

Yan Ru Pei

yanrpei@gmail.com | <https://github.com/PeaBrane>

ACADEMIC BACKGROUND

University of California, San Diego
PhD in Physics (Computational Science Specialization)

Advisor: *Massimiliano Di Ventra*
Sept 2017 – Sept 2021

University of California, Los Angeles
BS in Physics and Applied Mathematics (summa cum laude)

Advisor: *Robert Cousins*
Sept 2014 – Sept 2016

PUBLICATIONS/PREPRINTS ([GOOGLE SCHOLAR](#))

Physics:

- **Yan Ru Pei**, M. Di Ventra. (2022). Non-equilibrium criticality and efficient exploration of glassy landscapes with memory dynamics. *Physica A*.
- **Yan Ru Pei**, M. Di Ventra. (2022). A Finite-temperature Phase Transition for the Ising Spin-glass in $d \geq 2$. *preprint*.
- **Yan Ru Pei**, Olivier Coenen. (2024). TENNs-PLEIADES: Building Temporal Kernels with Orthogonal Polynomials. *preprint*.

Artificial Intelligence:

- **Yan Ru Pei**, Haik Manukian, M. Di Ventra. (2020). Generating Weighted MAX-2-SAT Instances with Frustrated Loops: an RBM Case Study. *Journal of Machine Learning Research*.
- Haik Manukian*, **Yan Ru Pei***, Sean Bearden, M. Di Ventra. (2020). Mode-assisted unsupervised learning of restricted Boltzmann machines. *Nature Communication Physics*.
- **Yan Ru Pei**, Sasskia Brüers, Sébastien Crouzet, Douglas McLelland, Olivier Coenen. (2024). A Lightweight Spatiotemporal Network for Online Eye Tracking with Event Camera. *CVPR*.
- Zuowen Wang, Chang Gao, Zongwei Wu, Marcos V. Conde, Radu Timofte, Shih-Chii Liu, Qinyu Chen, Zheng-jun Zha, Wei Zhai, Han Han, Bohao Liao, Yuliang Wu, Zengyu Wan, Zhong Wang, Yang Cao, Ganchao Tan, Jinze Chen, **Yan Ru Pei**, et al. (2024). Event-Based Eye Tracking. AIS 2024 Challenge Survey. *CVPR*.
- **Yan Ru Pei**, Ritik Shrivastava, FNU Sidharth. (2024). Real-time Speech Enhancement on Raw Signals with Deep state-space Modeling. *preprint*.

Other:

- **Yan Ru Pei**, Fabio L. Traversa, Massimiliano Di Ventra. (2019). On the Universality of Memcomputing Machines. *IEEE TNNLS*
- Yuki Wang, **Yan Ru Pei**. (2019). The Optimal Deterrence of Crime: A Focus on the Time Preference of DWI Offenders. *preprint*.

RESEARCH EXPERIENCE

Brainchip Inc. (neuromorphic AI startup)

Senior ML Research Engineer

Laguna Hills, CA

Oct 2021 – present

- **Developed a SOTA spatiotemporal network for computer vision at the edge** [[Paper](#)]
 - * Designed a spatiotemporal neural network interfacing with RGB and event cameras for efficient online object detection, [object tracking](#), and monocular depth estimation.
 - * Implemented complex data pipelines (pre-processing event signals) and wrote vectorized routines for object detection in functorch for evaluating the network, eliminating all CPU bottlenecks.
 - * Applied the network on the Prophesee event-camera road-scene dataset, and achieved a 30% increase in mAP over the previous SOTA network (ConvLSTM) with **100x fewer** parameters and compute.
 - * Combined the network with GNN and applied it for open-loop and closed-loop motion planning on the NuPlan dataset, reaching near first-place solutions on the leaderboard.
 - * Prepared four papers and three patents for an in-house hardware implementation of the network (Akida 2.0).
- **Developed a deep state-space model for sequence prediction**
 - * Designed a network that can be dual-configured as convolution and recurrent networks, with **a million times less parameters and compute** than transformer models for NLP tasks. Integrated concepts from modern SSMs (e.g. H3, hyena, [mamba](#)) to build an LLM model for the edge.
 - * Collaborated with SW engineers to train, quantize, and sparsify this network down to fixed-point 8-bit and 90% sparsity, and supported HW engineers with its RTL simulation.
 - * Wrote custom fused kernels with Triton to eliminate memory bottlenecks for training the model on Nvidia Ampere GPUs, yielding a training speed up of $\sim 20\%$.
 - * Coordinated multi-gpu training runs on the cloud using AWS EC2 with 8 A100 instances.
 - * Achieved near-SOTA results on audio denoising, vital signs prediction, speech recognition, and Wikitext-103.

Graduate Student Researcher

UCSD

Sept 2017 – Present

San Diego, CA

- Developed continuous dynamical approach to constrained optimization and simulation of spin glasses. Applied the approach for solving boolean satisfiability problems, encompassing problems such as prime factorization (a basis for RSA encryption).
- Developed a new pre-training method for RBMs based on modal sampling, significantly improving the stability and efficiency of the training process.
- Studied the possibility of applying machine learning techniques to quantum many-body systems. Drew a connection between neural networks and quantum computing benchmarks.
- Explored a mathematical approach for describing general computing architectures.

Research Support Associate

MIT

Sept 2016 – Sept 2017

Boston, MA

- Worked on a modular component for high precision magnetic field control for [Dysprosium](#) MOT chambers (under supervision of Nobel laureate [Wolfgang Ketterle](#)).

Undergraduate Research Assistant

UCLA

December 2015 – Sept 2016

Los Angeles, CA

- Analyzed the rigor of the method of data [unfolding](#) in high energy experiments in a Bayesian context (under supervision of Robert Cousins).
- Designed and simulated a voltage array for collimating ion beams (under the supervision of Eric Hudson).

RESEARCH INTERESTS

spin glass, artificial intelligence, memory, quantum computing

SKILLS

Theory: unsupervised learning, stochastic/dynamical systems, quantum/neuromorphic computing

Edge Computing: event-based algorithms, network quantization, online spatiotemporal inference

Libraries: PyTorch, TF, JAX, Hydra, wandb, scikit-learn, Numba, pandas, OpenCV

Engineering: signal processing, audio denoising, SLAM, circuit design, laser optics

CONFERENCE

International Conference on Learning Representations, 2025

IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, 2024

APS Physics March Meeting, Nashville, 2021

APS Physics March Meeting, Denver (Virtual), 2020

Harvard-MIT CUA Winter Retreat, 2017

REFEREED JOURNALS/CONFERENCES

IEEE Transactions on Neural Networks and Learning Systems

IEEE International Conference on Acoustics, Speech and Signal Processing

International Conference on Learning Representations

International Journal of Circuit Theory and Applications

International Joint Conference on Neural Networks

PLOS ONE

GRADUATE TA EXPERIENCE

Spring 2019: UCSD Physcs 212C - Quantum Mechanics III

Winter 2019: UCSD Physcs 200B - Theoretical Mechanics II

Fall 2018: UCSD Physcs 243 - Stochastic Methods