Victor Quétu

Third-Year PhD Student in Computer Science

Palaiseau 91120 victorquetu@gmail.com Victor Quétu victor-quétu VGCQ

Profile

Third-Year PhD Student specializing in Deep Learning applied to Computer Vision. Under the supervision of Gaël Richard and Enzo Tartaglione, my research delves into the development of efficient inference approaches for deep learning models, with applications aimed at improving computational efficiency and real-world deployment. I am actively seeking research internship opportunities for the year 2025.

Education

Télécom Paris, Institut Polytechnique de Paris Paris, France
PhD Degree in Computer Science, Advisors: Gaël Richard, Enzo Tartaglione 2022 – Current

- Subject : Constraining Deep Learning with Regularization.
- Current research interests: deep neural network compression, deep learning privacy, and generative models.

ENSEA - Engineering School Paris, France
Master's Degree in Signal Processing & Artificial Intelligence 2019 – 2022

- Major in Mathematics, Electronics, Signal Processing and Computer Science.

Experience

PhD Student Paris, France
LTCI, Télécom Paris, Institut Polytechnique de Paris Oct 2022 – Sep 2025

- Studied the sparse double descent phenomenon and proposed a learning framework to avoid it and improve performance.
- Developed deep learning algorithms using Knowledge Distillation and Pruning to optimize real-time performance on energy-constrained devices.
- Worked on removing private information from the data processed by deep learning models.
- Supervision of a master's student working on layer collapse to reduce on-device inference time.

Research Intern Gennevilliers, France
Thales Feb 2022 – Aug 2022

- Worked on communication systems operating on wide-band channels.
- Implemented Turbo Decision Feedback Equalization with Expectation Propagation using a Matlab simulator of the PHY layer to mitigate inter-symbol interference.
- Developed deep learning algorithms using Pruning to improve the performance-complexity trade-off.

Research Intern Chicago, USA
ECASP Research Laboratory - Illinois Institute of Technology May 2021 – Aug 2021

- Built an Object Recognition and Map Building Robot, controlled with ROS through an Android app.
- Implemented the YOLOv2 model and a SLAM algorithm on an embedded device (Jetson Nano).

Teaching

Teaching Assistant at ENSAE Fall 2024
Practical labs on numpy, pandas, geopandas, API, web scraping, neural networks.

Teaching Assistant at ENSAE — Fall 2024
Practical labs on paradigms and algorithms such as graphs, numerical optimization methods, sorting methods and dynamic programming.

Teaching Assistant at Télécom Paris — Spring 2023
Practical labs on SVMs and kernels methods, neural networks and deep learning.

Teaching Assistant at Télécom Paris — Fall 2022
Practical labs on transfer learning and recurrent neural networks.

Skills

Programming: Python (PyTorch, TensorFlow), Matlab, Cuda, C

Language: French (Native), English (Fluent), German (Beginner)

Mentorship & Service

Reviewer: IEEE TNNLS, GreenFOMO (ECCV 2024 workshop), SCEFA (ECML PKDD 2023 workshop)

Organizer: ELLIS Doctoral Symposium 2024

Research project supervisor: André Pereira e Ferreira (Layer collapse to reduce on-device inference time)

Publications

[1] **V. Quétu**, Z. Liao, N. Hezbri, F. Pizzati, E. Tartaglione, LaCoOT: Layer Collapse through Optimal Transport, *Submitted to CVPR 2025*.

[2] H. Wang, Z. Yu, G. Spadaro, C. Ju, **V. Quétu**, E. Tartaglione, FOLDER: Accelerating Multi-Modal Large Language Models with Enhanced Performance, *Submitted to CVPR 2025*.

[3] Z. Liao, N. Hezbri, **V. Quétu**, VT. Nguyen, E. Tartaglione, Till the Layers Collapse: Compressing a Deep Neural Network Through the Lenses of Batch Normalization Layers, in *AAAI 2025*.

[4] G. Pilo, N. Hezbri, A. Pereira e Ferreira, **V. Quétu**, E. Tartaglione, Layerfold: A Python Library to Reduce the Depth of Neural Networks, in *SoftwareX*. Article - Github Repo

[5] **V. Quétu**, E. Tartaglione, DSD^2: Can We Dodge Sparse Double Descent and Compress the Neural Network Worry-Free?, in *AAAI 2024*. Article - Github Repo

[6] **V. Quétu**, Z. Liao, E. Tartaglione, The Simpler The Better: An Entropy-Based Importance Metric To Reduce Neural Networks' Depth, in *ECML PKDD 2024*. Article - Github Repo

[7] M. Girard, **V. Quétu**, S. Tardieu, VT. Nguyen, E. Tartaglione, Memory-Optimized Once-For-All Network, in *CADL workshop at ECCV 2024*. Article

[8] **V. Quétu**, M. Milovanović, E. Tartaglione, Sparse Double Descent in Vision Transformers: real or phantom threat?, in *ICIAP 2023*. Article - Github Repo

Received the **Caianiello ICIAP Paper Award** for young researchers.

[9] **V. Quétu**, E. Tartaglione, Dodging the Double Descent in Deep Neural Networks, in *ICIP 2023*. Article

[10] **V. Quétu**, M. Milovanović, The Quest of Finding the Antidote to Sparse Double Descent, in *SCEFA workshop at ECML PKDD 2023*. Article

[11] Z. Liao, **V. Quétu**, VT. Nguyen, E. Tartaglione, Can Unstructured Pruning Reduce the Depth in Deep Neural Networks?, in *RCV workshop at ICCV 2023*. Article

[12] E. Tartaglione, F. Gennari, **V. Quétu**, M. Grangetto, Disentangling private classes through regularization, in *Neurocomputing, 2023*. Article