Samiha Mirza

Samiha.mirza1234@gmail.com | 346-599-6417 | Houston, TX | in/samiha-mirza | orcid/samiha | Website/Samiha | Github/Sam

Applied AI, Computer Vision & Scientific Imaging | PhD Candidate in CS (Aug 2026)

PhD candidate in Computer Science specializing in computer vision, deep learning, and scientific imaging. Experienced in developing image segmentation models, representation learning frameworks, and scalable ML pipelines for large-scale imaging datasets. Skilled in PyTorch-based model development, multi-GPU training, and data-centric AI methodologies for improving robustness and generalization across heterogeneous data sources.

WORK EXPERIENCE

Shell **June 2024 – Aug 2024**

Data Science/AI Research Intern *Houston, TX*

- Analyzed enterprise application source code and system interactions to identify interdependencies across digital delivery workflows.
- Developed graph-based dependency mappings & visualizations to analyze enterprise system interactions & operational workflows.
- Built ML pipelines on large-scale operational datasets including preprocessing, feature engineering, & model evaluation.
- Designed Power BI dashboards and reporting tools to communicate operational insights, workflow dependencies, and system-level trends to cross-functional teams.
- Collaborated with engineering and digital delivery teams to support operational decision-making and process visibility initiatives.

Shell | Quantitative Imaging Lab, University of Houston **Aug 2022 – Present**

Research Assistant | Aug, 2022 – Present *Houston, TX*

- Defined research problems, designed experimental protocols, and developed deep learning and StyleGAN-based generative frameworks for large-scale imaging applications.
- Designed data-centric evaluation frameworks to quantify how imaging variability, noise, and distribution shift impact segmentation performance, feature representations, and model reliability.
- Built reproducible, scalable ML pipelines in Python/PyTorch with experiment tracking and multi-GPU training for systematic evaluation and comparison.
- Published peer-reviewed research on data-centric AI approaches for improving robustness in real-world imaging systems.

Teaching Assistant | Aug, 2022 – Spring, 2024

- Taught *Digital Image Processing, Data Structures,* and *Computer Architecture*; mentored 100+ students on algorithm design and performance-aware coding practices.

EDUCATION

University of Houston **Aug 2022 – Aug 2026**

PhD, Computer Science *Houston, TX*

University of Houston **June 2025**

MS, Computer Science *Houston, TX*

University of Dammam **June 2022**

Bachelor's, Computer Science (Summa Cum Laude) *Saudi Arabia*

SKILLS & COMPETENCIES

- **Technical Skills & AI Modeling**: Python, PyTorch, TensorFlow, OpenCV, NumPy, Pandas, Linux, Deep learning, image segmentation, representation learning, generative modeling, multi-GPU training, predictive modeling, model evaluation, domain adaptation, robustness analysis, feature engineering
- **Professional:** Cross-disciplinary collaboration, reproducible research, agile development, scientific writing

SELECTED PROJECTS

Shell **Aug 2022 – Present**

Data-Centric Evaluation of Deep Learning Models for Large-Scale Imaging Data

- Designed controlled experiments to quantify how data variability impacts ML performance, improving cross-domain generalization (Accuracy improvement by ~4%).
- Studied model behavior under real-world variability to improve robustness in deployment-like conditions.

Large-Scale Deep Learning for Image Segmentation and Representation Learning

- Built CNN and 3D UNet models large scale image segmentation tasks using multi-GPU training pipelines.
- Investigated representation learning and model robustness across heterogeneous imaging datasets under distribution shift.

Risk Categorization and Predictive Analytics for Operational Systems

- Developed ML models on historical well-event datasets to categorize risks and analyze potential mitigation strategies.
- Investigated temporal patterns and feature relationships associated with downstream operational outcomes, emphasizing model interpretability and decision relevance in industrial workflows.