Théo Gigant

Multimodal Deep Learning PhD Candidate

✉ *theo.gigant@l2s.centralesupelec.fr*
gigganttheo.github.io
HuggingFace: gigant

Education

2022–2025 **Université Paris-Saclay / Centrale Supélec - L2S**,
Signal and Image Processing, PhD.
Designing and evaluating methods for automatic summarization of multimodal presentations based on LLMs and VLMs. Under the supervision of Frederic Dufaux and Camille Guinaudeau. Expected graduation Q3 2025.

2018–2021 **Institut Polytechnique de Paris - Télécom SudParis**,
Statistical Modeling, Master of Science.
Applied Mathematics, Machine Learning, Deep Learning, Computer Vision

Professional Experience

Jul 2022–Jul 2025 **PhD Researcher**, Centrale Supélec, Paris.
Conducting research on summarization of multimodal presentations using multimodal deep learning methods.

Jul 2022–Dec 2024 **PhD Researcher**, JustAI, Paris.
Building machine learning pipelines, for applications such as presentation summarization, keyframes extraction, image segmentation, tracking & recommender systems. Teaching deep learning.

Jan 2022–Jun 2022 **Machine Learning Engineer**, Freelance, Paris.
Developing image segmentation & detection models, teaching deep learning, contributing to open source projects.

Jul 2021–Dec 2021 **Computer Vision Research Intern**, Huawei R&D Center, Helsinki.
Developing multimodal models for long video understanding, video retrieval & image generation.

Feb 2021–Jun 2021 **Machine Learning Engineer Intern**, JustAI, Paris.
Proofs of concept with language models, developing data processing and data visualization pipelines.

Jul 2019–Aug 2019 **Machine Learning Research Intern**, Télécom SudParis, Evry.
Developing image classification datasets & models.

Projects

2025 **Multimodal Summarization Toolkit [Repo]**, Slide extraction, transcription, visualization and semantic search for slide-based multimodal presentations..

2022-Ongoing **Deep Learning blog posts [Blog]**, Insights, reflections and opinions on VLMs and LLMs.

Apr 2022–May 2022 **BigBio BioMedical Dataset project [Org]**, *BigScience*, paper published at NeurIPS 2022.

Apr 2022–May 2022 **WikiArt Diffusion Mini [Project]**, *HuggingFace & Paperspace*, HugGAN Challenge.

2022 **ASR Models for Romanian: Wav2Vec2 and Whisper**, *HuggingFace*, Speech Recognition Challenges (Winner).

July 2021 **T5-VAE [Project]**, *Google & HuggingFace*, JAX/Flax Community Week.

First-Author Publications

Apr 2025 **Summarization of Multimodal Presentations with Vision-Language Models: Study of the Effect of Modalities and Structure**, preprint, submitted at ACM TOMM journal.
Fine-grained evaluation of VLMs to devise optimal input settings for summarization of multimodal presentations.

Nov 2024 **Mitigating the Impact of Reference Quality on Evaluation of Summarization Systems with Reference-Free Metrics**, *EMNLP*.
Introducing a new metric for reference-free evaluation of the relevance of summaries. Evaluating the robustness of metrics with respect to the quality of references to develop a robust hybrid evaluation strategy.

Sep 2023 **TIB: A Dataset for Abstractive Summarization of Long Multimodal Videoconference Records**, *CBMI* .
Introducing a dataset of 9k multimodal presentations with a benchmark of fine-tuned state-of-the-art long document summarization models.

Skills and Research Interests

Languages Python (JAX/Flax, PyTorch), Bash, SQL, Java

Research Interests Vision Language Models, Large Language Models, Evaluation, Multimodality, Document Understanding

Spoken Languages French (native), English (Full professional proficiency : 990/990 TOEIC), Spanish (Notions)