Alexandre Araujo

+33 6 74752275 • aaraujo001@gmail.com

Summary & Research Interests

Accomplished machine learning researcher with over 5 years of experience developing, training, and evaluating large-scale neural network models. Skilled in distributed training techniques enabling state-of-the-art computer vision and natural language processing systems. Author of 10 papers at top-tier machine learning conferences, including NeurIPS, ICML, and ICLR. I am broadly interested in pushing the boundaries of neural networks while ensuring scalability, robustness and security aspects. My current focus is on the design of robust and efficient learned representation and on the robustness and security of foundation models.

Education

Ph.D. in Computer Science, PSL Research University, Paris, France 2017 – 2021
MS in Economics, SKEMA Business School, Lille, France 2013 – 2016
BS in Mathematics, University of Versailles, Versailles, France 2008 – 2011

Research Experience

Postdoctoral Researcher on Trustworthy Machine Learning 2023 – Present
New York University, New York, NY, US
- Advisors: Siddharth Garg, Farshad Khorrami
- Established theoretical connections between mathematical concept and areas of Trustworthy ML. Trained large-scale neural networks using PyTorch in a distributed fashion on a Slurm cluster.

Postdoctoral Researcher on Computer Vision 2021 – 2022
INRIA / École Normale Supérieure, Paris, France
- Advisors: Jean Ponce, Julien Mairal
- Research on Focus Stacking from Handheld Raw Image Bursts. Designed a large-scale computer vision dataset to improve recent advancements on Focus Stacking with supervised learning.

Ph.D. Candidate 2017 – 2021
PSL Research University, Paris, France
- Thesis: Building Compact and Robust Deep Neural Networks with Toeplitz Matrices
- Advisors: Jamal Atif, Yann Chevaleyre and Benjamin Negrevergne
- PhD in Deep Learning with a focus on compact and robust neural network with structured matrices.

Publications

Conference Papers

LipSim: A Provably Robust Perceptual Similarity Metric
S. Ghazanfari, A. Araujo, P. Krishnamurthy, F. Khorrami, S. Garg – ICLR (2024)

The Lipschitz-Variance-Margin Tradeoff for Enhanced Randomized Smoothing
B. Delattre, A. Araujo, Q. Barthélemy, A. Allauzen – ICLR (2024)

Novel Quadratic Constraints for Extending LipSDP beyond Slope-Restricted Activations
P. Pauli, A. Havens, A. Araujo, S. Garg, F. Khorrami, F. Allgöwer, B. Hu – ICLR (2024)

On the Scalability and Memory Efficiency of Semidefinite Programs for Lipschitz Constant Estimation of Neural Networks
Z. Wang, A. Havens, A. Araujo, Y. Zheng, B. Hu, Y. Chen, S. Jha – ICLR (2024)

Exploiting Connections between Lipschitz Structures for Certifiably Robust DEQ models
A. Havens, A. Araujo*, S. Garg, F. Khorrami, B. Hu* – NeurIPS (2023)

Diffusion-Based Adversarial Sample Generation for Improved Stealthiness and Controllability
H. Xue, A. Araujo, B. Hu, Y. Chen – NeurIPS (2023)

Certification of Deep Learning Models for Medical Image Segmentation
O. Laousy. A. Araujo, G. Chassagnon, M. Revel, M. Vakalopoulou – MICCAI (2023)

Towards Better Certified Segmentation via Diffusion Models
O. Laousy. A. Araujo, G. Chassagnon, M. Revel, S. Garg, F. Khorrami, M. Vakalopoulou – UAI (2023)

Efficient Bound of Lipschitz Constant for Convolutional Layers by Gram Iteration
B. Delattre, Q. Barthélemy, A. Araujo, A. Allauzen – ICML (2023)

A Unified Algebraic Perspective on Lipschitz Neural Networks
A. Araujo, A. Havens*, B. Delattre, A. Allauzen, B. Hu* – ICLR – Spotlight (2023)

A Dynamical System Perspective for Lipschitz Neural Networks
L. Meunier, B. Delattre*, A. Araujo*, A. Allauzen* – ICML – Oral (2022)

On Lipschitz Regularization of Convolutional Layers using Toeplitz Matrix Theory
A. Araujo, B. Negrevergne, Y. Chevaleyre, J. Atif – AAAI (2020)

Understanding and Training Deep Diagonal Circulant Neural Networks
A. Araujo, B. Negrevergne, Y. Chevaleyre, J. Atif – ECAI 2020 (2020)

Theoretical Evidence for Adversarial Robustness through Randomization
R. Pinot, L. Meunier, A. Araujo, H. Kashima, F. Yger, C. Gouy-Pailler, J. Atif – NeurIPS (2019)

Workshop Papers

R-LPIPS: An Adversarially Robust Perceptual Similarity Metric
S. Ghazanfari, S. Garg, P. Krishnamurthy, F. Khorrami, A. Araujo – ICML – Workshop (2023)

Advocating for Multiple Defense Strategies against Adversarial Examples
A. Araujo, L. Meunier, R. Pinot, and B. Negrevergne – ECML – Workshop (2020)

Compact Deep Learning Models for Video Classification using Circulant Matrices
A. Araujo, B. Negrevergne, Y. Chevaleyre, J. Atif – ECCV – Workshops (2018)

Preprints

PAL: Proxy-Guided Black-Box Attack on Large Language Models
C. Sitawarin, N. Mu, D. Wagner, A. Araujo – Preprint (2024)

Fine-grained Local Sensitivity Analysis of Standard Dot-Product Self-Attention
A. Havens, A. Araujo, H. Zhang, B. Hu – Preprint (2024)

Towards Real-World Focus Stacking with Deep Learning
A. Araujo, J. Ponce, J. Mairal – Preprint (2023)

Industry Experience

Data Scientist 2015 – 2017
Wavestone, Paris, France
- Collected 5 years of historical data for a mortgage broker and applied machine learning algorithms to predict mortgage application acceptance. Deployed the model into production.
- Gathered 3 years of historic data for an energy company with Hadoop to construct a 1 billion rows dataset. Applied machine learning algorithms to predict churn.
- Gathered 20 years of historic data for a European Railway Company and applied machine learning algorithms to predict train breakdown.

Data Engineer Intern dec. 2014 – may 2015
Amazon, Luxembourg
- Automated data pipelines to feed dashboards that showcase transportation and financial statistics.

Activities and Services

Teaching

New York University, New York, NY, US

Graduate Course: Adversarial Machine Learning	2023

PSL Research University, Paris, France

Executive Master: Adversarial Machine Learning	2020, 2021
Master IASD: Data Mining & Machine Learning	2019
Master ID: Data Mining & Machine Learning	2019

École Polytechnique, Paris, France

Data Science & Machine Learning	2016, 2017, 2018, 2019, 2020

Reviewer

Artificial Intelligence and Statistics (AISTATS)	2022, 2023
Association for the Advancement of Artificial Intelligence (AAAI)	2022, 2023
Computer Vision and Pattern Recognition Conference (CVPR)	2023
European Conference on Computer Vision (ECCV)	2024
International Conference on Computer Vision (ICCV)	2023
International Conference on Learning Representations (ICLR)	2023
International Conference on Machine Learning (ICML)	2023, 2024
Neural Information Processing Systems (NeurIPS)	2023

Supervised Students

Sara Ghazanfari: Ph.D. student, 2023 - Present
Blaise Delattre: Master student, Summer 2021 (Now Ph.D. student)
Alexandre Verine: Master student, Summer 2019 (Now Ph.D. student)

Invited Talks

University of Illinois Urbana-Champaign	October 2023
NYU – Center for Data Science	April 2022
INRIA / École Normale Supérieure de Paris	July 2021
École Normale Supérieure de Lyon	July 2021
INSIS – French National Center for Scientific Research	January 2021
PFIA – French AI conference	June 2019, 2020, 2021
International Cybersecurity Forum	January 2020
Limits of AI – BPI Conference	June 2019

Technical Skills

Programming Languages : Python, C++, SQL
HPC Job Schedulers : Slurm, IBM Spectrum LSF
Deep Learning Frameworks : TensorFlow, PyTorch
ML Libraries: XGBoost, LightGBM, Scikit-Learn
Data Science Framework : OpenCV, SciPy, NumPy, Pandas