Postgraduate Seminar Series

Venue: Graduate Seminar Room
Date & Time: October 31, 2026 at 2:30 PM

Speaker Information



Tasmiah Tamzid Anannya (Std No. 04123054002) is a full time Ph.D student in the department of CSE, BUET. She completed her undergraduate studies from Military Institute of Science and Technology (MIST) in 2013 and M.Sc from BUET in 2018. Her research interest lies in the field(s) of Machine Learning, Deep Learning, Agentic AI, Spatio-temporal Data. She is currently doing her doctoral thesis under the supervision of Prof. Tanzima Hashem. She will be speaking about her ongoing PhD research in this talk.

From Footprints to Futures: Evaluating VLMs for Physical, Functional, and Temporal Urban Change Detection

Detecting change in the urban built environment is a fundamental challenge with wide-ranging applications across urban planning, policy-making, and socio-economic studies: from monitoring neighborhood transformation and infrastructure development to understanding patterns of displacement and economic restructuring. While conventional machine learning and deep learning approaches can identify distinct physical changes through imagery, they fundamentally lack the capacity to interpret broader contextual relationships, such as associating new construction with emerging amenities or inferring underlying socio-economic shifts from spatial patterns. The advent of Large Language Models (LLMs), in particular Vision Language Models (VLMs), presents a compelling opportunity to bridge this gap by combining visual perception with language-grounded reasoning. In this study, we evaluate the reasoning capability of VLMs in understanding urban space and their analytical ability to solve urban change detection tasks through multimodal inputs combining visual representations of building footprints and Points of Interest (POI) with textual reasoning queries. We design a suite of urban tasks spanning physical, functional, and temporal change detection, and systematically evaluate the capability of VLMs to perform these tasks. Building on this, we further assess VLMs' decision-making capability by asking them to rank urban areas based on features extracted from the preceding task groups. To validate the generated neighborhood rankings, we conduct a downstream case study comparing them against real-world gentrification trends. Extensive experiments across a range of VLMs demonstrate that while VLMs can meaningfully interpret urban map representations, smaller models struggle to capture fine-grained microscopic features. Our findings highlight the broader potential of VLMs for domain-specific urban applications, including mitigating gentrification-led displacement and optimizing transit planning.