How2comm: Communication-Efficient and Collaboration-Pragmatic Multi-Agent Perception

Dingkang Yang[1]*, Kun Yang[1]*, Yuzheng Wang[1], Jing Liu[1], Zhi Xu[1], Rongbin Yin[2], Peng Zhai[1], Lihua Zhang[1]

[1]Academy for Engineering and Technology, Fudan University

[2]FAW (Nanjing) Technology Development Company Ltd

Outline

- **Background**

- **Proposed Framework**

- **Experimental Results**

Background

Inherent weaknesses of single-agent perception

Blind Spots

Unprotected left turn   Long-range perception
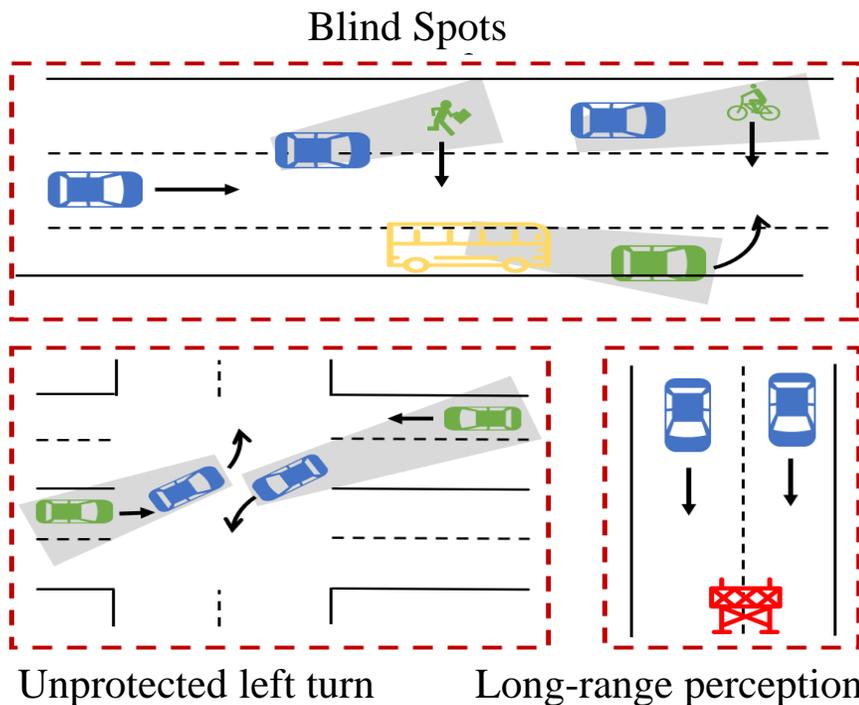
Collaborative perception scenarios
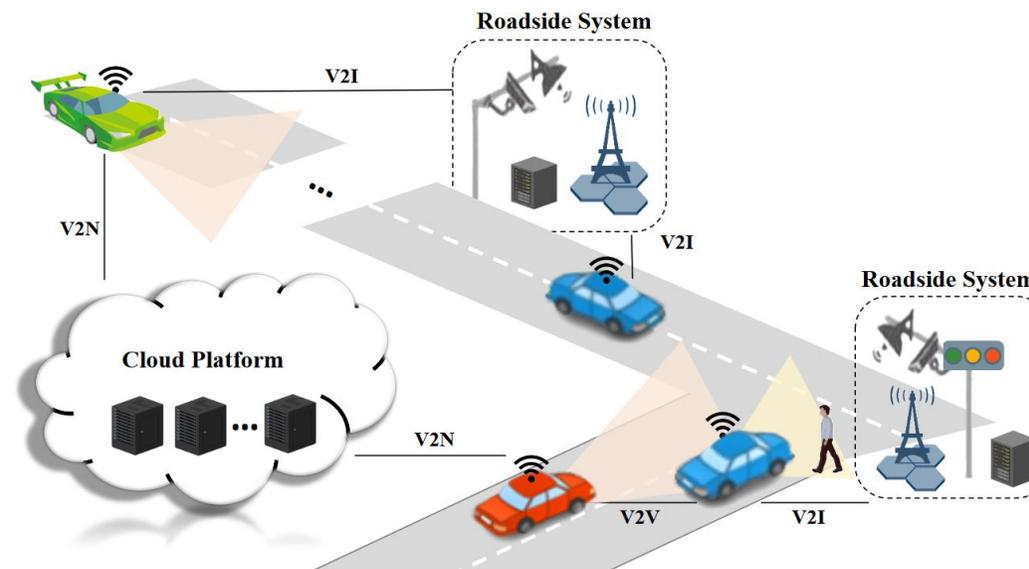
Liu, Si, et al. "Towards Vehicle-to-everything Autonomous Driving: A Survey on Collaborative Perception." *arXiv preprint arXiv:2308.16714* (2023).

Despite the advancements in previous efforts, challenges remain due to various dilemmas in the perception procedure, including **communication redundancy**, **transmission delay**, and **collaboration heterogeneity**.

Single-agent Perception v.s. Collaborative Perception

Proposed Framework
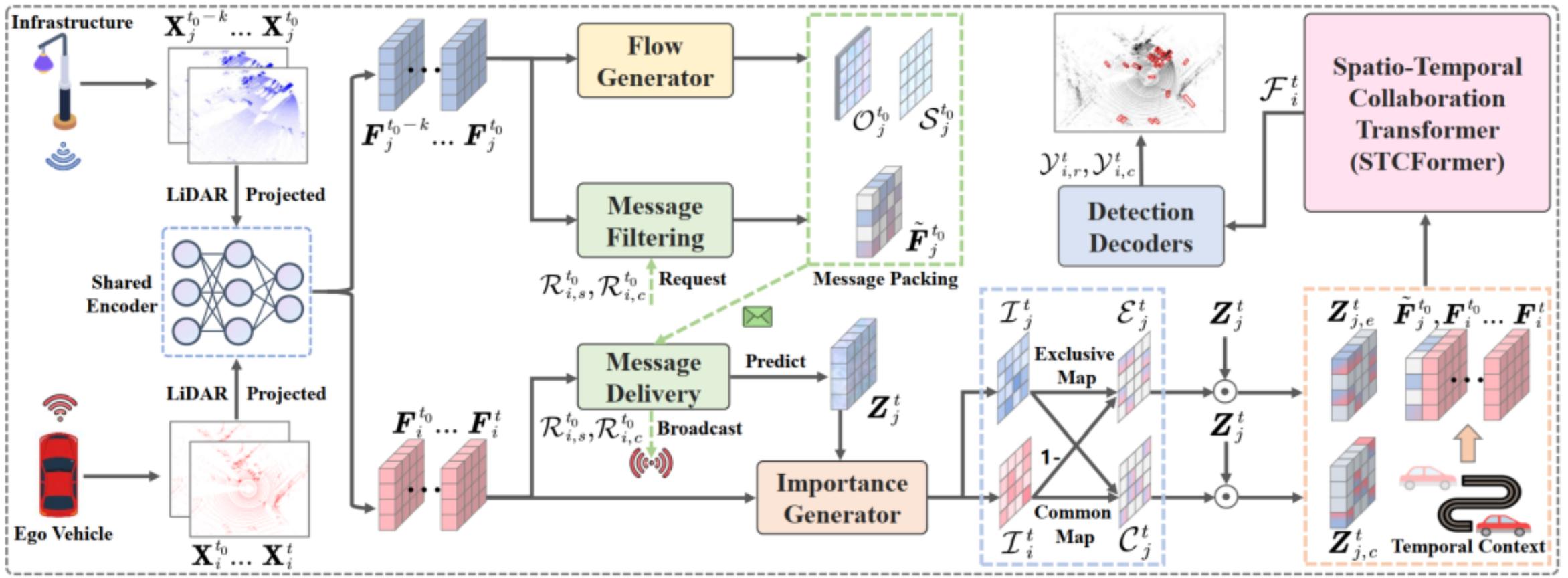
How2comm overview

Proposed Framework
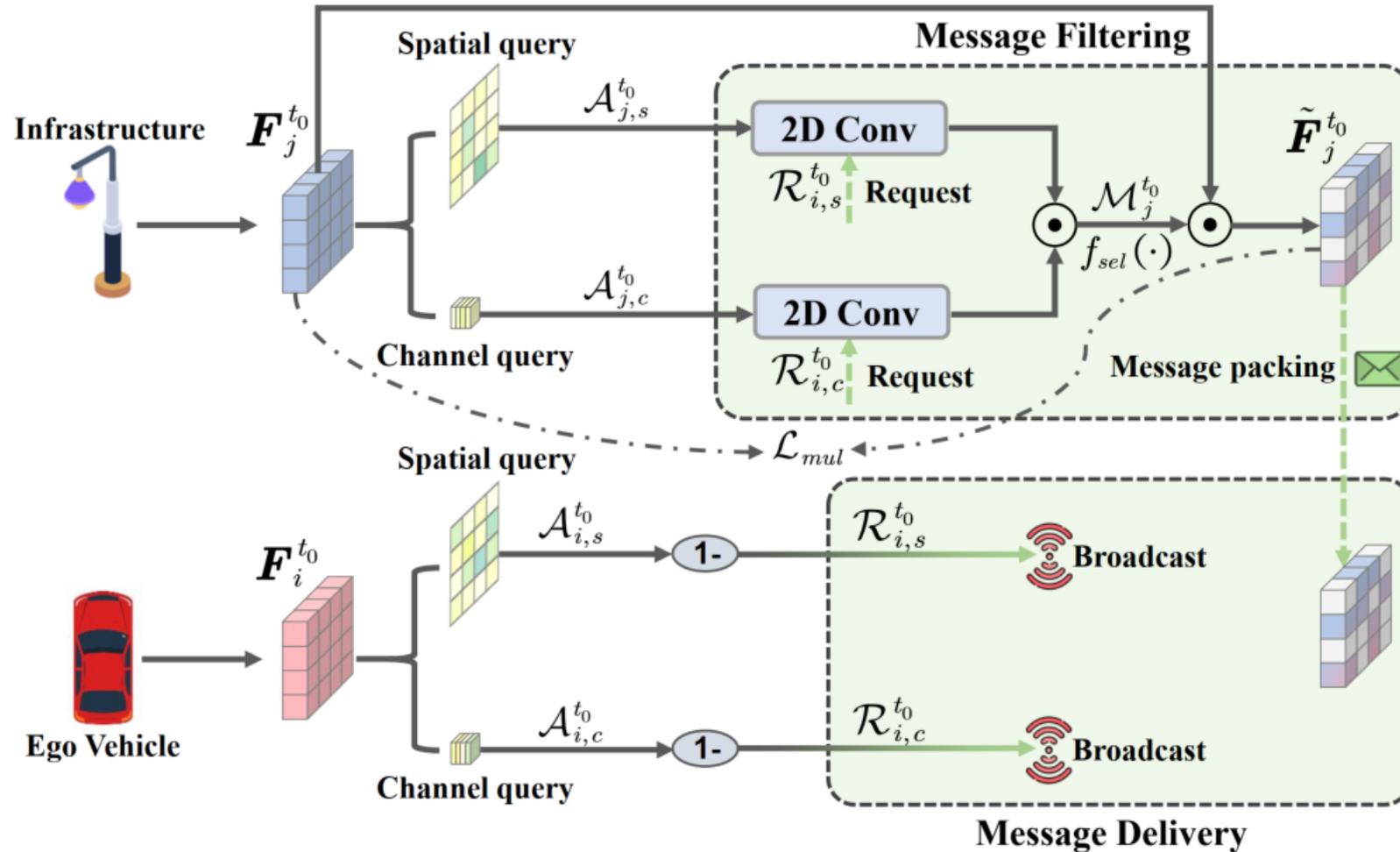
- We devise a **mutual information-aware communication mechanism** to preserve the beneficial semantics from vanilla characteristics in the transmitted messages.

Proposed Framework
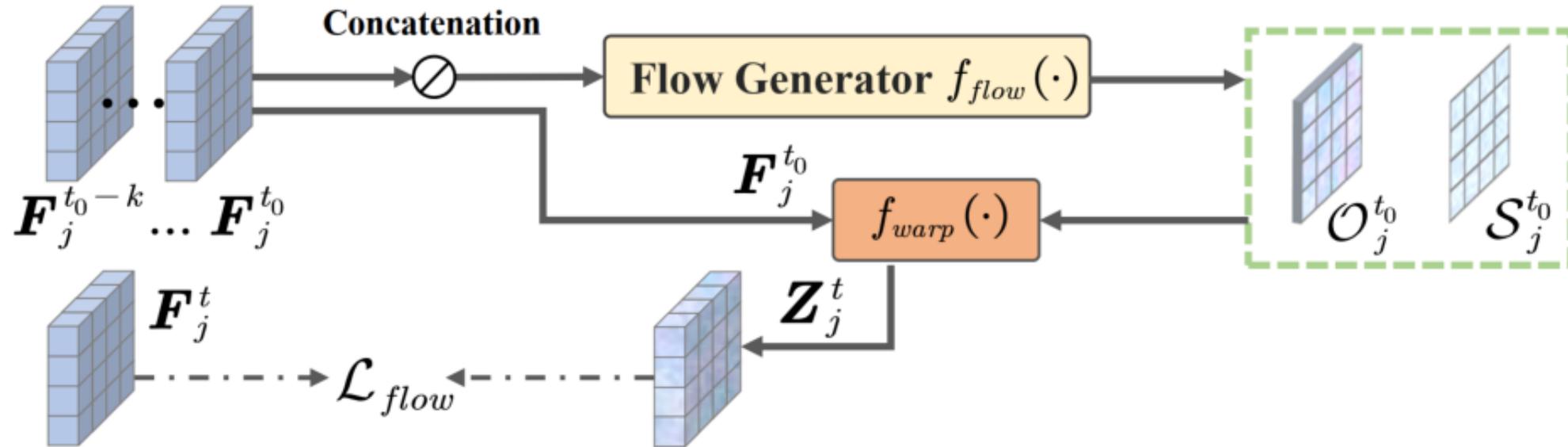
- We present a **flow-guided delay compensation strategy** to predict the future features of collaborators by mining contextual dependencies in sequential frames.

Proposed Framework
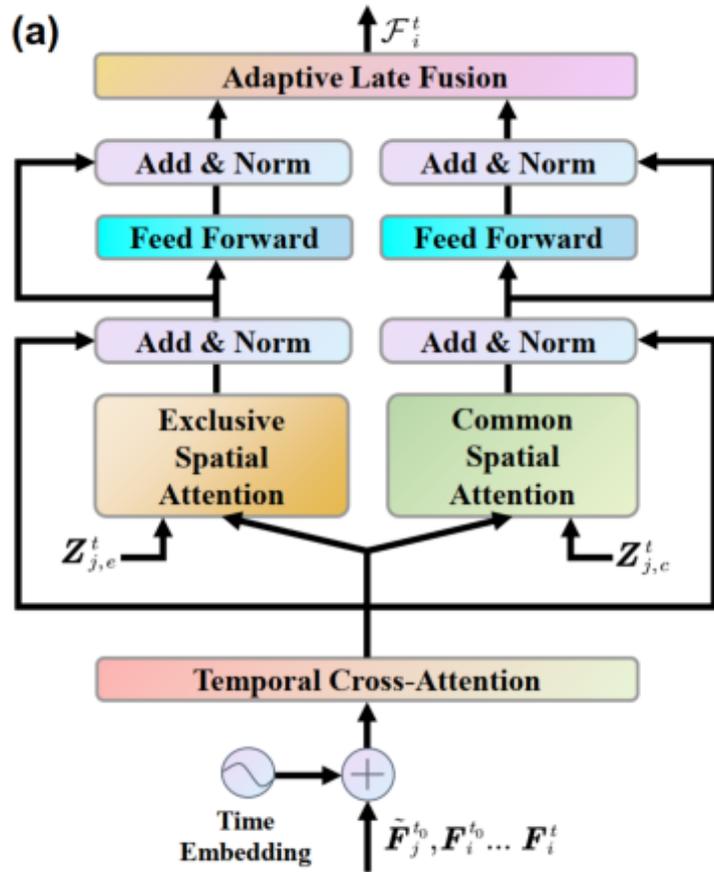
STCFormer
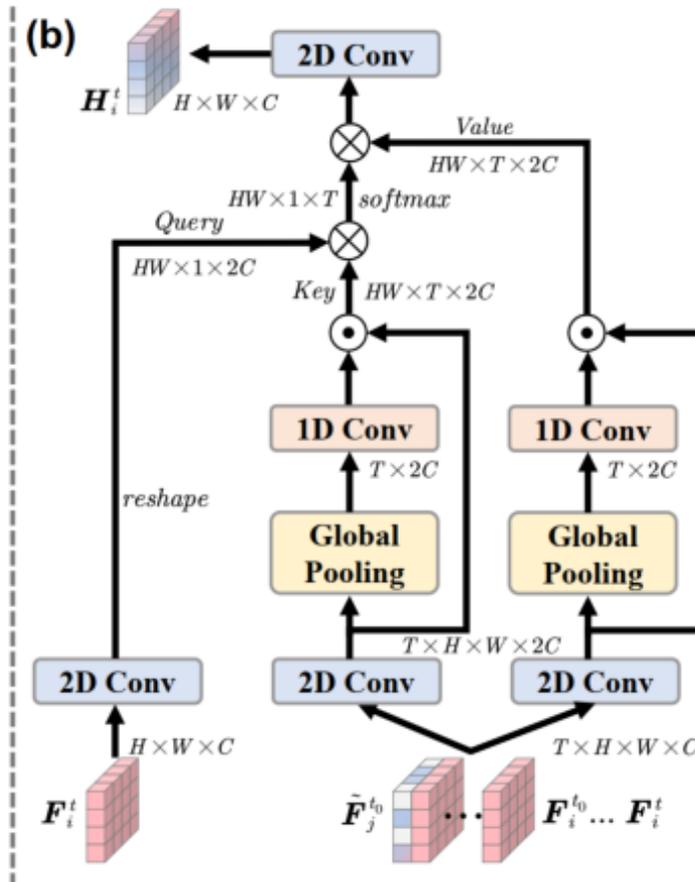
Temporal cross-attention
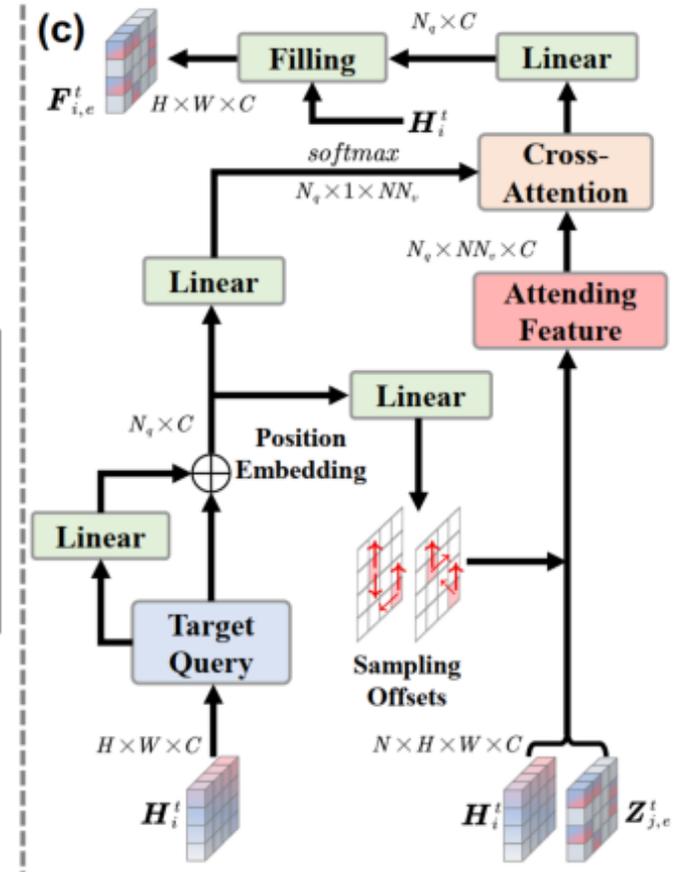
Exclusive spatial attention

- We introduce a pragmatic **spatio-temporal collaboration transformer (STCFormer)** to integrate holistic spatial semantics and temporal context clues among agents.

Experimental Results
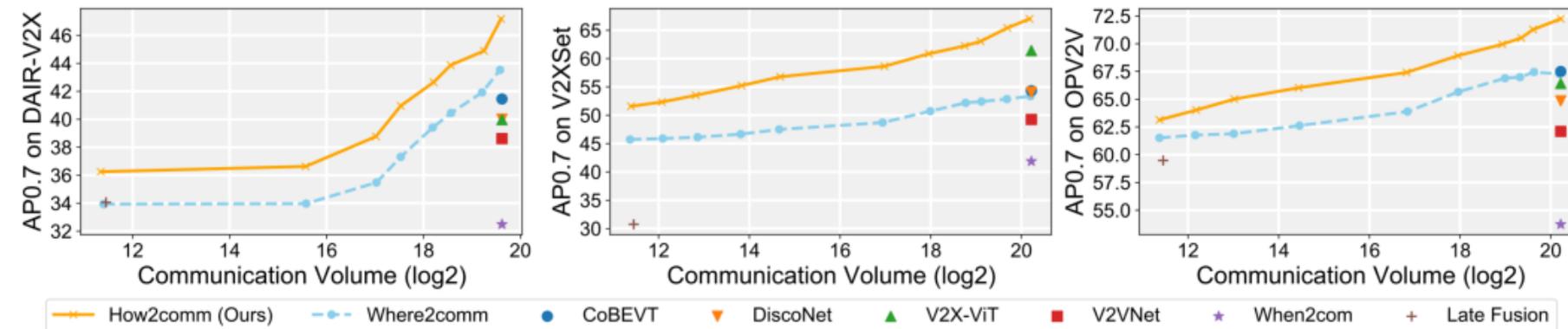
> ➤ **Collaborative perception performance comparison with varying communication volumes.**
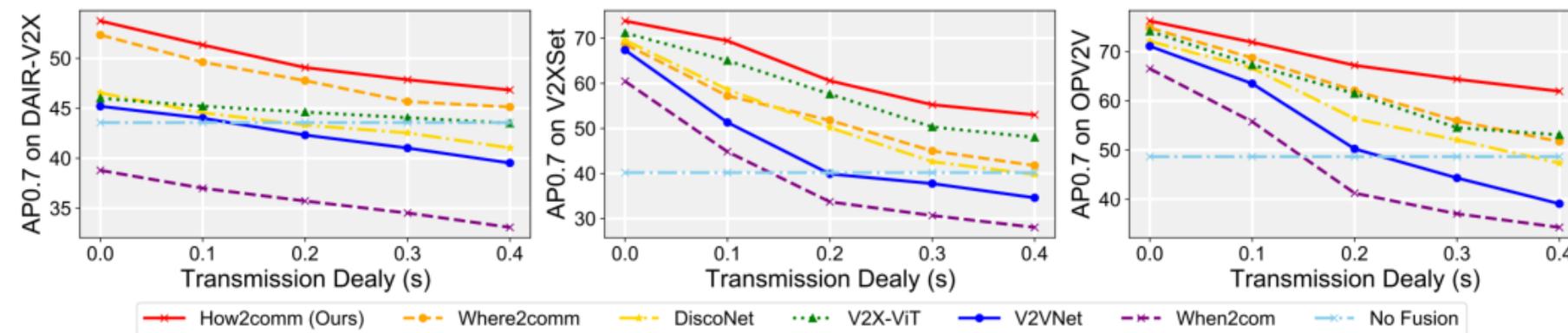
> ➤ **Robustness to transmission delay.**
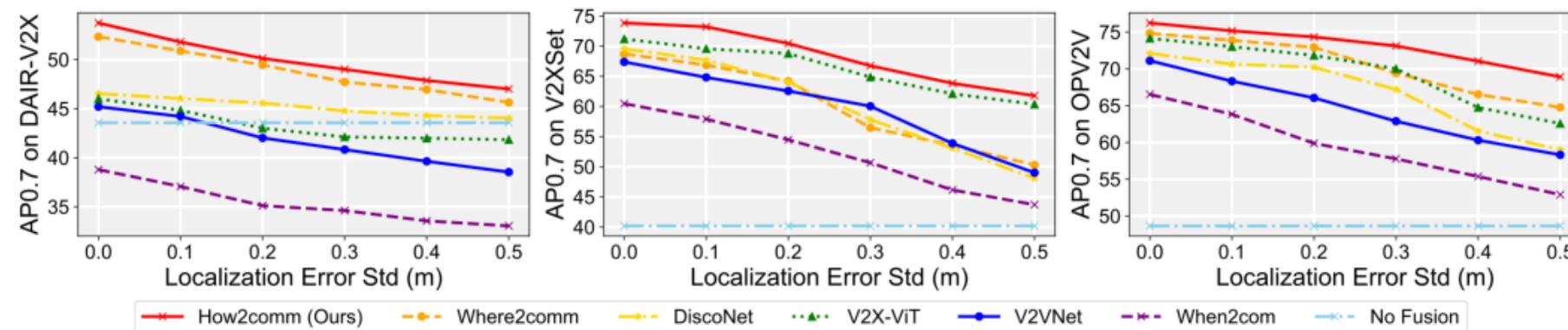
> ➤ **Robustness to localization error.**

Thanks!