SURYODAYA B. SHAHI

Graduate Researcher in AI & Computer Vision | Wearable & Assistive Systems
College Park, MD   sshahi20@umd.edu   (667) 445-7800
 SirAlchemist1   Portfolio   LinkedIn

EDUCATION

University of Maryland, College Park   *Expected Dec 2025*
M.S. in Data Science (GPA: 3.83/4.0)   *College Park, MD*

Harvard University, Division of Continuing Education   *Jun 2025 – Aug 2025*
DGMD S-14: Wearable Devices & Computer Vision (Grade: A)   *Cambridge, MA*

Delhi Technological University (formerly DCE)   *May 2023*
B.Tech in Software Engineering (GPA: 8.22/10.0)   *New Delhi, India*
Thesis: Deep-Learning Methods for Vehicle Trajectory Prediction (Springer ICICNIS 2023).

PUBLICATIONS

PUBLISHED

[P1] Shiwakoti, S., **Shahi, S.**, & Singh, P. (2024). *Deep-Learning Methods for Vehicle Trajectory Prediction: A Survey*. ICICNIS 2023, Springer. DOI.

IN PREPARATION

[I1] Yeh, J., **Shahi, S.**, & Wang, M. *VISTA: Action-grounded Egocentric Dataset & On-Device Assistive AI.*

[I2] Perception and Robotics Group UMD. VioPose: *Audio-Conditioned Pose Refinement for Violin Motion.*

RESEARCH EXPERIENCE

Perception & Robotics Group, University of Maryland   *Sep 2025 – Present*
Research Assistant (PI: Dr. Cornelia Fermüller)   *College Park, MD*

- Benchmarked **RoHM, FinePOSE, NLF** on **VioDat** dataset; produced MPJPE/PCK/AUC baselines with full diagnostics.

- Mapped failure modes (bow–hand occlusion, rapid wrist, off-axis drift) and linked to missing temporal/audio cues.

- Developing **audio-conditioned temporal refinement** (bi-GRUs + attention) for expressive violin motion.

Harvard Ophthalmology AI & Robotics Lab, Schepens ERI   *May 2025 – Sep 2025*
AI Research Intern (PI: Dr. Mengyu Wang)   *Boston, MA*
VISTA: Egocentric dataset & on-device assistive AI for low-vision users.

- Co-built **VISTA**, a multimodal (RGB, spatial audio, IMU, eye-tracking) dataset with precise cross-sensor sync.

- Designed **action-grounded annotations** enabling next-best-action assistance.

- Shipped a **privacy-first** pipeline (VRS chunking, IRB workflows, blur filters, Label Studio UI).

- Built on-device captioning/navigation with **sub-second latency** on Meta Aria glasses.

- Benchmarked Qwen-VL, LLaVA, SEED-LLaMA; mapped failure modes; tools open-sourced.

Macquarie University (Remote)   *Jul 2023 – Oct 2023*
Research Intern (Advisor: Asst. Prof. Usman Naseem)   *Sydney, Australia*

- Co-created **ENeMeme**: one of the first Nepali–English multimodal meme sentiment datasets (5,000+ items).

- Built annotation pipeline: text normalization, code-switch handling, visual filtering, sarcasm/toxicity tagging.

- Developed multimodal baselines (mBERT + CNN visual features), improving cross-lingual robustness by **17%**.

- Integrated **fairness audits** for dialectal bias, slang coverage, and cultural-context handling.

Delhi Technological University   *Jan 2022 – Jun 2023*
Undergraduate Researcher (Advisor: Asst. Prof. Priya Singh)   *New Delhi, India*

- Conducted a systematic review of 43 deep-learning methods for trajectory prediction.

- Designed a three-axis taxonomy (social-awareness, output type, DL technique).

- Contributed NGSIM-based comparative analysis and identified gaps in socially aware prediction.

SELECTED PROJECTS

Tiny-ACE: Self-Improving Small Language Models `[GitHub]` *2025*

- Implemented ACE-style reflection loops for small LMs; built reproducible benchmarking suite measuring accuracy, latency, and memory efficiency on edge hardware.

Aria Glasses + Qwen-VL: On-Device Assistive Captioning `[GitHub]` *2025*

- Delivered privacy-preserving, real-time scene captioning on Meta Aria (Gen 1), reducing end-to-end latency by **35%** with spoken feedback and fully on-device inference.

Clinical Notes Entity Search on AWS `[GitHub]` *2025*

- Built HIPAA-aligned, serverless NLP pipeline using AWS Lambda, Comprehend Medical, Athena, and a Streamlit dashboard; achieved p95 ≤ 1.5s latency and high recall on Problems/Medications.

AgentOps for Supply Chain `[Demo]` *2025*

- Designed deterministic control tower converting EDI/ERP error streams into incident cards with cause, impact, and next-best action; added replayable "flight recorder" for debugging.

AWARDS & HONORS

ICCR Scholar (Government of India) — Full B.Tech Scholarship (2019–2023)
Science Olympiad Gold Medalist — TechFest, IIT Bombay (2018)
Perplexity.AI Campus Partner — UMD Outreach Lead (2025)
Reviewer — ACM TheWebConf 2025 (MM4SG Workshop)
Silver Medalist in Shotokan Karate-Do International Federation SKIF, affiliated with Nepal Olympic Committee (2014-2015)

TECHNICAL SKILLS

Languages: Python, C++, C
AI/ML: PyTorch, TensorFlow, Transformers, scikit-learn, OpenCV, Meta Aria SDK
Systems: Docker, Git, AWS, GCP, Neo4j, SQL/NoSQL, Weights & Biases
Other: HTML/CSS/JS, LaTeX

REFERENCES

Dr. Mengyu Wang
Director, Harvard AI & Robotics Lab, Schepens ERI
mengyu_wang@meei.harvard.edu

Dr. Snehesh Shrestha
Postdoctoral Researcher, NIST | Adjunct Faculty, UMD
snehesh@umd.edu

Dr. Cornelia Fermüller
Research Scientist, Perception & Robotics Group, UMD
fermulcm@umd.edu

Dr. Usman Naseem
Assistant Professor, Macquarie University
usman.naseem@mq.edu.au

Last updated: November 2025