

Théo Gigant

Multimodal Deep Learning PhD candidate

✉ theo.gigant@l2s.centralesupelec.fr

📄 [giganttheo.github.io](https://github.com/giganttheo)

HuggingFace: [gigant](#)

Education

- 2022–2025 **Université Paris-Saclay / Centrale Supélec - L2S**,
Signal Processing, PhD.
Representation of encoded features for extractive and abstractive summarization of multimodal content. 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, Deep Learning

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 various applications such as Presentation Summarization, Keyframes Extraction, Image Segmentation, Tracking & Recommender Systems.
- 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.
Proof of Concepts 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]**.
- 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 [Demo]**, *Google & HuggingFace*, JAX/Flax Community Week.

First-Author Publications

- Apr 2025 **Study of the Individual and Combined Effects of Modalities and Structure in Summarization of Multimodal Presentations with Vision-Language Models**, preprint, submitted at ACM TOMM journal, 2025.
- Nov 2024 **Mitigating the Impact of Reference Quality on Evaluation of Summarization Systems with Reference-Free Metrics**, *EMNLP*, 2024.
- Sep 2023 **TIB: A Dataset for Abstractive Summarization of Long Multimodal Videoconference Records**, *CBMI*, 2023.

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)