

AAFIYA HUSSAIN

+1 (540) 934-8040 [✉ aafiyahussain@vt.edu](mailto:aafiyahussain@vt.edu) [in LinkedIn](#) [@ GitHub](#) [🎓 Google Scholar](#)

EDUCATION

Virginia Tech

Master of Science in Computer Science (Fully funded - **\$62,705/year** Scholarship), **GPA: 4.0/4.0** Blacksburg, Virginia
Aug 2024 - May 2026 (Expected)

- Graduate Teaching Assistant for CS5704, 5244, 5814, 1114 (Fall 24 – Spring 26).

K.J. Somaiya College of Engineering

B.Tech in Computer Engineering, **GPA: 9.44/10** Mumbai, India
Jul 2018 - Jul 2022

EXPERIENCE

Nokia Bell Labs

Computer Vision Intern (Industrial Metaverse) Jun 2025 - Aug 2025
Murray Hill, NJ, USA

- Led development of an end-to-end multimodal LLM and multi-view stereo pipeline to automate querying of complex 3D scenes by building structured scene embeddings from RGB images, **reducing manual annotation effort by ~95%**.
- Built an interactive visualization layer with 3D Gaussian splats enabling users to query scenes through a chat interface.
- Improved performance over diverse lighting conditions, heterogeneous sensors, and challenging occlusions (**10+ cases**), ensuring system robustness across diverse data-collection pipelines (**accuracy variance < 5%**).
- Enhanced the Gaussian Splat reconstruction framework using accurate ChArUco-board calibration, yielding **~25% higher geometric fidelity and ~18% better depth precision** (*patent submitted*).

Madan Bhandari University of Science and Technology

Research Scholar Jun 2023 - Jul 2024
Bagmati, Nepal

- Fine-tuned Whisper models for automatic speech recognition (ASR) of **13+ Indic Languages** to caption video-based explanations of undergrad computer science coursework, eliminating manual effort of translation.
- Integrated the ASR module into the University's learning management system, **improving student engagement by ~60%**.
- Built a BERT + classical NLP based dynamic quiz-generation system that extracts key content, performs negation and word sense disambiguation, and auto-creates customizable quizzes, **boosting average student performance by 23%**.

Deloitte

Software Development Engineer (Analyst) Nov 2022 - Jun 2023
Mumbai, MH, India

- Led the development of an automated fraud detection system processing **50K+ daily transactions**, achieving **94% accuracy** and flagging suspicious transfers for manual review.
- Deployed production-ready ensemble models to score client risk from financial profiles, achieving **92% accuracy** and strengthening end-to-end decision-making in compliance workflows.

SELECTED PUBLICATIONS

- G. Srivastava, **A. Hussain**, Z. Bi, S. Roy, P. Pitre, M. Lu, M. Ziyadi, and X. Wang. "BeyondBench: Benchmark-Free Evaluation of Reasoning in Language Models." [arXiv](#) | [leaderboard](#) . (ICLR 2026).
- P. Zhou, A. Asgarov, **A. Hussain**, W. Park, A. Paudyal + 7 authors. "Designing Multi-Robot Ground Video Sensemaking with Public Safety Professionals." [arXiv](#) (CHI 2026).
- **A. Hussain**, G. Srivastava, A.M. Ishmam, and C. Thomas. "SoundBreak: A Systematic Study of Audio-Only Adversarial Attacks on Trimodal Models." [arXiv](#) Under Review at **ACL 2026**.
- G. Srivastava, **A. Hussain**, C. Wang, Y. Celine and X Wang. "effGen: Enabling Small Language Models as Capable Autonomous Agents." Framework available at [effGen.org](#) | [arXiv](#) Under Review at **ICML 2026**.

*Complete list of publications - [Google Scholar](#) , citations: 176

SELECTED PROJECTS

LLMThinkBench Python | vLLM | PyTorch | Transformers

[GitHub](#) | [PyPI](#) | [Leaderboard](#)

- Developed a **dynamic math reasoning evaluation framework** to fairly measure LLMs' innate reasoning ability without benchmark contamination, covering **14 arithmetic tasks** with dual-backend inference system and multi-GPU parallelism.
- Engineered a production-ready PyPI package (**4.92k+ downloads**) with **17 configurable parameters**, **140+ regex-based answer extraction methods**; released leaderboard with **results from 55+ LLMs**.
- Proposed a novel **Overthinking Score** metric balancing accuracy with token efficiency using F1-harmonic mean.

DataSense - Multi-Agent Data Visualization LLM | Plotly | Streamlit

[GitHub](#)

- Built an **LLM-agent ensemble** using cost-effective small language models to analyze datasets, visualization patterns, and automatically create the **top 3 charts from 9+ options**, reducing manual effort spent in exploratory data analysis.
- Developed an end-to-end data pipeline, processing **10+ file formats** into interactive graphs using Plotly integration.
- Engineered a storytelling module that generates contextual insights and narratives, **reducing manual effort by ~75%**.

TECHNICAL SKILLS

Programming Languages

Python, Java, JavaScript

Libraries & Frameworks

PyTorch, TensorFlow, Transformers, vLLM, scikit-learn, OpenCV, LangChain

Technologies

Flask, MySQL

Tools

AWS, FastAPI, Docker