Suresh Raghu

+91-7904210795 | sureshraghu0706@gmail.com | linkedin.com/in/suresh-raghu | github.com/R-Suresh07

Education

Indian Institute of Technology, Madras — Tamil Nadu, India
B.Sc. in Programming & Data Science; CGPA: 3.3/4.0, Project CGPA: 4.0/4.0 — *Oct. 2020 – Jun 2026*

Vellore Institute of Technology, Bhopal — Madhya Pradesh, India
B.Tech in Computer Science and Engineering; CGPA: 3.6/4.0 — *Oct. 2020 – Jun 2024*

Research Interests

Primary Areas: Uncertainty Quantification (UQ), Reliability & Safety of Reasoning Models, Vision-Language Models (VLMs).
Focus: Calibration & Proper Scoring for Agentic UQ, Selective Prediction & Low-Cost Uncertainty Estimation, Hallucination Detection & Visual Grounding in Reasoning VLMs.

Publications & Pre-Prints

Proper Scoring Rules for Agentic Uncertainty Quantification.
Suresh Raghu*, *Satwik Pandey**, *Shashwat Pandey* — *(Under Review)*
arXiv:2605.24756 | **Accepted (poster) at CTB & FAGEN Workshops @ ICML 2026**

- Introduced a family of **strictly proper trajectory-level scoring rules** for evaluating uncertainty in LM agents, with a **censored-trace extension** and negative results showing that standard trajectory-level calibration metrics, including **ECE** and **Brier**, are not strictly proper in the agentic setting.

SELFDOUBT: Uncertainty Quantification for Reasoning LLMs via the Hedge-to-Verify Ratio.
*Satwik Pandey**, ***Suresh Raghu****, *Shashwat Pandey* — *(Under Review)*
arXiv:2604.06389 | **Accepted (poster) at FAGEN Workshop @ ICML 2026**

- Proposed the **Hedge-to-Verify Ratio (HVR)**, a single-pass $\mathcal{O}(1)$ uncertainty signal that scores a reasoning trace by how much it hedges versus self-verifies. Traces with no hedging language are correct **96.1% of the time** (at 25.4% coverage), and fusing HVR with the model's verbalized confidence beats sampling-based Semantic Entropy on discrimination ($p = 0.001$) at **10× lower inference cost**, across 7 models and 3 benchmarks.

Don't Blink: Evidence Collapse during Multimodal Reasoning.
Suresh Raghu*, *Satwik Pandey** — *(Under Review)*
arXiv:2604.04207 | * Equal contribution

- Identified **"evidence collapse"**, a universal (**9/9** model×dataset cells) decay of visual grounding during VLM reasoning that text-only entropy cannot detect, and designed a task-conditional **vision veto** that cuts selective risk by up to **1.9 pp** at 90% coverage on MathVista, HallusionBench, and MMMU_Pro.

Research & Engineering Experience

VFS Global — New Delhi, India
Manager – AI (Founding Lead, AI Research Engineering) — *May 2024 – Present*

- **Adaptive Query Routing:** Architected the document-extraction pipeline around an adaptive router that sends deterministic documents to lightweight parsers and ambiguous ones to reasoning VLMs, reaching **99.1% field-level extraction accuracy**.
- **Uncertainty-Gated Extraction Cascade:** Used a cheap token-logit trigger to run **Semantic Entropy** and Monte Carlo self-consistency only on suspect outputs, escalating flagged cases to a larger LLM and cutting critical extraction errors by **87%**.
- **Biometric Compliance Pipeline:** Proposed and built a two-stage visa photo-verification system consisting of a cheap image-quality gate (Laplacian blur detection) followed by a VLM stage for pose/occlusion checks, reducing cases requiring manual intervention by **78%**.

Adani Digital Labs — Gurugram, India
Applied Machine Learning Intern — *June 2023 – April 2024*

- **Self-Supervised Representation Learning:** Used Meta's **DINOv2** embeddings with an SVM over latent-space projections for few-shot image clustering, replacing manual tagging at **93% precision**.
- **Multimodal Search:** Built a CLIP + GPT-4 Vision search engine supporting both text and image queries over **proprietary retail inventory**.