Mostofa Rafid Uddin

412-626-2812 mru@andrew.cmu.edu, rafid.uddin95@gmail.com https://duranrafid.github.io
mostofa-rafid-uddin https://scholar.google.com/citations?user=csnAp98AAAAJ

Research Interests & Expertise

Research Topics of Interest: *High Expertise:* Unsupervised, Self-supervised Representation Learning, Contrastive Learning, 3D Computer Vision, Structural Bioinformatics, Object detection and Segmentation.
Moderate Expertise: Deep Generative Models, Probabilistic Graphical Models, Foundation Models, Domain adaptation, Pose Estimation.
Datasets of expertise: Scene-centric images, 3D tomographic images, Microscopy images, 3D data with various representations (mesh, point-cloud, SDF, voxels, Gaussian Splat), Biological Sequence Data (Protein, DNA, RNAseq).

Education

2020- Present	**Doctor of Philosophy (Ph.D), School of Computer Science**, Carnegie Mellon University, Pittsburgh, PA 15213, USA **Advisor: Min Xu** Distinctions: CMLH Fellowship for Digital Health 2023. **Relevant Coursework**: Machine Learning (Ph.D.) level, Computer Vision, Probabilistic Graphical Models.
2014 -2018	**Bachelor of Science in Computer Science and Engineering**, Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh. Academic distinctions: Deans list awards, University merit scholarships. **Relevant Coursework**: Computer Graphics, Object-oriented Programming, Structured Programming, Pattern Recognition, Digital Image Processing.

Selected Research Publications

A few representative publications are mentioned here. For a full and up-to-date list, please visit my google scholar link.

- **Mostofa Rafid Uddin**, Gregory Howe, and Min Xu. Harmony: A Generic Unsupervised Approach for Disentangling Semantic Content from Parameterized Transformations.*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022, pp. 20646-20655. (H5-index: 356). [paper link] [news link]. Skills: Unsupervised Learning, Representation Learning, Deep Generative Models, Bio-image Analyses.
  - In many real-life image analyses, particularly biomedical research domains, objects in the images undergo several parameterized transformations.
  - I developed an unsupervised method to disentangle the transformations from image contents and demonstrate that it significantly facilitates many downstream tasks.
- **Mostofa Rafid Uddin**, Sazan Mahbub, M Saifur Rahman, and Md Shamsuzzoha Bayzid. SAINT: Self-Attention Augmented Inception-Inside-Inception Network Improves Protein Secondary Structure Prediction. *Bioinformatics*, 2020 Nov 1; 36(17):4599-608. (H5-index: 136, Impact Factor: 6.937) [paper link]. Skills: Neural Machine Translation, Structural Bioinformatics, Dense Prediction
  - Predicting 8-state (Q8) secondary structure from amino acid sequences of protein is an important but challenging problem.
  - Developed a self-attention augmented inception-inside-inception network that improves state-of-the-art protein secondary structure prediction and recovers insights of protein folding through interpretable attention features.
- Sayali Onkar, Jian Cui, Jian Zou, Carly Cardello, Anthony R Cillo, **Mostofa Rafid Uddin**, April Sagan, Marion Joy, Hatice U Osmanbeyoglu, Katherine L Pogue-Geile, Priscilla F McAuliffe, Peter C Lucas, George C Tseng, Adrian V Lee, Tullia C Bruno, Steffi Oesterreich, Dario AA Vignali. Immune landscape in invasive ductal and lobular breast cancer reveals a divergent macrophage-driven microenvironment. *Nature Cancer*(Impact Factor: 23.18). [paper link] Skills: Statistical Image Analysis, Applied Science Research.
  - Developed a pipeline for spatial cell neighborhood analysis in multispectral tumor microenvironment images.
  - Demonstrated the role of different cell phenotypes in tumor microenvironment from spatial analysis.

Technical Skills

Languages: Python, Java, C, C++. **HPC Computing:** AMD Cluster, Oracle Cloud, AWS.
Frameworks: Pytorch, Detectron, Tensorflow, OpenCV, Numpy, Scipy, Scikit-learn.

Work Experience

2021 - Present	**Graduate Research Assistant**, Computational Biology Department, School of Computer Science, Carnegie Mellon University, PA, USA
2022	**Graduate Teaching Assistant**, Computational Biology Department, Spring 2022 CMU 02-620: Machine Learning for Scientists Fall 2022 CMU 02-740: Bioimage Informatics
2019- 2020	**Lecturer**, Department of Computer Science and Engineering, East West University, Dhaka, Bangladesh.

Grants, Awards, & Services

- **Center for Machine Learning and Health (CMLH) fellowship in Digital Health, 2023** CMLH fellowships, around 100,000 USD worth, are awarded each year to several (around 10) outstanding digital health-related research proposals by CMU PhD students. I received the award in 2023 with my proposal *"Leveraging Cryo-ET Imaging Technology to Improve Patient Care for Neurodegenerative Diseases by Identifying Subcellular Biomarkers"*. [link] Skills: Grant Writing, Independent Research.
- Regularly serve as a reviewer in top AI and vision conferences such as CVPR, ICCV, ECCV, WACV, AAAI, etc.
- Worked as a mentor in CMU AI Mentoring Program, where I mentor CMU undergraduate students coming from underrepresented communities interested in AI research
- Gave research talk on IEEE Applied Imagery Pattern Recognition (AIPR) Workshop (virtual), October 2021, Washington, DC, US. [*link*]
- Won best poster award at 3rd International Conference on Networking, Systems and Security (NSysS 2017). Poster Title: *Archiving Medical Records in DNA Sequence*[*pdf*] Skills: Precision Health, Electronic Health Records

Mini-Research Projects

♢ **Design of Phase-separated Protein Sequences using Adaptive Sampling and Active Learning** We address the problem of in silico protein design with a high propensity for liquid-liquid phase separation (LLPS) and droplet formation. We developed an effective method consisting of multiple components, including a relaxed "energy" based sequence generator, a biochemical condition-aware attention-neural network-based surrogate model, a Bayesian acquisition function, and its optimizer. [*github*] Skills: Probabilistic Graphical Models, Protein Design, Optimization.

♢ **Pytorch Implemented Local Energy Minimizer** Implemented the local energy minimizer module of OpenMM software by modifying pytorch autograd mechanics. [*github*]

♢ **Edge prediction: Predicting Edge in Academic Citation Networks** Predicted how likely an academic article is to cite another particular article using an intelligent and novel feature engineering pipeline that could generate highly accurate predictions with relatively simpler models. Around 95% F1 score was achieved with random forest classifier with our engineered features, which largely outperformed the graph neural network-based model on 3 different academic citation network datasets. [*github*]

♢ **Predicting age from lung single cell data** Applied multiple feature extraction models and classifiers to predict biological age from scRNA-seq data of multiple control patients. [*github*]

♢ **Onubadok: Bangla to English Machine Translation Using Seq2Seq Model with Attention Mechanism.** In this project on Neural Machine Translation(NMT), I observed that using Bahdanau's attention with a vanilla encoder-decoder model improves BLEU score for Bangla to English translation. [*github*]

♢ **Arduino based Posture Corrector.** Developed a posture corrector android application that could detect unusual bending of user wearing a device containing flex sensor. [*youtube*]