Abbaas Alif Mohamed Nishar

Applied Scientist | Multimodal AI | Computer Vision | Vision-Language Models | LLMs
470-601-4788 | abbaasalif6998@gmail.com | LinkedIn | GitHub | Google Scholar | Portfolio

Summary

Applied Scientist and Computer Science Ph.D. specializing in multimodal AI, computer vision, vision-language model evaluation, event-based vision, and large language models. Developed real-world screen-camera communication, neuromorphic-camera systems, video-language stress tests, image-segmentation pipelines, and GPU-accelerated visual decoders. Published across computer vision, multimodal learning, optical wireless communication, NLP, and AI systems.

Core Skills

Multimodal and Vision-Language AI: Video-language models, video understanding, temporal grounding, visual question answering, multimodal evaluation, cross-modal robustness, synthetic benchmark generation, LLM-as-a-judge
Computer Vision: PyTorch, OpenCV, image segmentation, object detection, YOLO, GANs/DCGAN, PoseNet, feature engineering, model evaluation, data annotation, DALL-E inpainting
Event-Based and Optical Vision: Neuromorphic/event cameras, optical camera communication, screen-camera systems, visible-light communication, non-line-of-sight sensing, multimodal sensing, rolling-shutter channels
LLMs and Systems: Transformers, Hugging Face, PEFT, LoRA/QLoRA, RAG, constrained decoding, vLLM, FastAPI, React, Kubernetes, Helm, Amazon EKS, Docker, AWS S3, multi-GPU training/inference
Programming and Data: Python, C/C++, CUDA C, SQL, MATLAB, JavaScript/TypeScript, NumPy, Pandas, Scikit-learn, TensorFlow, JAX, Numba, Plotly

Experience

Applied Scientist June 2025 – Present
Capital One – AI Foundations *San Jose, CA*
- Develop LLM-powered understanding, personalization, model evaluation, and agentic workflows for an enterprise conversational assistant.
- Fine-tuned Llama-based models with QLoRA across 300+ intents; integrated retrieval, constrained decoding, and multi-turn context to improve long-tail accuracy by approximately 12 percentage points.
- Built large-scale evaluation and failure-analysis systems for tens of thousands of conversational records, including semantic judging, synthetic-data generation, and ambiguity analysis.
- Built a Kubernetes/Helm model platform for deploying Hugging Face and multimodal-capable models as GPU-backed vLLM services with S3 streaming and automated validation.

Founding Engineer July 2023 – June 2025
Revelio Communications Inc. *Atlanta, GA*
- Built a real-world screen-camera communication system that embeds visually imperceptible metadata in video and decodes it through standard mobile cameras.
- Designed spatially adaptive OKLAB flicker encoding, symbol packing, frame localization, and neural decoding for robust operation across viewing distances, angles, noise, and screen-camera asynchrony.
- Accelerated the GPU encoding pipeline by 100× and developed a multi-stage neural decoder that improved mobile decodability by 60%.

Graduate Research Assistant Aug 2021 – June 2025
Georgia State University *Atlanta, GA*
- Conducted research in event-camera communication, joint sensing and communication, video watermarking, optical tags, LLM-generated network simulation, and environmental sensing.
- Developed non-line-of-sight visible-light communication using asynchronous event streams and adaptive pulse modulation across reflective everyday objects.
- Built data-ingestion, visualization, and ML pipelines for multimodal IoT, radon sensing, cosmic-ray measurements, and network-traffic classification.

Data Science Intern May 2022 – Aug 2022
American Family Insurance *Remote*
- Developed image-segmentation models for home-inspection underwriting; reduced the training codebase by 80% and implemented automated multi-GPU training.

Research Intern Dec 2019 – Jun 2020
Tata Consultancy Services *Chennai, India*
- Developed AR/VR and computer-vision prototypes for clothing-pattern detection, PoseNet-based surveillance, and text-to-avatar generation.

Selected Multimodal and Computer Vision Research

REVEAL: Stress Testing Video-Language Models | *VidLMs, Temporal Grounding, VQA* 2026
- Co-developed a diagnostic benchmark with five controlled stress tests for visual evidence use, temporal expectation bias, language-only shortcuts, video sycophancy, camera-motion sensitivity, and spatiotemporal occlusion.
- Built scalable perturbation and evaluation pipelines showing that leading video-language models often rely on linguistic priors rather than observed visual and temporal evidence.

Revelio Screen-Camera Communication | *PyTorch, OpenCV, CUDA, Video Watermarking* 2025
- Designed a complete video-to-mobile communication pipeline using imperceptible high-frequency color modulation and neural symbol recognition.
- Prioritized visual quality and robust 48-bit metadata recovery for interactive television, content authentication, and second-screen experiences.

Non-Line-of-Sight Communication with Event Cameras | *Neuromorphic Vision, VLC* 2025
- Demonstrated passive optical communication through reflected light using high-temporal-resolution event streams; evaluated modulation, material, color, size, and multipath effects.

Data-Agnostic Image Annotation with Optical Tags | *Vision, Localization, Communication* 2021
- Developed an optical-tag approach for automated image labels and localization, reducing dependence on manual annotation for object detection and segmentation datasets.

hAIrmony and Custom Visual Models | *Segmentation, DALL-E, YOLO, GANs* 2020–2023
- Built a hairstyle visualization system using hair segmentation and generative inpainting; trained a custom YOLO ambulance detector and implemented GAN-based synthetic face generation.

Education

Ph.D. in Computer Science May 2025
Georgia State University *Atlanta, GA*

M.S. in Computer Science May 2022
Georgia State University *Atlanta, GA*

B.Tech. in Electronics and Communication Engineering May 2020
Vellore Institute of Technology *Chennai, India*

Selected Publications and Patent

- **Stress Tests REVEAL Fragile Temporal and Visual Grounding in Video-Language Models** — arXiv 2026; presented at two ICML 2026 workshops.
- **Revelio: A Real-World Screen-Camera Communication System with Visually Imperceptible Data Embedding** — IEEE ICASSP 2025.
- **Non-Line-of-Sight Optical Wireless Communication Using Neuromorphic Cameras** — EWSN 2025.
- **Poster Abstract: Joint Optical Wireless Communication and Sensing Using Neuromorphic Cameras** — ACM/IEEE IPSN 2024; Third Place Best Poster Award.
- **Data Agnostic Image Annotation** — NeurIPS 2021 Data-Centric AI Workshop.
- **Machine-Learning-Based System for Real-Time Screen-to-Camera Communication** — U.S. Patent Application 20250030811.

Awards and Leadership

Best Student Paper Award, IEEE SoutheastCon 2024; Third Place Best Poster Award, ACM/IEEE IPSN 2024; Third Place Oral Presentation, GSU Graduate Research Conference. President, IEEE at Georgia State University (2024–2025); guest lecturer for Advanced Computer Vision and Introduction to Robotics.