Md Marufi Rahman

Dallas, Texas | mdmarufirahman@my.unt.edu | (940) 268-4062 | Github: https://mdmarufirahman.github.io |

LinkedIn | Authorized to work in the United States for any employer (U.S. Permanent Resident)

Career Summary

Ph.D. candidate in Computer Vision with 7+ years of experience designing scalable, real-time perception systems for autonomous and safety-critical applications. Proven track record of developing motion estimation, SLAM-inspired depth matching, and robust vision algorithms under real-world constraints. Strong programming skills in C++ and CUDA for high-performance deployment, with a research focus aligned with onboard UAV/drone perception, sensor fusion, and multi-view geometry.

Education

University of North Texas (UNT), Texas, USA   *Aug 2019 - present*

- Ph.D. candidate, Computer Science, GPA: 3.85/4.0

Rajshahi University of Engineering & Technology (RUET), Bangladesh   *Mar 2013 - Dec 2017*

- B.Sc., Electronics and Telecommunication Engineering, GPA: 3.32/4.0

Skills

Languages & Tools: C++, Python, R, Matlab, SQL, CUDA, HIP, OpenCL, Git, Docker, CI/CD, GCP, Azure, AWS, Agile/Scrum

Frameworks & Libraries: PyTorch, TensorFlow, Keras, GTSAM, Ceres Solver, MediaPipe, ROS (basic familiarity), OpenCV, HuggingFace, ONNX, MediaPipe, Pandas, NumPy

Core Expertise: SLAM, Multi-view Geometry, Sensor Fusion (GPS/IMU), Camera Calibration, Real-time Perception, Motion Estimation, Model Optimization, Vision-Language Models (BLIP, CLIP, VQA), Generative AI (Diffusion, GANs), Familiarity with UAV/drone systems and onboard real-time vision deployment; experience simulating aerial perturbations for vision robustness testing.

Experience

Research Assistant, *Multimedia Information Group - UNT, TX*   *Aug 2020 - Present*

- Led the development of scalable deep learning pipelines for medical video analysis using PyTorch and CUDA, achieving 4× speedup in training and 2.5× reduction in inference latency over baseline TensorFlow-based workflows on GCP and Azure clusters.
- Designed a monocular motion estimation pipeline using lightweight affine modeling to enhance scene coverage in video sequences (a method analogous to visual odometry and onboard drone perception). The system supports real-time inference with GPU acceleration and improves frame-to-frame coverage without dense 3D reconstructions.
- Developed a stereo disparity-based similarity detection algorithm (depth-aware content retrieval), achieving a 92.5% F1-score in image matching and reducing redundancy by 38%. This work shares principles with SLAM and autonomous navigation pipelines in robotics
- Developed a BLIP-2-based vision-language model fine-tuned on the MedVQA dataset for image-to-text prompting, improving ROUGE-L score by 12% and reducing data labeling needs by 30% through uncertainty-based active sampling.
- Implemented real-time AR face filters using diffusion models (DreamBooth, ControlNet) aligned via MediaPipe landmarks, deployed as a 60 FPS OpenCV-CUDA application - demonstrating real-time CV deployment for constrained platforms.
- Wrote custom CUDA kernels and optimized inference pipelines using Nsight and PyTorch Profiler with strict real-time performance constraints, mirroring embedded CV stack development for UAV/autonomous platforms.
- Simulated real-world UAV camera perturbations (zoom, warp, elastic deformation) to test perception model robustness under environmental variance - a crucial step for autonomous aerial navigation validation.
- Collaborated across multi-institutional research teams, mentored junior researchers, and contributed to research proposals and patent drafts in motion-guided video interpretation and encoder optimization.

Teaching Fellow, *University of North Texas, TX*   *Aug 2019 - Present*

- Instructed and supported undergraduate and graduate students across core CS courses including Database Systems, Data Structures,
- Mentored 40+ students per semester and created hands-on programming assignments in C++, SQL and algorithm design topics.

Selected Publications

1. Y. Li, J. Baik, **M. M. Rahman**, I. Anagnostopoulos, R. Li, and T. Shu, Pareto optimization of CNN models via hardware-aware neural architecture search for drainage crossing classification on resource-limited devices, *Proceedings of the SC '23 Workshops of The International Conference on High Performance Computing, Network, Storage, and Analysis*, Denver CO USA, 2023.
2. M. M. Rahman, JH Oh, Wallapak Tavanapong and Piet C. de Groen, Content Based Image Retrieval Using Depth Maps for Colonoscopy Images, *Proceedings of the 16th International Joint Conference on Biomedical Engineering Systems and Technologies*, Lisbon, 2023.

Selected Award

- National Science Foundation (NSF)-sponsored Online Workshop on Deep Learning Systems in Advanced GPU Cyberinfrastructure (DL-GPU) scholarship (2023).