Suryodaya B. Shahi

sshahi20@umd.edu • (667) 445–7800 • College Park, MD
LinkedIn • GitHub • Portfolio

EDUCATION

University of Maryland, College Park — *Expected Dec 2025*

M.S. Data Science — Focus: Multimodal Perception & Vision–Language Systems

Harvard University (DCE) — *Aug 2025*

Course: Wearable Devices & Computer Vision — Grade: A (4.0)

Delhi Technological University (DTU) — *May 2023*

B.Tech in Software Engineering — Govt. of India ICCR Scholar (2019–2023)

EXPERIENCE

Perception & Robotics Group, UMD — *Sep 2025 – Present*

Research Assistant — College Park, MD

- Enhanced **VioPose** 4D audiovisual estimator; improved motion accuracy by **15%**.
- Processed **120+ multimodal recordings**, reducing SMPL-X retargeting drift by **30%**.
- Built **LLM-driven acoustic–pose linking module** for expressive violin motion understanding.

Harvard Ophthalmology AI & Robotics Lab — *May – Sep 2025*

AI Research Intern — Boston, MA

- Benchmarked **Qwen-VL, DeepSeek, SEED-LLaMA, LLaVA** on Meta Aria Gen-1; achieved **23% higher grounding accuracy**.
- Co-developed the **VISTA Dataset** (50+ egocentric scenes) introducing **Action–Truth** annotations for assistive AI.

Macquarie University (Remote) — *Jul – Oct 2023*

Research Intern — Sydney, Australia

- Processed **5,000+ Nepali memes** for multilingual sentiment analysis.
- Fine-tuned **BERT, VGG19, ResNet50**, boosting multimodal accuracy by **17%**.

Delhi Technological University (DTU) — *Jan 2022 – Jun 2023*

Undergraduate Research Intern — New Delhi, India

- Co-authored Springer paper *"Deep Learning Methods for Vehicle Trajectory Prediction"*; achieved **14% accuracy gain**.

SELECTED PROJECTS

Aria Glasses on Qwen (Wearable AI) — Built a real-time **VLM pipeline** for Meta Aria Gen-1 using Qwen-VL, BLIP, and OWL-ViT; achieved **35% faster inference** and **60% less manual input**.

Tiny-ACE (Self-Improving Tiny LMs) — Adapted Stanford's ACE framework for sub–3B models (Phi-2, TinyLlama), enabling continual on-device learning via contextual reflection without fine-tuning.

Clinical Notes Entity Extraction (AWS NLP) — Designed a **serverless medical NLP pipeline** (Lambda, S3, DynamoDB) automating entity tagging of diagnoses and tests, cutting review time by **40%**.

PUBLICATIONS & LEADERSHIP

- **In Preparation:** "VISTA: Vision for Impaired Smartglass Tactile Actions" — with Harvard Ophthalmology AI & Robotics Lab, targeting *CVPR / Lancet Digital Health 2026*.
- Shiwakoti, S., **Shahi, S.**, & Singh, P. (2024). *Deep Learning Methods for Vehicle Trajectory Prediction*. In *ICICNIS 2023*. Springer. DOI: 10.1007/978-981-99-6586-1_37
- **Reviewer:** ACM TheWebConf 2025 (MM4SG Workshop).
- **Founder:** Kaushala (EdTech Startup) — 6 partner schools, 3K+ users.
- **Awards:** ICCR Scholar, Gold Medal — IIT Bombay TechFest.

TECHNICAL SKILLS

Programming: Python, C++, R, MATLAB, SQL   **ML/DL:** PyTorch, TensorFlow, Scikit-learn, Transformers, LoRA, RLHF
Vision/AI: OpenCV, SMPL-X, Meta Aria SDK, VLMs, VLAs   **Data/Cloud:** AWS, GCP, MySQL, MongoDB, Neo4j, Docker, W&B   **Tools:** Git, VS Code, PyCharm, Linux, macOS, LaTeX