracks from 22 ENCODE experiments across five assay families (ATAC-seq, histone ChIP-seq, PRO-cap, eCLIP, RNA-seq), 32,768 bp input, per-base coverage over the central 12,288 bp. Reproduced the NTv3-650M teacher (Poisson–multinomial shape + scale loss; mean Pearson 0.606 over 3 seeds, per-track s.d. $\leq$0.003, matching the published model) and distilled it into compact NTv3-architecture students (8M and 100M; joint student with a shared trunk and 34 per-track heads, plus single-track and per-assay specialists).
- ▷ **Correlation-compatible distillation loss:** designed a resolution-aligned track-matching objective – the student's output is adaptive-average-pooled to the teacher's resolution, and each track is z-normalized by its training-set mean and s.d. so high-magnitude assays do not dominate the loss while the scale-invariant per-track Pearson evaluation is unchanged; MSE to the frozen, stop-gradient teacher tracks. Under an identical recipe, distillation lifts joint 34-track Pearson from 0.475 to 0.505 (8M) and 0.518 to 0.545 (100M), beats the baseline on every assay family and all seven single-track experiments (79–91% of the teacher), and improves every per-assay specialist (+0.020 to +0.038). A random-init ladder (4M–300M) gains +0.026 to +0.034 at every size, with distilled random-init students matching pretrained-init baselines; the 8M student is ∼16x faster with 8x less GPU memory than the teacher (136 s / 1.2 GB vs. 2,122 s / 9.3 GB).

MUGO: Differentiable Combinatorial Optimization for Causal Variant Discovery in the Non-coding Genome 2025 - 2026

Contributing researcher (ran benchmark experiments). Published in the KDD 2026 proceedings, pp. 12042–12053.

- ▷ Contributed to MUGO, a framework that recasts non-coding causal-variant discovery as differentiable combinatorial optimization: a sparse variant-selection mask is relaxed with a Gumbel-Softmax distribution and a straight-through estimator, keeping DNA inputs discrete for a frozen sequence-to-signal oracle (Borzoi, 524kb context) while gradient ascent searches the combinatorial variant space, with a 10-head Monte-Carlo consensus for variance reduction.
- ▷ Ran benchmark experiments across 5 molecular modalities (RNA-seq, CAGE, ATAC-, ChIP-, DNase-seq) and 7 tissues against ISM (greedy/exhaustive), saliency maps, random search, CADD and FunSeq2; MUGO delivered the largest predicted expression gain (e.g., 174.8 vs. 58.2 saliency on blood RNA-seq) at near-constant ∼2 min/locus, while a separate memory analysis showed the ensemble fit within 24GB, versus the intractable $O\left(\binom{N}{k}\right)$ subset search of exhaustive ISM.
- ▷ Evaluated prioritized variants against GTEx v8 cis-eQTLs, where MUGO achieved 2.50x enrichment over random versus 1.94/0.97/1.20 for the reported baselines; separately evaluated GWAS Catalog variants, confirmed cross-backbone robustness (Borzoi vs. Enformer, r=0.56), and ran an LD-resolution stress test separating PIP>0.9 causal variants from high-LD (r^2>0.8) proxies.

Team YXZ at BioLaySumm: Adapting Large Language Models for Biomedical Lay Summarization Apr 2024 - June 2024

- ▷ Evaluated prompt-only OpenBioLLM-Llama3-70B, whose test-set LENS score was 81.2052 versus 48.0963 for the BART baseline (68.8% relative improvement); the team ranked first on the shared task's readability measure.
- ▷ Fine-tuned Llama3-8B using LoRA; on the reported validation comparison, LENS increased from 80.1033 to 80.7012, AlignScore from 0.6316 to 0.70628, and SummaC from 0.5282 to 0.5816.
- ▷ Implemented semantic-similarity K-shot prompting; at K=1, the reported automated factuality metrics improved by 15.0% relative for AlignScore and 7.9% relative for SummaC.

REINVENT-Transformer: Molecular De Novo Design through Transformer-based Reinforcement Learning May 2023 - May 2024

Advisors: Tianfan Fu (then tenure-track Assistant Professor at Rensselaer Polytechnic Institute; now tenure-track Associate Professor at Nanjing University) and Jimeng Sun (Professor of Computer Science, UIUC).

- ▷ Implemented a transformer-based autoregressive molecular generator with oracle-feedback reinforcement learning; in the public paper's AUC Top-10 comparison, the method achieved the highest aggregate score (12.197) and the best or tied-best result on 8 of 20 tasks.
- ▷ Applied Oracle-feedback reinforcement learning on the downstream tasks to reach higher performance than pretrained model.
- ▷ Carried out ablation study and investigation into loss curve and conditional probability over the next token as a function of previously chosen ones according to the model.

Hardware-aware Latency Predictable, Differentiable Search for Faster Config and Convergence of Distributed ML Pipeline Parallelism Feb 2023 - May 2023

Advisor: Indranil Gupta, Professor of CS UIUC | Advanced Distributed Systems | Researcher.

- ▷ Implemented a predictive model that considers communication cost, model computational cost, and hardware information to predict

latency and resources of parallel configurations, saving time on pre-profiling before searching the parallel configuration.
- ▷ Proposed a differentiable parallel configuration search space inspired by DARTS, can potentially reach optimal configuration faster than the original dynamic programming.
- ▷ Employed parallel random initialization using sampling algorithms like Bayesian Optimization for faster train loss convergence.

Multi-modal target detection with zero-shot depth estimation, Multi-modal NAS May 2022 - Sep 2022

Shanghai AI Lab | Intelligent Photoelectric Department | Multi-modal Cognitive Computing Algorithm Intern.

- ▷ Literature review on nature/science papers on multi-modal learning, and conference papers on Neural Architecture Search, multi-modal learning, image and sound feature fusion.
- ▷ The model inputs image and 2 audio channels, and outputs the behavior category and distance. Improve the estimation of distance using zero-shot monocular depth estimation like MiDaS added to the model structure.

MIRACLE: Multi-task Learning based Interpretable Regulation of Autoimmune Diseases through Common Latent Epigenetics | Research Assistant Sep 2021 - Dec 2023

Advisor: Hongyi Xin, Associate Professor of UM-SJTU Joint Institute, Shanghai Jiao Tong University.

- ▷ Explored an adaptable, interpretable neural network for learning common DNA-methylation patterns across approximately 480K CpG sites from hundreds of samples.
- ▷ Designed a site–gene–pathway constrained neural architecture whose learned weights can prioritize candidate sites, genes, and pathways for follow-up.
- ▷ Implemented a variational autoencoder with a pathway-level latent embedding shared across disease datasets for multi-task learning.
- ▷ Optimized the reported pretrain–fine-tune scheme, increasing the six-dataset mean accuracy from 84.10 to 87.20 (+3.10 percentage points, approximately 3.7% relative), and contributed to writing the public 2023 arXiv v2 preprint.

Augmented reality simulation of cardiovascular interventional surgery | Research Assistant Mar 2020 - May 2021

Advisor: Lixu Gu, Professor of Biomedical Engineering, Shanghai Jiao Tong University.

- ▷ Developed the framework of an augmented reality surgery training assistant system for medical student and surgery.
- ▷ Predicted the operation trajectory using LSTM and used KD-Tree to calculate the distance for operation safety warning.
- ▷ Displayed vascular model in AR with OpenGL and designed the UI interface to support translation.
- ▷ Used the aruco library in OpenCV to coordinate positioning of the QR code.
- ▷ Published Vascular Intervention Training System Based on Electromagnetic Tracking Technology on ICVRV as second author.

Normal fundus image generation based on Generative Adversarial Network | Research Assistant Apr 2019 - Oct 2019

Advisor: Lisheng Wang, Professor at SEIEE, Shanghai Jiao Tong University.

- ▷ Utilized GAN to design an image generation system to generate normal fundus image for comparisons to identify lesion areas.
- ▷ Used datasets including MESSIDOR-2 for data preprocessing like normalization, contrast enhancement, de-noising, clipping.
- ▷ Implemented a U-Net-based segmentation network to identify and remove fundus blood vessels as part of de-noising process.
- ▷ Built a CycleGAN framework based on segmentation to generate the normal fundus image to identify the lesion area.

PROFESSIONAL EXPERIENCE

Pandita AI Inc. (diagram.ai) | Large Language Model Research Scientist Intern | Verifiable RL for Scientific Diagram Generation July 2026 - Present

Lead author of **TexGemma-D1** (to be submitted): an LLM training pipeline with a verifiable symbolic reward – the Diagram Semantic Verifier (DSV) – for scientific-diagram generation.

- ▷ Designed DSV, the verifiable symbolic reward at the core of TexGemma-D1: an 18-dimensional evaluator (23 with render checks) spanning 11 metric families that maps a TikZ program to a typed relational graph and scores what a diagram *asserts*, not how it looks – exact decision procedures where the semantics are decidable (Schwartz–Zippel genericity, DDAR-style geometric consistency, product-automaton search), explicit abstention where they are not, and a dual-modality witness that localizes each failing aspect to a source line/column *and* a pixel region of the render.
- ▷ Validated the metric without human labels: on 2,904 known-transform pairs from 728 real DaTikZ-V4 items, DSV separates semantics-preserving redraws from semantics-breaking edits at abstention-penalized AUROC 0.785 (95% CI $[.769, .802]$) versus 0.421 for a code-edit-distance baseline (*below* chance), records **0.0% strict inversions over 905,000+ pairwise comparisons** where DreamSim prefers a real corruption 35.1% of the time and code edit distance 56.1%, and beats two VLM judges (Gemini 2.5 Flash, Claude Sonnet 4.5) on the image side.
- ▷ Built the training pipeline: distilled explicit `<think>` reasoning traces into a Gemma-4-12B SFT checkpoint (LoRA, multi-GPU) so the policy plans structure before emitting LaTeX, then post-trained it with DAPO/Dr.GRPO against the DSV+RenderGraph verifiable reward behind a hard compile gate – raising DaTikZ-984 compilation from **61.3% to 91.4%**.
- ▷ Applied the same symbolic reward to a published Qwen2-based 3B DaTikZ SOTA policy (continued post-training): image similarity (DeTikZify-V2) rose from 0.260 to 0.353 and strict compilation from 93.8% to 98.8%, confirming the reward transfers across model families and lifts both correctness and validity.
- ▷ Enforced scientific rigor with matched controls and pre-registered ablations, reporting the ties and nulls: adding an R_{Sim} image-similarity term to DSV+RenderGraph improved matched-description R_{Sim} (0.2674→0.2859) and compilation (94.7%→96.0%) but the gain *vanished* under prompt-distribution shift, and witness-gated per-token credit showed no measurable benefit – evidence that verifiable, distribution-robust reward, not perceptual similarity or handcrafted credit, is what drives the gains.

XtalPi Inc (QuantumPharm Inc) | XAB-Antibody IDD | Research Scientist Intern Aug 2024 - July 2025

- ▷ Designed and implemented AI models for antibody binding affinity prediction. Reproduced training pipelines for binding-ddg-predictor and GearBind, developed data visualizations for the SKEMPI dataset, and established robust dataset splits to prevent leakage by ensuring complexes remain within the same fold. Conducted comparative analyses of focal loss (mutation-based), weighted MSE/BCE, balanced sampling, and differentiable ranking strategies.
- ▷ Developed a balanced sampling strategy for binding-ddg-predictor to ensure equal representation of mutations with and without alanine. Achieved performance improvements of approximately +10% in Pearson correlation and +5% in Spearman correlation.
- ▷ Integrated Carbondesign encoder into the binding-ddg-predictor framework to enhance encoder feature representations. Augmented the SKEMPIv2 dataset by generating multi-mutation variants, increasing sample size by ~10x. Successfully reproduced GearBind and conducted benchmarking experiments on the augmented SKEMPIv2 dataset, demonstrating improved model generalization.

Amazon Web Services | VMware Cloud on AWS (Brio) | Software Engineer Intern | Seattle, Washington May 2023 - Aug 2023

- ▷ Designed UI to automate the workflow to update Rate Card for Brio resource console using React, TypeScript and JavaScript.
- ▷ Integrated UI with backend API using Java, API Gateway and Amazon Lambda. Supported validation and creating and updating rate cards through csv. Preview visualizes modifications and identifies pricing errors. (5k+ lines of code in total).
- ▷ Designed anomaly detection algorithms for disbursement generator to identify usage spikes or subscription anomalies using Java, SQL, Amazon Athena, S3 Buckets, DynamoDB Table. Utilized CloudWatch to create tickets and alarms to engineers.

4Paradigm Co., Ltd | OpenMLDB | GitLink Code Camp 2022 | Open-Source Developer | Shanghai July 2022 - Oct 2022

- ▷ Developed an automated feature engineering pipeline, including feature generation and selection, based on AutoX and OpenMLDB sql with Python (600+ lines of code). See pull request: `https://github.com/4paradigm/OpenMLDB/pull/2381`.
- ▷ Saving data scientists in dozens of companies' weeks of time by automatic feature selection pipeline.
- ▷ The data is transformed by OpenMLDB sql to get time series and statistics features, and then we select the most important K features based on algorithms like Adversarial Validation, GRN feature selection or Reinforcement Learning.

Intel Co., Ltd | DL Model Optimization Department | Deep Learning Software Engineer Intern | Shanghai Nov 2021 - June 2022

- ▷ Implemented new features to Intel® Neural Compressor, Cross Layer Equalization, a data free quantization to rescale different layers' weight range to reduce over 5% in the drop in accuracy after quantization like FP32 to INT8, and edited documents.
- ▷ Studied open-source software like Nvidia Triton Inference Server (around 70k lines of code) and AI Model Efficiency Tool and gave presentations to around 100 colleagues in whole department to introduce the design and technique detail of them.
- ▷ Cooperated to design and implement class of AI inference server software by C++, which enables the team to deploy trained AI models from multiple frameworks and deploy more models on GPU or CPU based infrastructure to simplify AI inferencing.

ShuKun Technology Co., Ltd | R & D Department | Algorithm Intern | Shanghai Dec 2020 - Apr 2021

- ▷ Implemented multi-node and multi-GPU training with horovod framework, NVIDIA clara train sdk, OpenMPI, and NCCL2; 4GPU (2 nodes each 2GPU) can reach 2.5x speed up compared to 1GPU (Discovered that communication is bottleneck).
- ▷ Compared the efficiencies of models like 3D-UNet when multi-node training with different GPU configurations in Python and wrote the training process documents, saving algorithm scientists days in each training job (~0.65n times faster with n nodes).
- ▷ Used Java to add multi-machine and multi-GPU training functions to the company's back-end web page for algorithm scientists.

Teaching Experience

ICS 6B Boolean Logic & Discrete Structures | Professor: Wayne Hayes | Teaching Assistant Winter 2025
University of California, Irvine.

ICS 6D Discrete Mathematics | Professor: Jing Zhang | Teaching Assistant Spring 2025
University of California, Irvine.

VE370 Intro to Computer Organizations | Professor: Gang Zheng | Teaching Assistant Sep 2021 - Dec 2021
University of Michigan - Shanghai Jiao Tong University Joint Institute.

Personal Projects

AgenticVBench: Long-Horizon Video-Agent Tasks Jul 2026 - Present

- ▷ Shipped a 238-minute Minecraft action-ledger benchmark (1,995 timed events) with machine-exact ground truth, deterministic order/time verification, anti-shortcut tests, and three-agent calibration [Issue #74; merged PR #108].
- ▷ Built a 55-minute, 12-race SuperTuxKart telemetry task with reproducible generation, machine-exact ground truth, coverage-hardened scoring, and multi-agent calibration; under upstream review [Issue #73; PR #106].

Photon 47 – Bilingual 3D Browser Game with Multiplayer and a Serverless Postgres Backend (Solo) 2026 - Present
TypeScript/JavaScript, WebGL (three.js), PostgreSQL (Supabase), Deno Edge Functions, Python (Playwright), Blender; live at photon47.com; 2,000+ commits; shipped to 5 platforms.

- ▷ Designed and built, solo, a free no-install 3D creature RPG and multiplayer arcade (English/Chinese) with seven game modes sharing one roster of 45 playable heroes and 6 bosses – a 7-chapter open-world story campaign, 5v5 MOBA, first-person arena, 285-card real-time deck battler, auto chess, turn-based squad brawl, and online ELO-matched 3D chess with 10 AI levels – plus an arcade and a living homestead hub, with cross-play PvP, cross-device cloud saves, and full touch controls.
- ▷ Designed a 13-table PostgreSQL/Supabase backend with 38 `SECURITY DEFINER` stored procedures (profiles, leaderboards, ELO

matchmaking, payments, analytics), secured with Row-Level Security and idempotent migrations; concurrency-safe matchmaking via `SELECT ... FOR UPDATE SKIP LOCKED` and a transactional ELO update (K=32).

- ▷ Built an event-analytics layer – append-only JSONB event ingestion plus time-windowed SQL rollups (`GROUP BY` / `COUNT(*) FILTER`) for mode popularity, character pick rates, item sales, and per-mode win rates.
- ▷ Shipped a Stripe payments pipeline as Deno/TypeScript Edge Functions with JWT/JWKS-verified checkout and an idempotent, signature-verified fulfillment webhook, deployed via an env-driven test/live script.
- ▷ Built a verification layer of ∼600 scripted Playwright/Node checks (including GPU rendering runs on SLURM) behind a one-source release pipeline for web, Steam, CrazyGames, Poki, and itch.

COURSE PROJECTS

CS 425 Distributed Systems: Distributed ML Inference System (C++ and Python) (100/100) Aug 2022 - Dec 2022

- ▷ Implemented scheduling algorithm to obtain fair-time inference, making each ML job query rate within 20% difference.
- ▷ Design a distributed file system to maintain datasets with SWIM-like membership protocol (nearly 1e-2 false positive rate).
- ▷ Supported fault-tolerance and can recover when N members and coordinator leave or fail. Leader re-election is within 10s.

CS 543 Computer Vision (A): (Python) Aug 2022 - Dec 2022

- ▷ Channel Alignment: Designed multi-scale algorithm to align 3 channels of images, evaluating with NCC and Fourier Transform.
- ▷ Laplacian Blob Detection: Implemented a scale-space blob detection with a Laplacian scale space using Laplacian of Gaussian filter and performed non-maximum suppression in scale space.
- ▷ Stitching: Stitching pairs of images using OpenCV, SIFT descriptors and RANSAC to estimate homography mapping of images.
- ▷ 3D Shape from shading: Estimate the albedo and surface normal and compute the surface height map by integration to reconstruct 3D surface from image of shading.

CS 438 Communication Network (A+): TCP Protocol (C++) Aug 2022 - Dec 2022

- ▷ Implemented transport protocol with properties equivalent to TCP based on unreliable UDP, which can tolerate packet drops, allow other concurrent connections fair chance (0.5x to 2x of TCP), doesn't give up entire bandwidth to other connections.

Sonar detection water pipe state system based on machine vision and neural network compression Sep 2021 - Dec 2021

ECE4710 Advanced Embedded Systems (MDE, Thesis). | Instructor: Zou An, Assistant Professor of UM-SJTU Joint Institute.

- ▷ Processed the source data of sonar as a matrix and use Python to generate polar images in real time.
- ▷ Used OpenCV to perform subtraction and corrosion operations to complete hole or bulge detection.
- ▷ Designed a compressed and binarized YOLO architecture to realize defect area detection with accuracy 99%.
- ▷ Used HLS to deploy to FPGA, matrix parallel operation, real-time embedded system is implemented.

HONORS & AWARDS

Kaggle ARC Prize 2026: ARC-AGI-3 – Bronze medal tier; rank 310 of 4,923; public score 32.53 (team "Ryan #3," account "explcre"; provisional public leaderboard, Nov 14, 2026) 2026

Shanghai Jiao Tong University Outstanding Graduate (School Level) Aug 2022

Microsoft Imagine Cup 2021, China Regional Final – Third Prize Jan 2021

Mathematical Contest in Modeling – Meritorious Winner, team 2006402 [paper PDF] [official results] Apr 2020

"Jidong Cup" CCVR China Virtual Reality Competition, Future Virtual Reality Product Group – Second Prize Sep 2020

Academic Year Undergraduate Excellence Scholarship (2018–2019 & 2020–2021) Nov 2019 / 2021

SKILLS

▷ Programming: C/C++ (Proficient), Python (Proficient), MATLAB, R, Verilog, Java, RISC-V Assembly, SQL, Shell, TypeScript, CUDA Programming.

▷ Frameworks & Libraries: PyTorch, TensorFlow, horovod, Keras, Sk-learn, Pandas, NumPy, OpenCV, Matplotlib, Selenium.

▷ Developer Tools: Docker, Git, LaTeX, PSpice, Linux.

▷ Languages: English (Fluent, Full Professional Proficiency), Chinese (Native Speaker), German (Beginner).