Nitesh Sekhar

niteshsekhar@gmail.com — +1-630-596-6875 — github.com/nitesh1813 — linkedin.com/in/nitesh-sekhar

Applied AI Research Scientist specializing in large-scale multimodal foundation models, vision-language pretraining, and representation learning. Experience developing large-scale models from data curation and distributed training to production deployment.Published at ICLR and CVPR and shipped ML systems across consumer devices, AR platforms, and large-scale cloud services.

Technical Expertise

- **Research Areas:** Multimodal LLMs, Representation Learning, Vision-Language Models, Scaling Laws, Retrieval Embeddings, 3D Perception
  Architectures: Transformers, Vision Transformers (ViT), CLIP-style Models, Contrastive Learning, Autoregressive Models
  Frameworks: PyTorch, Nemo, Distributed Training, CUDA, ONNX, Megatron, vLLM

Research & Industry Experience

Amazon
Applied Scientist 3 – AGI (Amazon Nova) *Mar 2025 – Present*
- Conducted research on multimodal foundation model scaling behavior, designing large-scale pretraining experiments to study compute–data–model tradeoffs.
- Architected and trained a unified multimodal embedding model spanning video, images, documents, audio, and text, including dataset curation, objective design, and distributed training pipeline

Amazon
Applied Scientist 2 – AGI (Amazon Nova) *Aug 2023 – Mar 2025*
- Developed vision-language pretraining framework combining contrastive representation learning with autoregressive language modeling to enable high-resolution multimodal training on long-form documents
- Designed a resolution-agnostic Vision Transformer supporting variable image resolutions and aspect ratios, improving downstream visual understanding accuracy by 23.4%
- Proposed token-merging strategy compressing visual token sequences by 75% before LLM ingestion while maintaining downstream task accuracy and reducing compute
- Introduced parameter-efficient adaptation (LoRA) to add vision capabilities for custom finetuning of Nova models

Amazon
Applied Scientist 2 – Alexa AI *May 2022 – Aug 2023*
- Shipped real-time expressive face avatar system for conversational interaction on Alexa devices
- Developed active speaker detection model for multi-speaker conversations deployed publicly as part of the Alexa "Let's Chat" feature
- Performed demographic performance analysis and guided targeted data collection to mitigate bias and improve robustness

Amazon
Applied Scientist 2 – Physical Stores *Jul 2021 – Apr 2022*
- Built and deployed 3D multi-view object verification system for automated planogram compliance in retail stores within 4 months
- Developed customer action-location recognition using temporal 3D convolutional networks for in-store behavior understanding

Magic Leap, Inc
Senior Computer Vision Engineer *Jun 2019 – Jul 2021*
- Developed real-time 3D object recognition system for indoor augmented reality environments deployed on head-mounted devices
- Designed panoptic segmentation pipeline for scene understanding in spatial computing platforms
- Improved on-device object detection and instance segmentation models (RetinaNet, Mask-RCNN, EfficientDet)
- Built cloud simulation platform to accelerate testing and validation of AR perception systems

Selected Publications

Avoiding Spurious Correlations via Logit Correction ICLR 2023

Multi-user Scalable 3D Object Detection in AR Cloud CVPR Workshop on AR/VR 2020

Octree Representations of 3D Volume Radiology: Artificial Intelligence

Education

University of California, San Diego La Jolla, CA
M.S. Computer Science (GPA: 3.96/4.0) *2017 – 2019*
Indian Institute of Technology Kharagpur India
B.Tech Computer Science and Engineering (GPA: 8.82/10.0) *2013 – 2017*