

Minh Dinh Trong

(+1) 802-296-1616 | minh.t.dinh.gr@dartmouth.edu | [Personal Website](#)

About me (TLDR): I am an MSCS student at Dartmouth with research experience in multimodal encoding, generative models, and privacy-preserving AI. My recent work focuses on building robust, trustworthy systems across vision and language, leveraging diffusion models and self-supervised methods.

EDUCATION

Dartmouth College

New Hampshire, United States

Masters in Computer Science — Tuition Scholarship. GPA: High Pass (A+/ 4.0).

Expected December 2025

- Relevant coursework: Robotics, Machine Learning, Deep Learning, Computer Vision

Aalto University

Espoo, Finland

Bachelor of Data Science with Honours — Tuition Scholarship, Dean's List. GPA: 4.95/ 5.0 August 2021 - July 2024

- Related coursework: Reinforcement Learning, Object-Oriented Programming (C++), Computational Photography

EPFL - Swiss Federal Institute of Technology Lausanne

Lausanne, Switzerland

Exchange Student @ School of Computer and Communication Sciences

September 2023 - July 2024

- Semester project in the IVRL lab
- Graduate coursework: Computer Graphics, Modern NLP, Applied Data Analysis

RESEARCH EXPERIENCE

Research Student

September 2024 – Present

Dartmouth Computer Science Dept. — PI: SouYoung Jin

Hanover, NH, United States

- Developed a privacy-preserving dataset generation pipeline using VLM models and efficient diffusion-based editing. Under review for CVPR 2026.
- Designed ViT-MAE-based autoencoders with structured masking for anonymization while preserving classification performance; optimized for low-latency inference.
- Developed a compact instruction-tuned multimodal model for alignment across vision, audio, and video; evaluated on 13 visual-language benchmarks.

Semester Student

July 2023 – January 2024

EPFL - School of Computer and Communication Sciences - IVLR Lab

Lausanne, Switzerland

- Investigated diffusion-based deepfake detection as part of the ELSA Deepfake Detection Challenge; reproduced and extended the DIRE method using residuals from reverse diffusion steps.
- Built training pipelines on EPFL's GPU cluster and conducted a literature review on GAN/diffusion forgeries and detection vulnerabilities under compression, lighting shifts, and domain variation.

Research Assistant

June 2022 – August 2023

Aalto University – PI: Stéphane Deny

Espoo, Finland

- Research domains: **Self-supervised, Learning Invariant representation, Signal Augmentation**
- Developed self-supervised learning models with robustness to noise, scale, and affine transformations using learnable encoders and reversible operators.
- Designed pipelines for generalization across out-of-distribution domains, including adaptation to high-density EMG signals with real-time classification and interactive visualization. Paper investigating the impact of out-of-distribution physical shifts on EMG-based prediction accepted at EMBC 2025.
- Conducted ablation studies on masking strategies, latent smoothness, and contrastive loss variants for representation alignment under distribution shift.
- Granted **funding** from Helsinki Institute for Information Technology (HIIT).

Bachelor Thesis

January 2023 – April 2023

Aalto University - Advisor: Alexander Ilin

Espoo, Finland

- Topic: **Recent advances in scene representation.** [Thesis](#)
- Independent research on implicit 3D scene representation.
- Studied and reviewed the advancements of frameworks NeRF and SRT.

WORK EXPERIENCE

Machine Learning Engineer

Pokedata.io

December 2025 – Present

New York City, United States

- Productionized the visual search pipeline by packaging detection, embedding, OCR, and retrieval modules into deployable mobile components.
- Optimized models for on-device inference (quantization, pruning, caching), reducing latency from seconds to sub-200ms on mid-range devices.
- Integrated the pipeline into iOS/Android apps and improved real-world robustness (lighting, glare, motion) through capture-side enhancements.

AI Engineering Intern

Pokedata.io

June 2025 – September 2025

New York City, United States

- Led the end-to-end development of a fine-grained visual search system, integrating DINOv2 for retrieval, Grounding DINO for detection, and OCR for metadata alignment.
- Designed and executed a progressive fine-tuning strategy (category → product → variant), achieving 94% final retrieval accuracy on a database of 50K items and reducing dominant error categories by >20%.
- Built scalable training and data pipelines in PyTorch with reference caching and variant-aware hard-negative sampling, and deployed a Gradio demo showcasing real-time detection and retrieval.

TEACHING EXPERIENCE

Teaching Assistant

Dartmouth Computer Science Dept. & Thayer

September 2024 – Present

Hanover, NH, United States

- Courses: Deep Learning, Human-centered LLMs, Artificial Intelligence, Machine Learning and Statistical Data Analysis, Applied Machine Learning, Foundations of Applied Computer Science.
- Hosted weekly online and in-person office hour sessions. Prepared lecture materials and led tutorials explaining complex AI/ML concepts to students across various backgrounds.
- Coordinated assignment grading for over 30 students every course.

Teaching Assistant

Aalto University

September 2022 – June 2023

Espoo, Finland

- Courses: Artificial Intelligence, Machine Learning, Databases, Object Oriented Programming with C++
- Relevant topics and tools PostgreSQL, Numpy, Pandas, scikit-learn, Cmake, Github/ Gitlab, DBeaver,

PROJECTS

Autonomous Indoor 3D Mapping | *ROS2, RTAB-Map, Gaussian Splatting*

April-June 2025

- * Integrated RTAB-Map and 3D Gaussian Splatting into a frontier-based exploration pipeline; debugged multi-sensor fusion, pose alignment, and reconstruction failures to stabilize mapping performance in real environments.
- * Ran extensive Gazebo and ROSbot 2 experiments, diagnosing odometry drift, RGB-D noise, and motion-blur artifacts, and validating the system across simulated and physical indoor spaces.

Uncertainty-Aware Zero-Shot Classification | *Vision-Language, BLIP-2, LLM Q&A*

Jan-Mar 2025

- * Built an uncertainty-sensitive image classifier that defers ambiguous predictions to a secondary LLM-based reasoning module, improving decision robustness on challenging samples.
- * Integrated LLM-generated clarifications into the visual embedding space, yielding a measurable performance gain with a 3% accuracy improvement on misclassified images.

TinyMCQA: Answering MCQ with Finite Resources | *SFT, RAG, DPO, Quantization*

Mar -June 2024

- * Developed a lightweight QA system with TinyLlama fine-tuned using DPO and RAG; integrated semantic passage retrieval to improve relevance and answer grounding across long-form content. Improved answer reliability under tight resource constraints, achieving 24.6% accuracy on MMLU with 2× model compression.

CLIP-Guided Diffusion for Lightweight Visual Generation | *Diffusion Models, CLIP*

Mar -June 2024

- * Designed a text-to-image pipeline that synthesizes compact 2D Gaussian splats from natural language prompts; trained with CLIP loss for semantic alignment. Benchmarked against LIVE/DiffVG, showcasing improved compression-performance tradeoffs for on-device use.

- Rendering Interactive Random Island with WFC** | *Computer Graphics, GLSL, JavaScript* Mar -June 2024
- * Utilized Wave Function Collapsed (WFC) for generating 2D maps with varied terrain attributes.
 - * The maps are interpolated with 3D noise function, resulted in 3D interactive landscapes.
- 2D Gaussian Splatting for Image Representation and Compression** | *Computer Vision* Mar -June 2024
- * Compared performance between 2D Gaussian Splatting and LIVE/DiffVG in compressing images.
 - * Proposed Text-to-2DGS model, which generates 2D Gaussian points for efficiently rendering pictures that match textual description following CLIP loss.
- Distributed Information Systems project series** | *Pytorch, NLP, Pandas* Sept - December 2023
- * Developed effective methods for Information Retrieval, Recommender Systems, and Named Entity Recognition by employing deep learning techniques and efficient pipelines.
 - * Built a TF-IDF + MiniLM search pipeline with 75% recall on 1.4M documents in 8 minutes. Developed an attention-based recommender (MSE 0.861) and a graph-based entity linker simulating session-aware search.
 - * Consistently achieved a high ranking on Kaggle leaderboard by almost-perfect score on MSE and F-1 metrics.
- Bayesian Modelling of Housing Prices in Finland** | *Bayesian Data Analysis, R, BRMS* January 2023
- * Different approaches of Bayesian models, both linear and non-linear, with various prior choices.
 - * Best-performed model gives prediction with lower RMSE than frequentist method.
- Mispronunciation Detection with Wav2Vec2 & Forced Alignment** | *PyTorch, NLP* Nov–Dec 2022
- * A deep leaning framework applicable for foreign language learners in detecting pronunciation mistakes.
- Understanding company risks in Finnish municipalities** | *Time-series prediction, Dash* Sep–Dec 2022
- * A cohesive tool that facilitates client’s discovery of Finland’s municipalities with financial risk and those with investment potentials by using time-series analysis and a proposed risk function
 - * Delivered as a Dash application with interactive 2D and 3D visualisation
- Exploratory Visualization for Sanitary Data** | *D3, JavaScript* Sep–Dec 2022
- * Demo on Personal Website. A HTML-based, interactive data Storytelling utilizing D3 and folium map, comprehensively covering aspects of Usability, Time, and Space.
- Applying Soft Actor-Critic for novel environments** | *Deep Reinforcement Learning* Nov–Dec 2022
- * Implementing Soft Actor-Critic, from scratch, from original to refined version, select suitable hyperparameters for various environment.

SERVICE, COMMUNITY, & LEADERSHIP

Reviewer: International Journal of Computer Vision

Volunteer

June - July 2022

IEEE International Symposium on Information Theory (ISIT)

Espoo, Finland

- Active volunteer in all Plenary Sessions and in 5+ tutorial sessions.
- Prepared stage and assisted speakers for their presentation.

Head of Public Relations

October 2019 – August 2020

CSP Dance Team

Hanoi, Vietnam

- Manage social media platforms (Facebook page, Instagram, Youtube); plan & conduct weekly agenda to reach averagely 10.000+ audiences.
- Vice Head of Organization Board of annual "Le Début" event, which attracted 200 attendants.
- Director of Photography for the crew’s choreography videos (reached 100.000+ YouTube audiences).
- Lead the promoting campaign when the crew joined HBDC Competition + post live updates during the Finale.

Organizing Committee

August 2018 – August 2020

PTC Media

Hanoi, Vietnam

- Organizing Committee member of Fiesta A Cielo - the greatest chain of event in High School For Gifted Students, which attracts 2000+ attendees.

Editor

September 2017 – February 2018

TechPrimeVN

Hanoi, Vietnam

- Conduct research about technology-related topics and write in-depth analysis.
- Publish updates from technological field by translating from creditable sources.
- Experience in managing websites and embed Search Engine Optimization to articles