

Om Rastogi

 github.com/omrastogi |  linkedin.com/in/om-rastogi |  omrastogi@gmail.com |  +1-857-376-8169 |  Boston

Applied AI Research | GenAI, Diffusion models, Multimodal systems | 3+ Experience

EDUCATION

Northeastern University

Master of Science in Artificial Intelligence — **GPA 3.7/4.0**

Relevant Coursework: Deep Learning, Robotic Sensing

Boston
Sep '25 - May '27

JSS Academy Of Technical Education

Bachelor of Technology (B.Tech.) in Electronics and Communication

Delhi
Aug '17 - Aug '21

TECHNICAL SKILLS

GenAI	Diffusion Models, Transformers, Multimodal LLMs, LoRA/PEFT
Deep Learning	PyTorch, TensorFlow, HuggingFace, Transformers, Diffusers, Accelerate
Cloud/Compute	Google Cloud, Amazon Web Services (Certified), CUDA, SLURM
Computer Vision	OpenCV, FFmpeg, Feature Matching, Optical Flow, CNN, ViT
Tools	Git, Bash, Docker, Vim, Conda, Jupyter, TensorBoard/W&B, BigQuery(sql)

RESEARCH PROJECTS

MLLM-Based Short-Form Video Clipping from Long-Form Ads

Augmented Cognition Laboratory at Northeastern University

Nov '25 - Present
Boston

- Formulated the task of selecting short-form ad clips from long-form video as a temporal classification problem and designed IoU-based metrics to evaluate predicted vs. ground-truth clip alignment, enabling standardized MLLM benchmarking.
- Built a video processing pipeline for shot segmentation, metadata extraction, and model inference; evaluated QwenVL3, and MiniCPM on a curated dataset of 800 long-short video pairs. Submitted to IJCAI '26.

Finetuning Diffusion Transformers for Depth Estimation

Vision and AI Lab at IISc

Jan '24 - Oct '25
Bangalore

- Adapted [PixArt-Sigma's Diffusion Transformer](#) architecture for monocular depth estimation, modifying the denoising pipeline to predict dense depth maps instead of RGB outputs.
- Trained on the [MatrixCity dataset](#), a synthetic urban dataset spanning street-level to aerial viewpoints; achieved 0.107 absolute relative error and 0.88 δ_1 accuracy, demonstrating generalization across extreme scale variations.

EXPERIENCE

Dashtoon | Research Engineer

Bangalore | Mar '25 - Jul '25

Developed Character LoRA Models and Video Models for Animated Storytelling

- Worked with production and design teams to review 500+ generated samples, finding texture fidelity and identity consistency issues in 150 low-quality outputs and identifying them as key workflow bottlenecks.
 - Integrated [Hidream-I1 \(MoE\)](#) into the generation pipeline, **reducing low-quality outputs by 90%** and improving production reliability.
 - Owned development of character-generation components for animated storytelling, training, and deploying Flux LoRA adapters for 20+ characters, ensuring **identity consistency** across 8 distinct artistic styles.
 - Led **data curation of 10,000+ samples** from 100+ films to build a regionally representative training set tailored to Indian user contexts.
- Keywords:** Lora finetuning, architecture, optimization, Proof of Concept, GCP, Vertex AI, mixture-of-experts

Collablens | Senior Research Engineer

Gurgaon | Dec '22 - Jan '24

Vision pipelines for video analytics and precision visual inspection

- Led collaboration with ITC on a truck-loading monitoring system for freight damage mitigation; engineered rough-handling detection using **YOLO on superimposed frames**, achieving 70% accuracy with <5% false detections.
- Architected and deployed an [automated edge-to-cloud inference pipeline](#) integrating three intelligent modules in near real-time.
- Developed a visual inspection system for Smartivity, achieving **97% accuracy** on missing-part detection and **5x speedup** in inspection time, driving measurable business impact on production quality.

Keywords: Object Detection, Action Recognition, Optical Flow, SIFT, AWS, GCP, Bigquery

EZ Works | Data Scientist

Gurgaon | Jul '21 - Nov '22

Machine Translations using language models, vision solutions for document parsing

- Led development of an English-Arabic NMT system through dataset curation (10k pairs), mBART fine-tuning, and **ONNX**-optimized deployment, achieving a BLEU score of 50+.
- Collaborated with the translation team to evaluate model quality, identified subjective gaps, and augmented training data to improve the internal quality score from 300 to 360/400.
- Presented product recommendations to the implementation team, aligning model capabilities with business use cases; deployed on 10% of translations, saving **\$1M+ annually**.
- Built translation memory retrieval using **BERT** embeddings with custom batch similarity, achieving **10x speedup** and reducing turnaround time for the translation workflow.
- Designed a tool achieving 80% accuracy in detecting tabular structures using object detection and wireframe parsing.

Keywords: Training, BERT, GPT, Model optimization, Semantic Search, Deep Learning