Shubham Gajjar

Portland, Maine • gajjar.shu@northeastern.edu • linkedin.com/in/shubhamgajjar • github.com/ShubhamGajjar

Education

Northeastern University, MS in Artificial Intelligence – Portland, Maine — Expected May 2027

- GPA: 4.0/4.0

LDRP Institute of Technology and Research, BE in Computer Engineering – Gandhinagar, India — May 2025

- GPA: 8.41/10.0

Experience

Northeastern University, Roux Institute, Research Assistant, Vision-Language Models in Biomedicine – Portland, Maine — Feb 2026 - Present

- Selected and set up MedGemma 1.5 4B, LLaVA-Med and Gemma 4 31B as base models, then fine-tuned MedGemma with 4-bit QLoRA on an NVIDIA H100, preserving the model's reasoning under a strict output format
- Designed a five-type visual question answering schema that turned expert annotations into supervised training data across ∼2M microscopy tiles
- Built the evaluation that drove checkpoint selection, reaching a case-level F1 of 0.83, and traced over-detection to a training-distribution shift
- Ran a linear-probe sweep to separate what the model could not perceive from what it had miscalibrated, and used it to choose the tiling scheme, worth 8 points of per-cell accuracy over whole-field downscaling
- Implemented Group Relative Policy Optimization from scratch in PyTorch, with a verifiable reward computed against ground-truth cell counts

BigCircle (UPSAAS Technologies LLP), AI Engineer – Gandhinagar, India — Mar 2025 - Jul 2025

- Diagnosed redundant API calls between evaluation agents in an AI research platform and re-architected the workflow, cutting report generation from 20 minutes to 5
- Traced server timeouts in a client-facing web app to a 20,000-row table loading at once, and built paging that cut login time from 3 minutes to 2 seconds

BigCircle (UPSAAS Technologies LLP), Developer Intern – Gandhinagar, India — Jan 2025 - Feb 2025

- Prototyped the tooling that became the research platform, and was promoted to AI Engineer

Projects and Publications

VGG16-MCA UNet for Brain Tumor Segmentation (arXiv preprint, first author) — Sep 2026

- Designed the decoder's Multi-Channel Attention module through a study of ten architecture and loss variants, reaching a pooled pixel-level Dice of 95.10% on BraTS 2020

MorphoCLIP: Image-Text Contrastive Learning for Cell Microscopy (arXiv preprint) — Aug 2026

- Built the text pipeline, CellCLIP baseline, evaluation scripts and multi-GPU training infrastructure; 37-41% image-to-text Recall@10 on held-out CPJUMP1 against ∼10% chance

Hybrid ResNet-ViT for Skin Cancer Classification (IEEE AIC 2025, first author) — Jul 2025

- Paired a frozen ResNet50 with a Vision Transformer block, reaching 96.3% accuracy and macro F1 0.961 on HAM10000

GMSM-USA Community Platform — Jun 2026 - Present

- Solo-built a platform used by 340 members across 45 states, with authorization enforced in-database by 47 row-level security policies

Technical Skills

Languages: Python, JavaScript, TypeScript, SQL

Deep Learning: PyTorch, TensorFlow, HuggingFace Transformers, CUDA, QLoRA, GRPO, DDP

Multimodal: MedGemma, LLaVA-Med, SigLIP, DINOv3, CLIP, OpenCV

MLOps & Web: MLflow, PySpark, PostgreSQL, Docker, HPC/Slurm, React, Next.js, FastAPI