

Vy Ai Vo, Ph.D.

<https://vyaivo.github.io> | vyaivo@gmail.com | Portland, OR 97005 | 760-689-2486

RESEARCH AND WORK EXPERIENCE

AI/ML Research Scientist 05/2019 - 11/2024

Brain-Inspired Computing Lab at Intel Labs, Intel Corporation, Hillsboro, OR

- Developed research proof-of-concepts and software to transfer technology to business units
- Published in top-tier conferences (e.g. ICLR) on natural language processing (NLP), graph neural networks (GNNs), language models (LMs, LLMs) for code generation
- Worked on customer-facing, open-source SOTA performance library for vector similarity search. Benchmarked use in retrieval augmented generation (RAG) on high-performance CPU/GPU systems
- Granted 5 patents for technologies that use machine learning or AI
- Organized 3 cognition and ML workshops at key conferences (e.g. MemARI NeurIPS 2022)
- Division Recognition Award for risk-taking and fast, excellent execution of research showcase demo

Ph.D. research 09/2013 – 03/2019

Neurosciences Graduate Program, University of California, San Diego, La Jolla, CA

- Applied ML and signal processing techniques to diverse data types (fMRI, EEG, human behavior) to investigate how human brains process visual information under cognitive demand (6 publications)
- Optimized analysis on large, noisy datasets (e.g. 400-500% increase in efficiency of model fitting)
- Taught graduate-level advanced data analysis methods and statistics
- Structural MRI segmentation, multi-band fMRI acquisition, eye-tracking, psychophysics

Lab manager & research assistant 07/2011 – 07/2013

Brain & Cognitive Sciences Program, University of Rochester, Rochester, NY

- Investigated how children learn number concepts using fMRI and behavioral tests (1 publication)
- Interviewed, mentored, and managed research assistants and lab volunteers

EDUCATION

University of California, San Diego. Ph.D., M.S. in Neurosciences (Computational). 2013 - 2019

Swarthmore College. B.A., High Honors, double major in Cognitive Science, Biology. 2007 - 2011

SKILLS

Programming: in order of fluency: Python, R, MATLAB, C, Javascript, LabVIEW.

Machine learning: Transformers, recurrent neural networks, dimensionality reduction, supervised and unsupervised learning, generalized linear models, model regularization, big data, preprocessing. Deep neural network optimization, pre-training, fine-tuning, and inference on heterogeneous HW.

Frameworks/Tools: PyTorch, Docker, Tensorboard, Kubernetes, pandas, scikit-learn, git/GitHub/GitLab.

HuggingFace, Langchain, OpenAI API. HPC architectures, some cloud computing, some SQL.

Statistics: Null-hypothesis testing, Monte Carlo methods, some Bayesian models. R, SPSS, JMP.

SELECTED PUBLICATIONS

Kadosh, T., Hasabnis, N., **Vo, V.A.**, Schneider, N., Krien, N., Capotă., M., Wasay, A., Tamir, G., Willke, T., Ahmed, N., Pinter, Y., Mattson, T., Oren, G. (2024). MonoCoder: Domain-Specific Code Language Model for HPC Codes and Tasks. Outstanding paper award at *IEEE High Performance Extreme Computing*.

Tang, J., Du, M., **Vo, V.A.**, Lal, V., Huth, A.G. (2023). Brain encoding models based on multimodal transformers can transfer across language and vision. *Neural Information Processing Systems (NeurIPS)*.

Jain, S., **Vo, V.A.**, Wehbe, L., Huth, A.G. (2023). Computational language modeling and the promise of in silico experimentation. *Neurobiology of Language*.

Mahto, S., **Vo, V.A.**, Turek, J.S., Huth, A.G. (2021). Multi-timescale representation learning in LSTM language models. *International Conference on Learning Representations (ICLR)*.

Jain, S., **Vo, V.**, Mahto, S., LeBel, A., Turek, J., Huth, A. (2020). Interpretable multi-timescale models for predicting fMRI responses to continuous natural speech. *Neural Information Processing Systems (NeurIPS)*.

Vo, V.A., Sprague, T.C., and Serences, J.T. (2017). Spatial tuning shifts increase the discriminability and fidelity of population codes in visual cortex. *Journal of Neuroscience*.