

Nishka Katoch

[Portfolio](#) | nishkakatoch@gmail.com | [Linkedin](#) | [Github](#)

AI researcher with expertise in deep learning, generative models, and quantum-enhanced algorithms, driving innovations in scalable RAG systems and robust deepfake detection. Skilled in building production-ready pipelines with Docker, Milvus, and modern ML frameworks, delivering measurable improvements in accuracy, efficiency, and resilience. Proven leader in research and cross-functional collaboration, with publications, scholarships, and hands-on experience across computer vision, NLP, ASR, and medical imaging.

TECHNICAL SKILLS

Languages: Python, C/C++, Java, SQL, HTML/CSS, JavaScript

ML/DL Frameworks: PyTorch, TensorFlow, Scikit-learn, Monai, nnUNet, AxonDeepSeg, fastai

Tools: Linux, Git, Jupyter, SSH, ROS, Unity, Docker, Ivadomed, MuJoCo, Django, PostgreSQL, Milvus

Domains: Computer Vision, Generative Models, Natural Language Processing, Multilingual ASR, Medical Imaging, Domain Adaptation, Self-Supervised Learning, RAG, Quantum Computing

Soft Skills: Research Communication, Analytical Thinking, Technical Writing, Team Collaboration, Problem-Solving

EDUCATION

M.Sc., Computer Science (ML & AI Thesis) – University of Montréal (Mila) Sep 2021 – Aug 2025

- **Supervisor:** Prof. Julien Cohen-Adad and Prof. Guy Wolf
- **Thesis:** Multi-Contrast Image-to-Image Translation for Axon and Myelin Segmentation

B.Tech., Computer Science – Banasthali University Jul 2017 – May 2021

EXPERIENCE

AI Research Intern | Team Lead (Intern Cohort) - Tech Mahindra, Hyderabad Aug 2025 – Oct 2025

- Leading two deep learning research projects: benchmarking embedding models & vector quantization for scalable RAG systems, and enhancing GAN-based deepfake detection with quantum algorithms.
- Develop and test pipelines with Docker and Milvus, evaluating trade-offs in accuracy, latency, and memory to optimize enterprise-scale deployment.
- Built a novel Vision Transformer–PQC pipeline with 4- and 8-qubit circuits for quantum-enhanced binary classification of synthetic images.
- Addressed challenges in dimensionality reduction, quantum encoding, and circuit depth to evaluate hybrid systems for real-world deepfake detection.

Graduate Researcher - NeuroPoly Lab, Polytechnique Montreal Jan 2022 - Aug 2025

- Developed deep learning models for axon–myelin segmentation, morphometry, and modality translation, boosting automation in neuron imaging workflows.
- Built MONAI-based medical imaging pipelines, reducing training and inference overhead for multiple projects.
- Contributed to Ivadomed and AxonDeepSeg, improving reproducibility and usability for neuroimaging researchers.
- Conducted research behind Unpaired Modality Translation for Pseudo-Labeling and Multi-Contrast Axon Segmentation via Latent Diffusion, advancing cross-contrast imaging methods.
- Collaborated with medical researchers to refine outputs, enabling more reliable neuropathology analysis.

AI Intern – Custom Quality Build (Remote) Apr 2024 – Mar 2025

- Built a computer vision pipeline using the Segment Anything Model (SAM) and diffusion models, generating novel interior design variations from raw house images.
- Designing a text-to-image control interface to allow interactive, natural language–based customization of design outputs.

ML Research Intern – Maker’s Lab, Tech Mahindra (Remote) Dec 2023 – Jan 2024

- Designed and deployed a Hindi ASR system with >90% accuracy on diverse regional datasets. Fine-tuned Indic language models for multilingual speech-to-text tasks, improving robustness across dialects.
- Conducted a comparative review of 20+ ASR and transformer-based architectures, identifying optimal approaches for multilingual speech-to-text.

Web Application Development Intern – CRIS, Delhi

Jun 2020 – Nov 2020

- Developed 4 modules in a Django web app, digitizing workflows, increasing efficiency by 70%.
- Optimized databases by reducing redundancy 60%, enhancing scalability, and reporting.
- Mentored junior developers on software engineering best practices.

Research Assistant – IIIT Delhi, India

May 2018 – Jul 2018

- Built a ROS–Unity interface for VR-based human control of UAV swarms, achieving <80ms latency for real-time responsiveness.
- Simulated predator–prey swarm dynamics to evaluate multi-agent coordination under VR-driven human guidance.

SELECTED PROJECTS & PUBLICATIONS

Unpaired Modality Translation for Pseudo-Labeling – MICCAI 2024 | NeuroPoly Lab

Domain Adaptation, Diffusion Models, nnUNet, VQ-VAE, Medical Image Segmentation

- Created a SyncDiff-based domain adaptation pipeline, improving segmentation accuracy on out-of-distribution microscopy data by 12%.
- Applied VQ-VAE compression, accelerating training while retaining 95% image fidelity.

Multi-Contrast Axon Segmentation via Latent Diffusion – Thesis | NeuroPoly Lab

Privacy-Preserving Machine Learning, Latent Diffusion Models, Healthcare AI

- Designed privacy-preserving latent diffusion models for histological images of axons & myelin, achieving 80% Dice score using synthetic contrast data.
- Demonstrated domain adaptation without requiring access to real patient data.

SpeechBrain Scaling Study – Academic Project

Self-Supervised Learning, Model Scaling, Speech Recognition, Benchmarking

- Benchmarked wav2vec2, RAVE, and SpeechBrain across dataset sizes, validating theoretical scaling laws.
- Informed efficient model deployment strategies for multilingual ASR in low-resource environments.

Single-Cell RNA Analysis of Alzheimer's Disease – Academic Project

Transcriptomics, Bioinformatics, Gene Expression, Alzheimer's Disease

- Analyzed single-cell transcriptomics to identify genetic and environmental factors.
- Visualized gene expression patterns to explore biomarkers for Alzheimer's.

SCHOLARSHIPS & AWARDS

- Awarded a Study, Development Scholarship for my research with Prof Julien Cohen-Adad at Mila.
- Awarded an International Student Scholarship to pursue my master's at the University of Montreal.
- Awarded "Top Student" in "Data Analysis Using Python" workshop by CETPA Infotech 2019.