

# Md Mohaiminul Islam

✉ [mmiemon@cs.unc.edu](mailto:mmiemon@cs.unc.edu)

🌐 [Webpage](#)

🌐 [LinkedIn](#)

🔍 [Google Scholar](#)

🐙 [GitHub](#)

## Research Interest

- Machine Learning    📖 Deep Learning, Multi-Modal Learning, Self-Supervised Learning, Unsupervised Clustering, Large Language Modeling, Natural Language Processing.
- Computer Vision    📖 Video Understanding, Long-Range Video Modeling, Image and Video Recognition, Object Detection and Tracking, Vision-Language Modeling.

## Education

- 2021 – Present    📖 **Ph.D. in Computer Science**, 3rd Year, UNC Chapel Hill  
Advisor: [Gedas Bertasius](#)  
Courses: *Deep learning, Advanced Topics in Video Understanding, Self-Supervised Visual Representation Learning, Visual Recognition with Transformers, Connecting Language to Vision and Robotics.*
- 2014 – 2018    📖 **B.Sc. in Computer Science and Engineering**  
*Bangladesh University of Engineering and Technology*  
CGPA 3.86/4.00 (3.93 final 2 years)

## Experience

- May 23 – Aug 23    📖 **Research Scientist Intern**, FAIR Accel, [Meta AI](#)  
Video Understanding, Egocentric Vision
- May 22 – Aug 22    📖 **Research Intern**, [Comcast AI](#)  
Efficient Movie Scene Detection using State-Space Transformers. [\[pdf\]](#)
- Apr 19 – Dec 20    📖 **Lecturer**, University of Asia Pacific
- Nov 18 – Mar 19    📖 **Software Engineer**, Samsung Research

## Research Publications

- 4 Efficient Movie Scene Detection using State-Space Transformers  
**M. Islam**, M. Hasan, K. Athrey, T. Braskich, G. Bertasius  
In *Computer Vision and Pattern Recognition (CVPR)* 2023. [\[pdf\]](#)
- 3 Long Movie Clip Classification with State-Space Video Models  
**M. Islam**, G. Bertasius  
In *European Conference on Computer Vision (ECCV)* 2022. [\[pdf\]](#)
- 2 Object State Change Classification in Egocentric Videos with Divided Space-Time Attention  
**M. Islam**, G. Bertasius  
In *1st Ego4D workshop, CVPR* 2022. [\[pdf\]](#)
- 1 COVID-DenseNet: A Deep Learning Architecture to Detect COVID-19 from Chest Radiology Images  
**M. Islam**, T. Hannan, L. Sarker, Z. Ahmed  
In *International Conference on Data Science and Applications (ICDSA)* 2022. [\[pdf\]](#)

## Selected Projects

ViS4mer	■	An efficient video classification model for long-range videos utilizing a transformer encoder and a novel-proposed multi-scale state-space decoder.
TranS4mer	■	An efficient model to detect scene boundaries in long-range movies that uses a novel state-space self-attention (S4A) building block.
VidSum	■	Summarizing hour-long videos utilizing pretrained vision-language (CLIP) and large language models (GPT).
EgoOSCC	■	A transformer architecture using the divided space-time attention mechanism that can detect object state changes in egocentric videos.
Audio2Video	■	A conditional GAN approach to generate plausible videos from natural sounds. Proposed a 3D CNN model for video and CNN-RNN model for audio.
COVID-DenseNet	■	A CNN-based model to detect COVID-19 from chest radiology images which can localize the critical regions of affected images using a GRAD-CAM-based module.
ProteinSeq	■	A CNN-LSTM model with an attention mechanism to predict the three-dimensional structure of a protein from its amino acid sequence.
BenglaAI	■	A CNN-based model for Bengali Handwritten Digit Recognition.
eMed-DNA	■	Presented a proof-of-concept for efficient management of Electronic Health Records (EHRs) of a person inside his DNA sequence.
UniShare	■	An academic resource-sharing platform developed using PHP, CodeIgniter framework, MySQL, JavaScript, HTML, CSS, and XAMPP server.
PoliceBox	■	An interactive server-client system developed using Java Swing, JavaFX, MySQL, and Socket Programming to efficiently manage various services provided by the police.
C compiler	■	A simple C compiler using Flex and Bison.

## Skills

ML tools	■	PyTorch, TensorFlow, Keras, Jupyter, Distributed training.
Coding	■	Python, Java, C, C++, Java HTML, XML, PHP, JavaScript, 80x86 Assembly.
Databases	■	Oracle, MySQL, PostgreSQL, SQLite.
Web Dev	■	Html, CSS, JavaScript, Android, Apache Web Server.
Misc.	■	Lex, Yacc, Shell Script, LaTeX, Git, Docker.

## Miscellaneous Experience

### Awards and Honors

2022	■	<b>2nd place</b> , <i>Ego4D: Object State Change Classification Challenge</i> , Ego4d Workshop, CVPR 2022.
2018	■	<b>Champion Student Poster Award</b> , International Conference on Networking, Systems and Security, NSysS 2017.
	■	<b>9th place</b> , <i>Bengali Handwritten Digit Recognition Challenge</i> , Kaggle Competition 2018.
2016-18	■	<b>Deans List Award</b> , Bangladesh University of Engineering and Technology.
	■	<b>University Merit List</b> , Bangladesh University of Engineering and Technology.

### Leadership

2014-18	■	<b>Class Captain</b> Bangladesh University of Engineering and Technology.
---------	---	---