



CYCLO: Cyclic Graph Transformer Approach to Multi-Object Relationship Modeling in Aerial Videos

Trong-Thuan Nguyen¹, Pha Nguyen¹, Xin Li², Jackson Cothren¹, Alper Yilmaz³, Khoa Luu¹

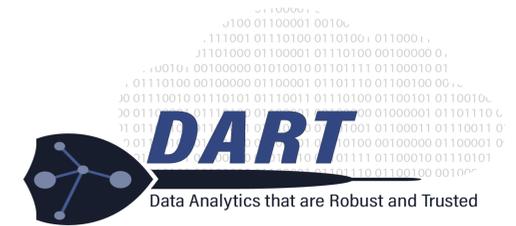
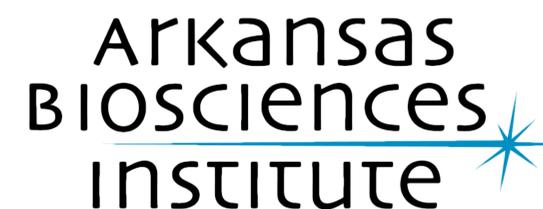
¹University of Arkansas

²State University of New York at Albany

³Ohio State University

<https://uark-cviu.github.io/projects/CYCLO/>

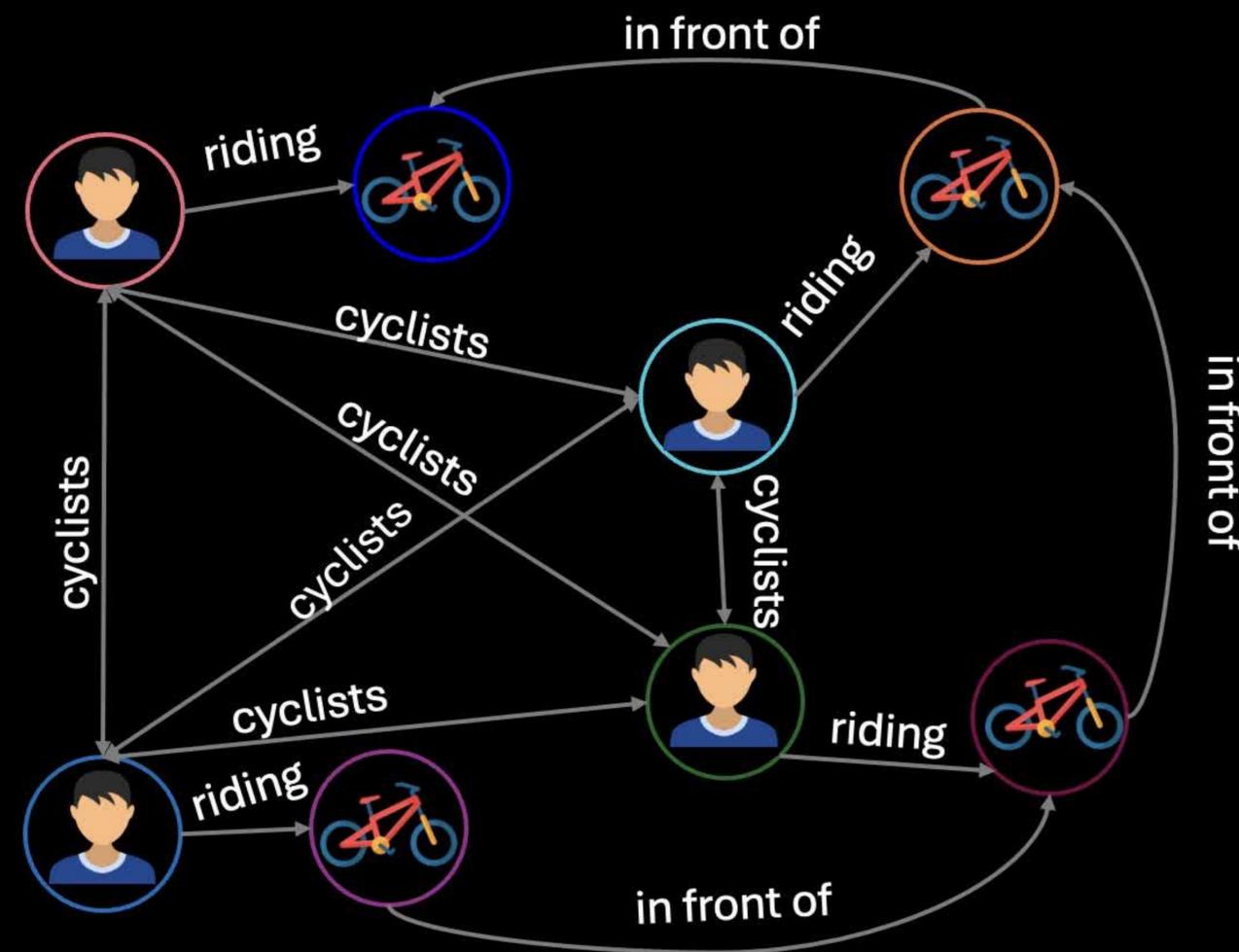
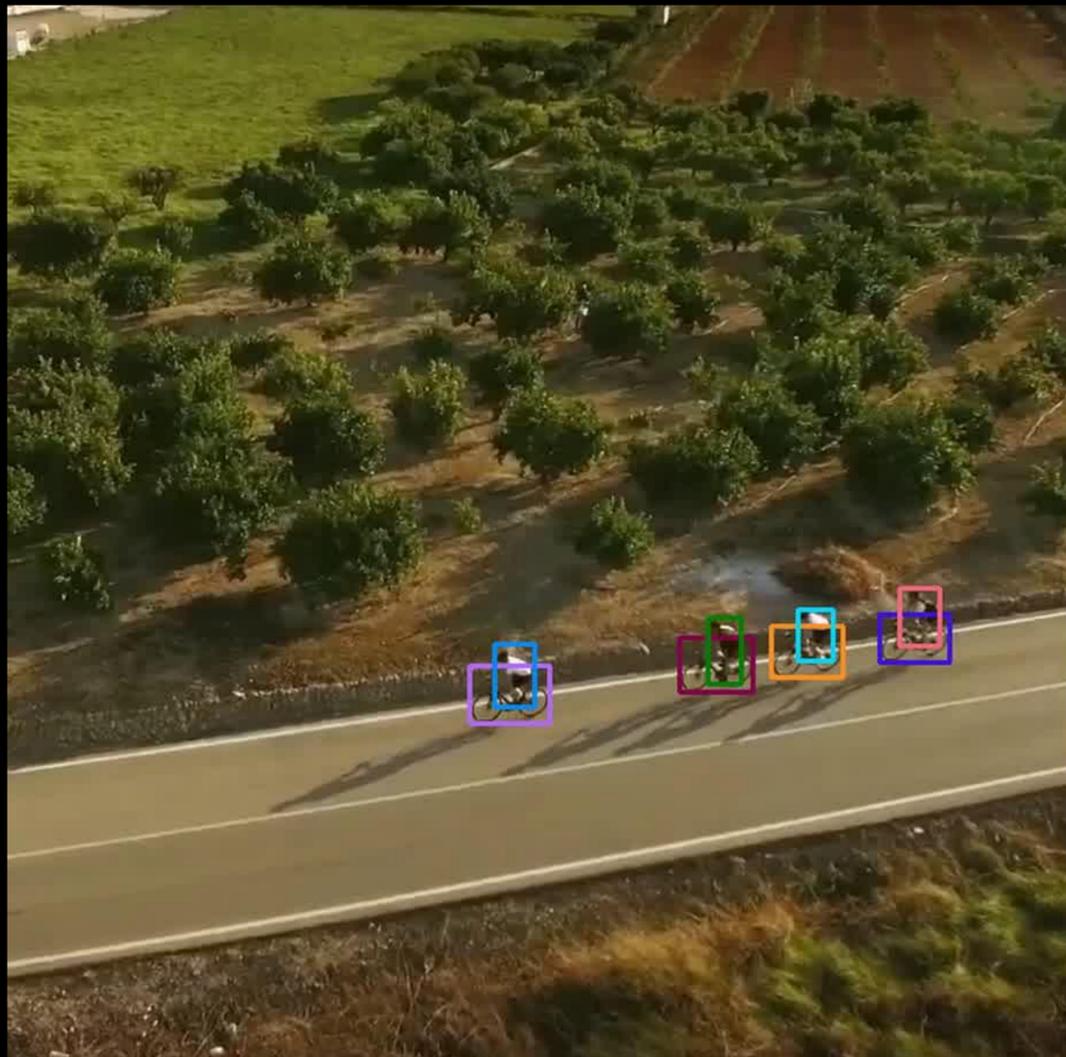
This work is partly supported by J.B. Hunt Transport Services (JBHunt), NSF Data Science and Data Analytics that are Robust and Trusted (DART), NSF SBIR Phase 2, and Arkansas Biosciences Institute (ABI) grants.



Problem Statement

Video Scene Graph Generation (VidSGG):

- ✓ *relationships between objects*
- ✓ *temporal transition of these relationships* across video frames.



Motivations

1

Complex Interactions in Aerial Videos:

- High object **density** and **diverse object categories**.

2

Dynamic and Diverse Environments:

- Varied settings from **urban to disaster-stricken areas**.

3

Limitations of Existing Solutions:

- Ground-level focus of current datasets and methods.
- Inadequate **temporal dependency modeling**.

4

Importance of Enhanced Scene Understanding:

- Critical applications in **surveillance, disaster response, and traffic management**.
- Improved **decision-making** through detailed scene graphs.

Contributions

1

AeroEye Dataset:

- First **VidSGG** dataset for aerial videos.
- Includes **2,260** videos, **261,503** frames, **57** object categories.
- Includes **384** predicates capturing **positions and relations**.

2

CYCLO Approach:

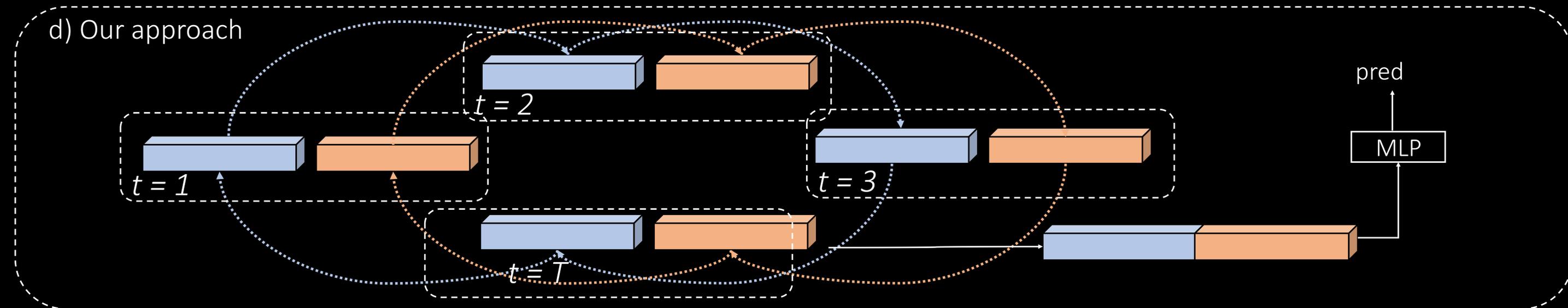
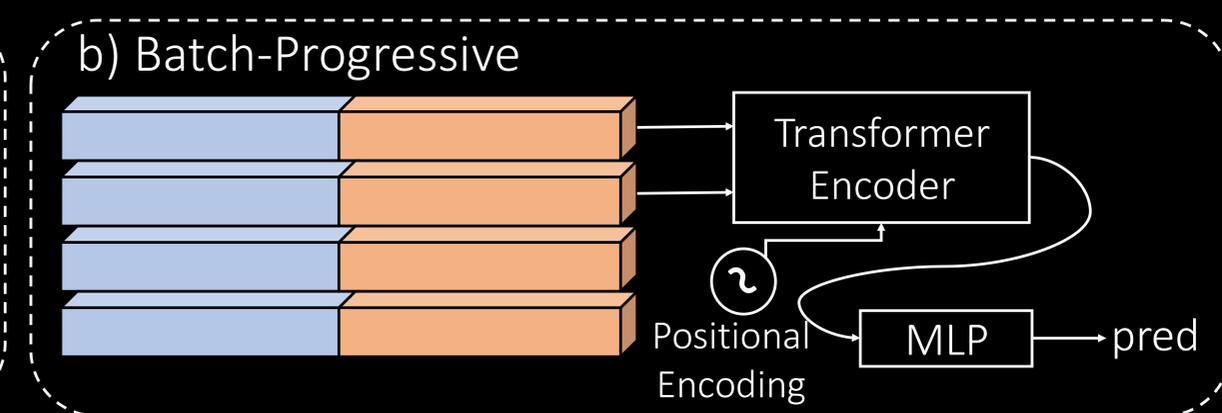
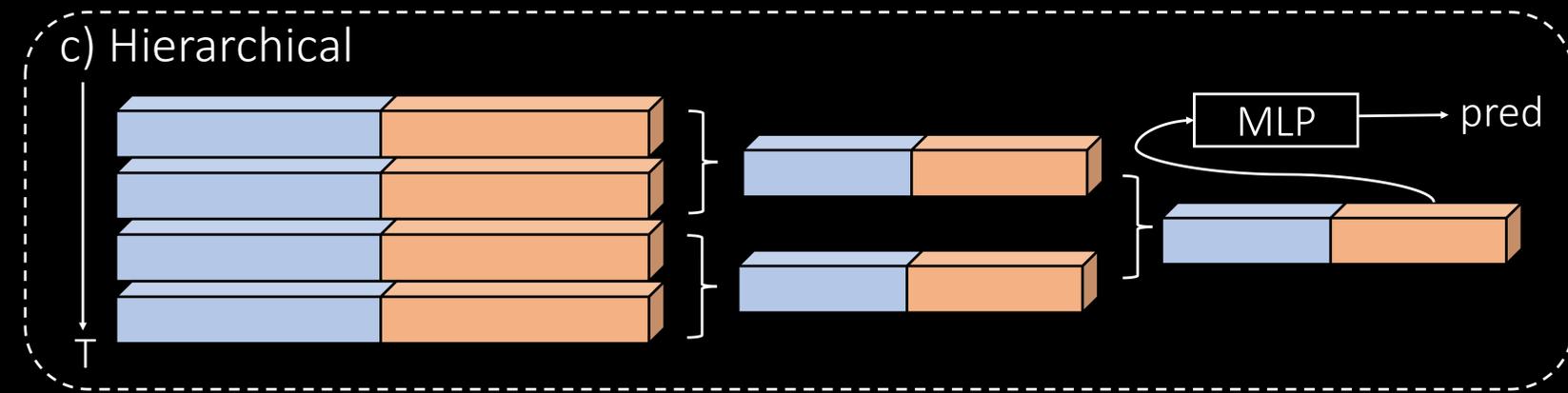
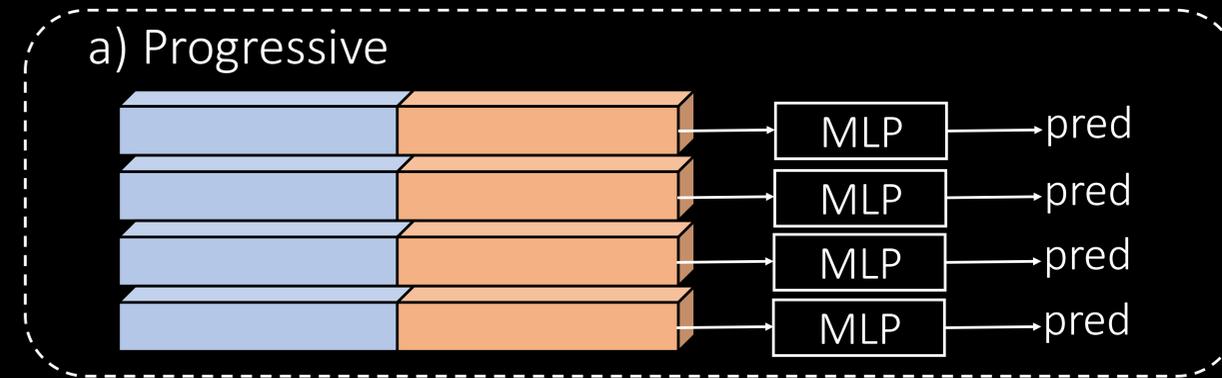
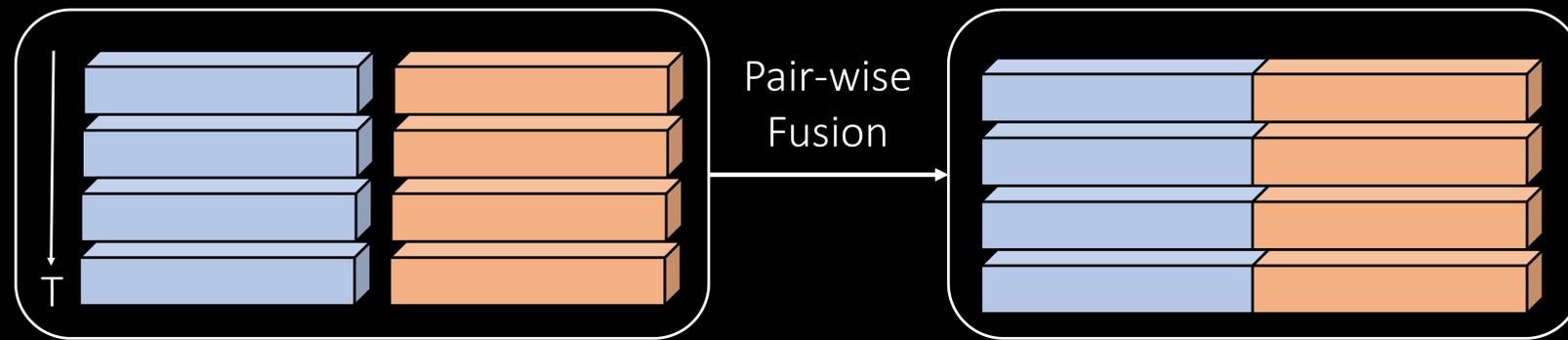
- Handles **periodic** and **overlapping relationships**.
- Captures **direct** and **long-range temporal dependencies**.
- Achieves **State-of-the-Art** performance on **aerial** and **in-the-wild** video datasets.

Our AeroEye Dataset

Dataset	#Frames	Annotations		Viewpoints				
		BBox	Relation	3rd	ego	aerial	oblique	ground
Action Genome	234.3K	✓	✓	✗	✗	✗	✗	✗
ASPIRe	1.6M	✓	✓	✗	✗	✗	✗	✗
SportsHHI	11.4K	✓	✓	✗	✗	✗	✗	✗
PVSG	153K	✓	✓	✓	✓	✗	✗	✗
VisDrone	261.9K	✓	✗	✗	✗	✓	✗	✗
UAVDT	40.7K	✓	✗	✗	✗	✓	✗	✗
UAVid	300	✓	✗	✗	✗	✗	✓	✗
MARVEC	537K	✓	✗	✗	✗	✗	✓	✓
AeroEye (Ours)	261.5K	✓	✓	✓	✓	✓	✓	✓

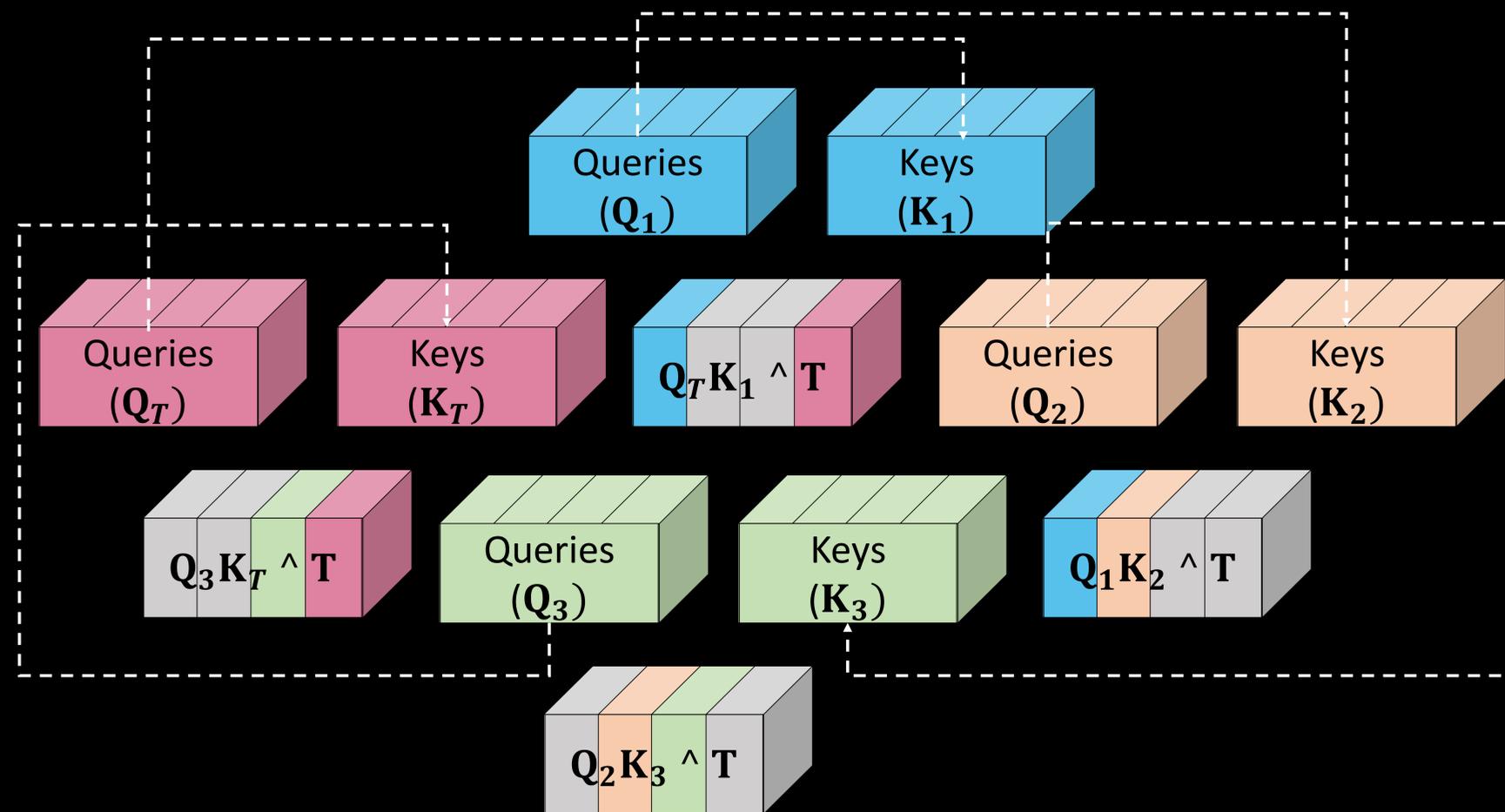


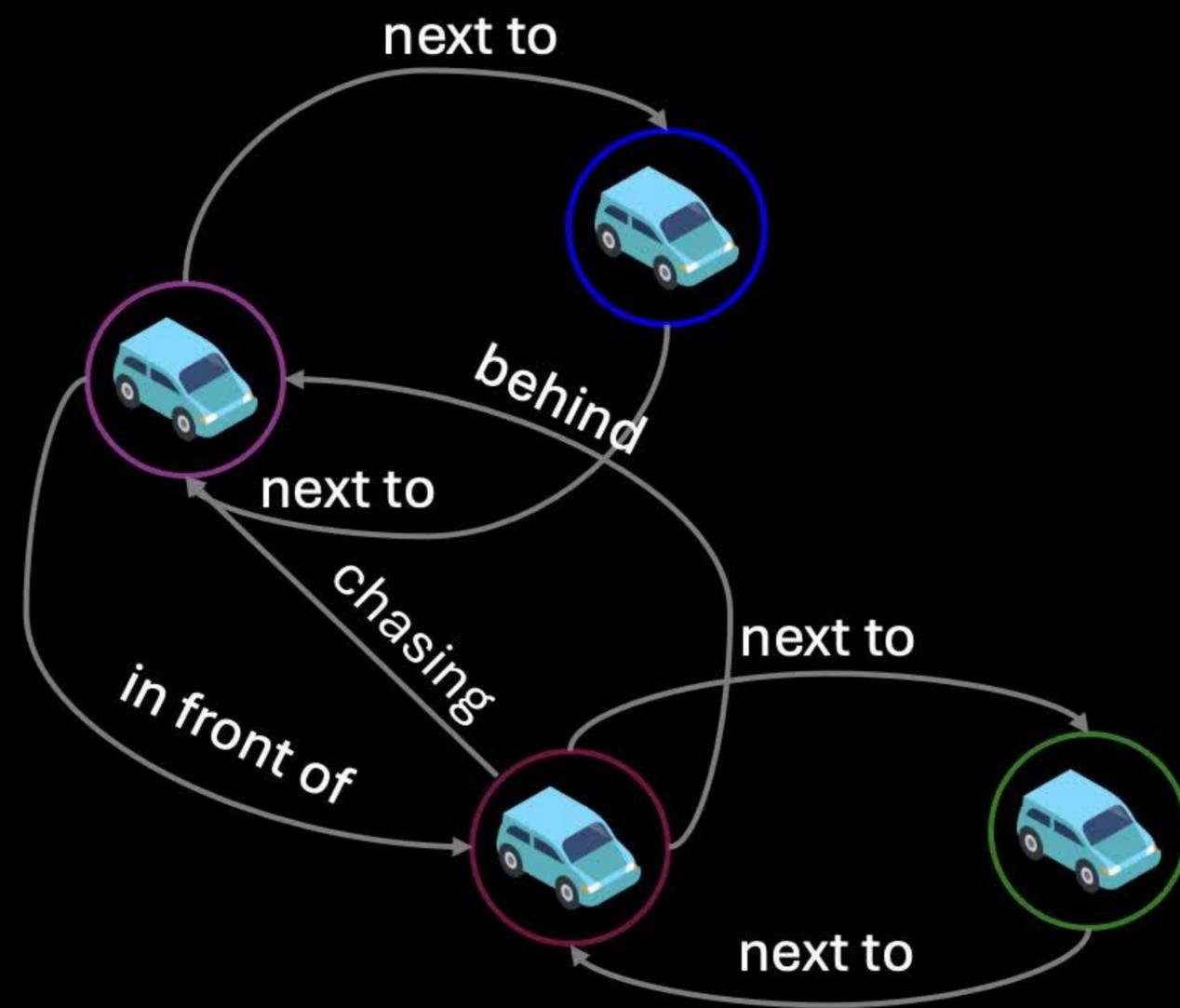
Our Approach

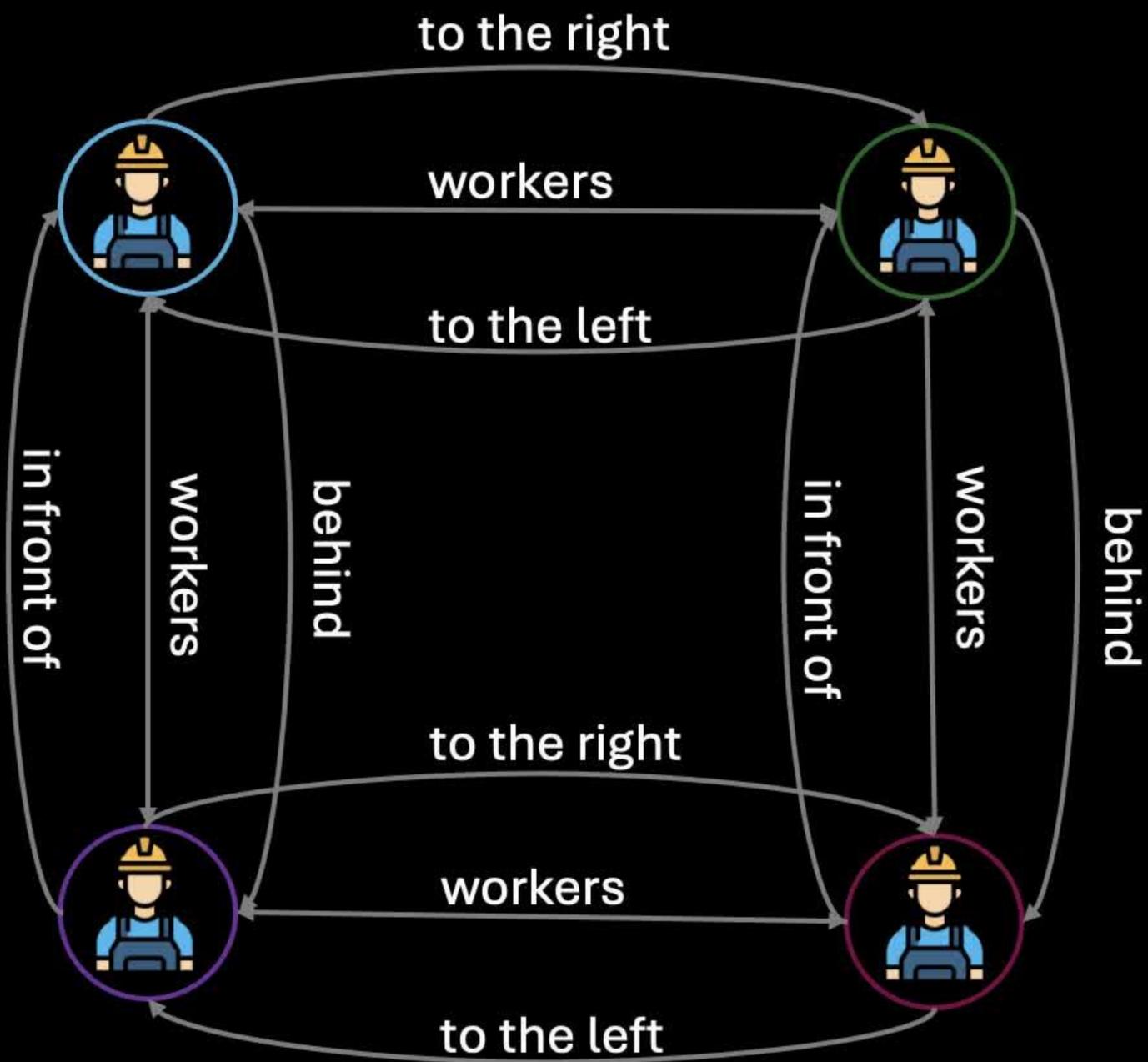


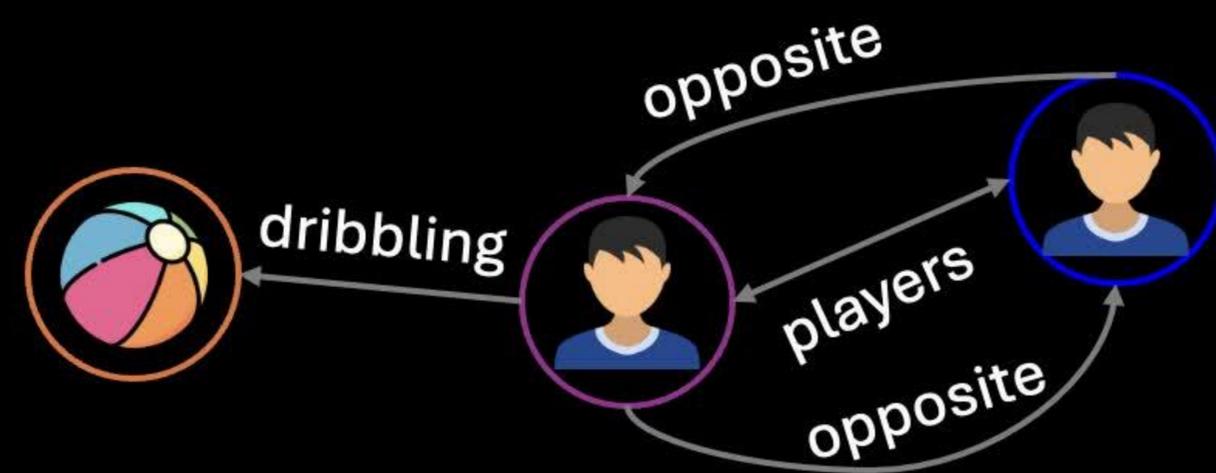
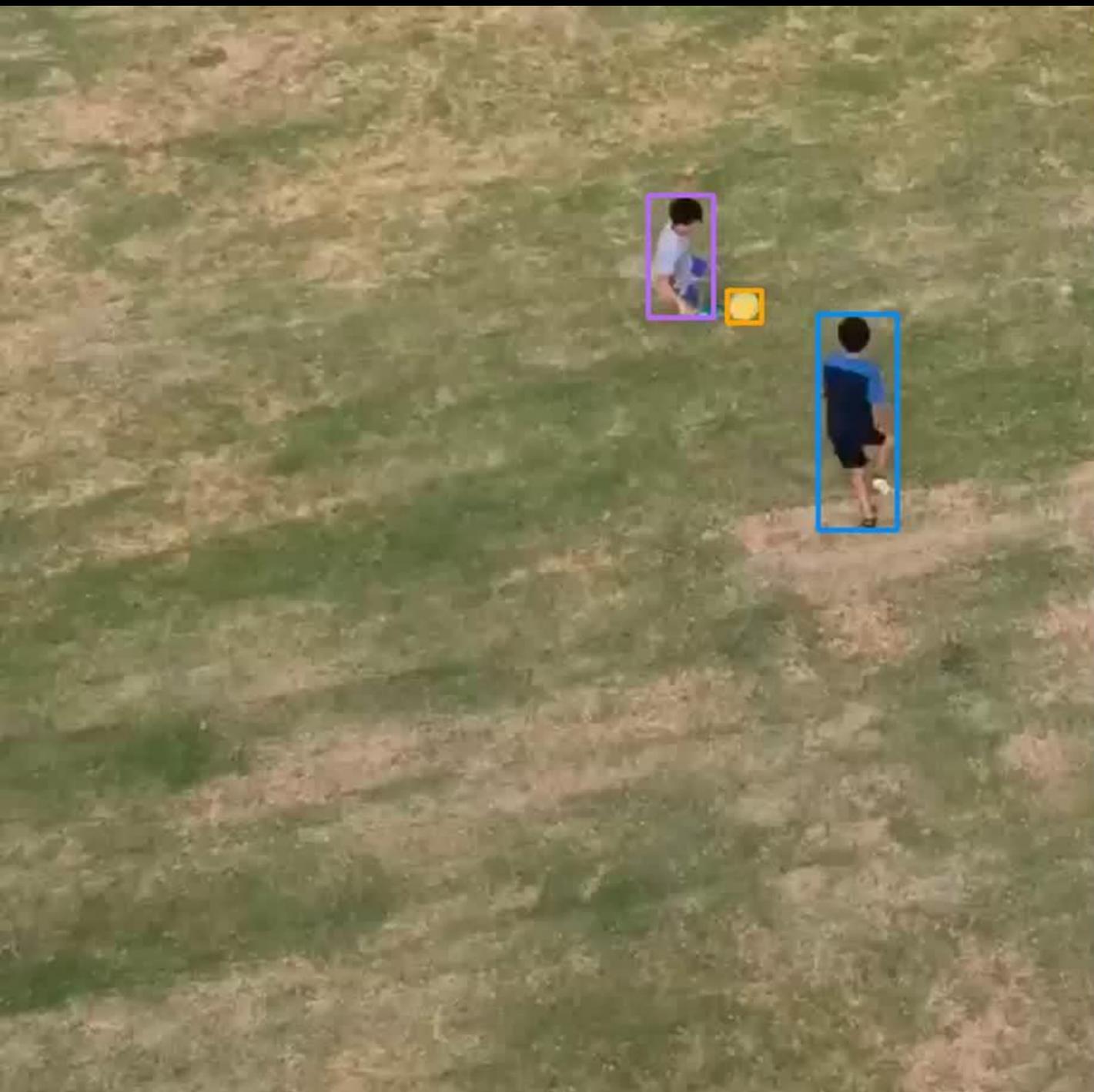
Cyclic Temporal Graph Transformer

- Cyclic Attention Mechanism









Thank you!

Contact: uark-cviu.github.io