IDD-X: A Multi-View Dataset for Ego-relative Important Object Localization and Explanation in Dense and Unstructured Traffic

[1]IIIT Hyderabad, [2]IIT Mandi

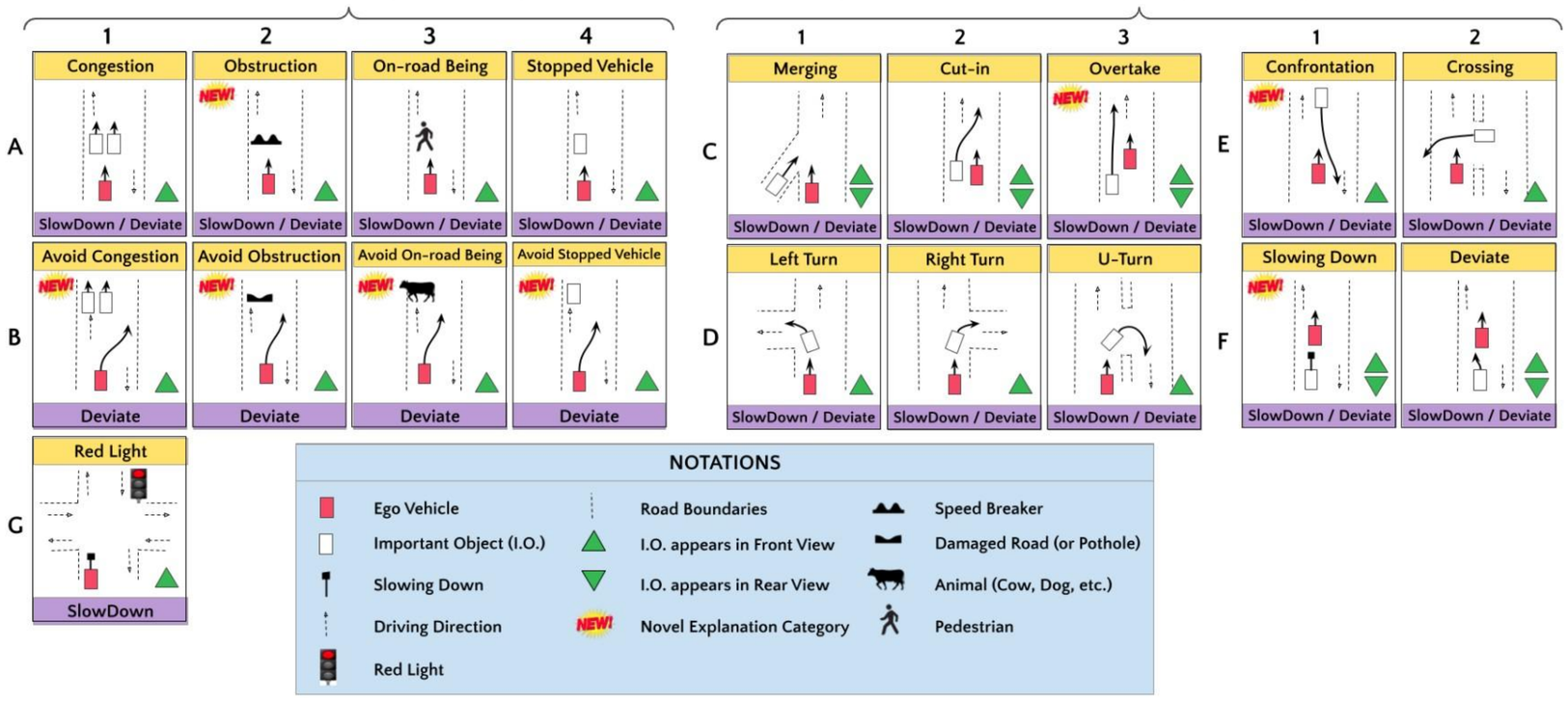
Chirag Parikh[1], Rohit Saluja[2], C.V. Jawahar[1], Ravi Kiran Sarvadevabhatla[1]

DATASET MOTIVATION & OVERVIEW

- Intelligent vehicle systems require a deep understanding of the interplay between road conditions, surrounding entities, and the ego vehicle's driving behavior for safe and efficient navigation.
- Existing datasets, predominantly geared towards structured and sparse traffic scenarios, fall short of capturing the complexity of driving in unstructured dense, and heterogeneous traffic environments in developing countries.
- To fill this gap, we present **IDD-X**:
  - A large-scale (> **85 hours**) dual-view (front and **rear**) driving video dataset
  - Captured in **Dense**, **Heterogeneous**, and **Unstructured** Traffic
  - With **697K** Important Object Bounding Boxes (**9K** Tracks) of **10** different object categories.
  - Containing **1-12** Important Objects (Per Video) with **19** Ego-relative Explanation Categories (depicted below)

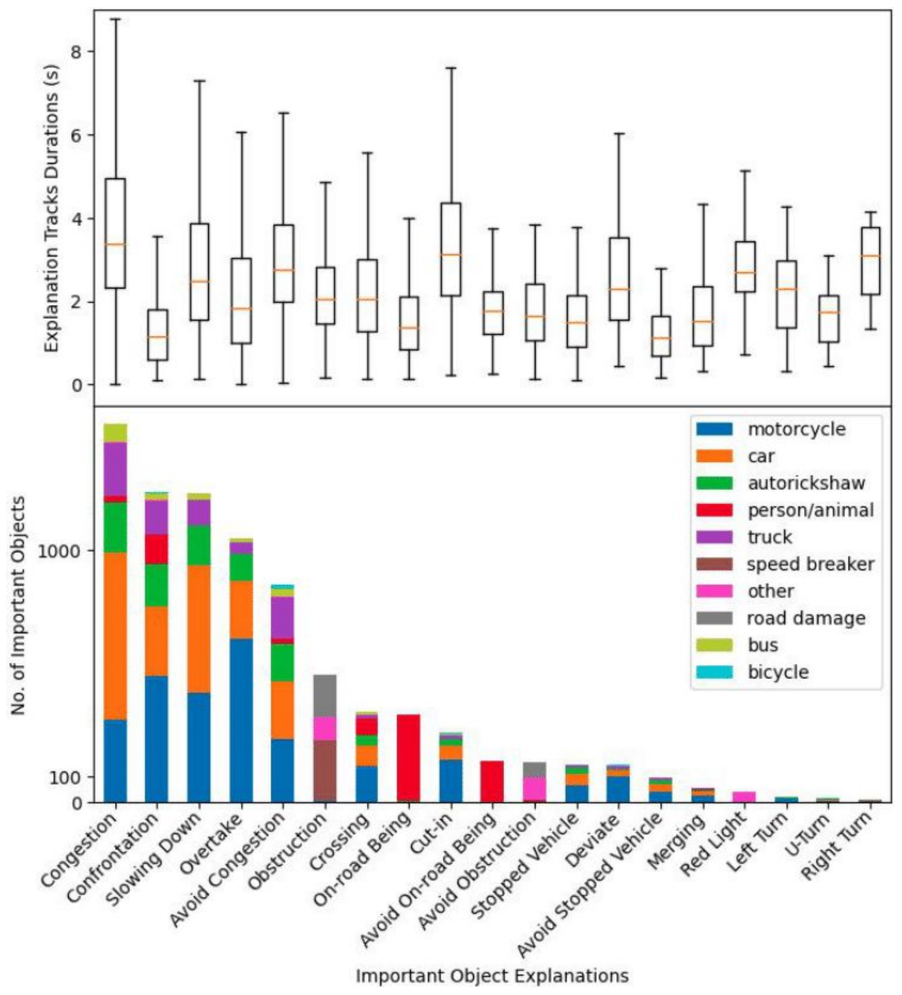
(I) **Passive influence** on ego's driving decision

(II) Ego-relative **maneuvering styles** of Important Objects



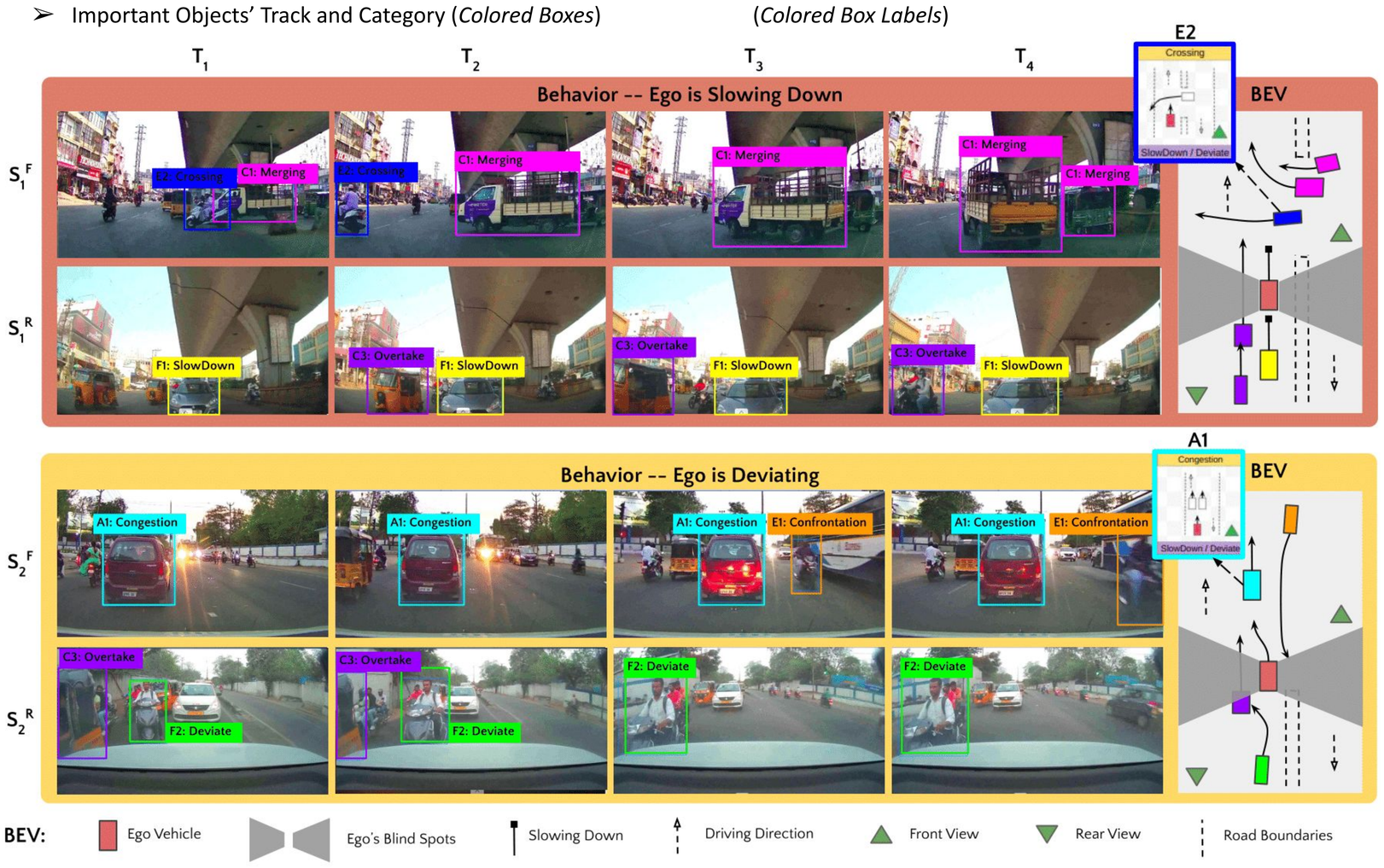
DATA STATISTICS

- Important Object Explanations' Duration and Heavy-Tail Distribution



IDD-X DATASET ANNOTATIONS

- Ego Vehicle's Driving Behavior (e.g. *Slowing Down*)
- Important Objects' Track and Category (*Colored Boxes*)
- Important Objects' Ego-relative Explanation Category (*Colored Box Labels*)



- Figure shows Front ([F]) and Rear ([R]) Views of Sample Driving Scenarios (S_1, S_2) at Timeframes (T_1-T_4) with Important Object Location and Explanation annotations in colored boxes and the Ego-vehicle's Driving Behavior label at the top.

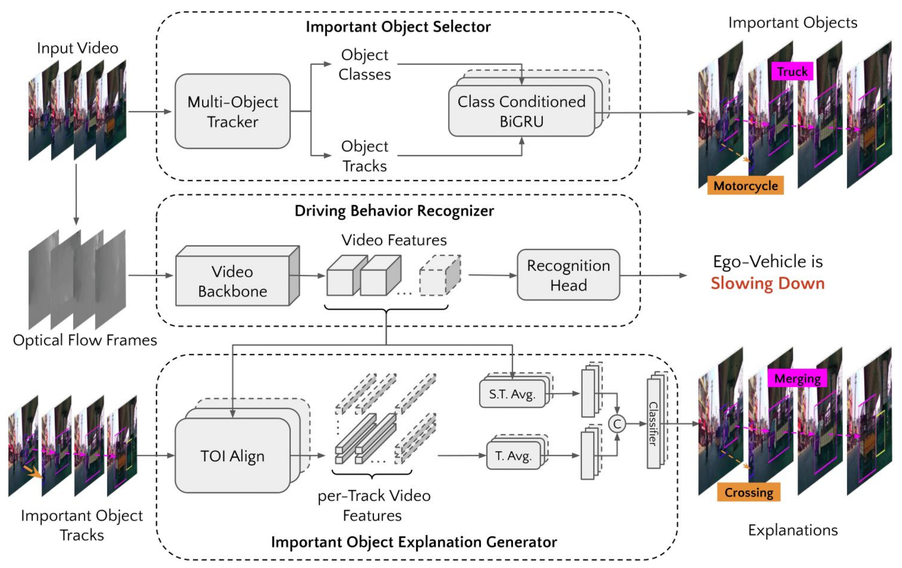
EXISTING DATASETS COMPARISONS

Dataset	Multi-View	I.O. Location	I.O. Track	Irregular Road Surface Description	Dense Traffic	Unstructured Traffic	#I.O. Bounding Boxes	#I.O.s Per Video	#I.O. Explanation Categories	#I.O. Categories	#I.O. Explanation Type
IDD-X	✔	✔	✔ (9K)	✔	✔	✔	**697K**	**1-12**	**19**	**10**	Categorical
DRAMA [3]	-	✓	-	-	-	-	17K	1	-	3	Textual
H3D [1]	-	✓	-	-	-	-	8K	1-3	-	4	-
BDD-OIA [4]	-	-	-	-	✓	-	-	-	8[1]	6	Categorical
OIE [2]	-	✓	-	-	-	-	4K	1-2	-	2	-
HDD [6]	-	-	-	-	-	-	-	1	5	4	Categorical
BDD-X [5]	-	-	-	-	-	-	-	-	-	-	Textual
METEOR [7]	-	-	-	-	✓	✓	-	-	-	-	-

* I.O. = Important Object

APPROACH

- Introduced New Methods for:
  1. **Important Object Track Identification**
  2. **Important Object Explanation Prediction**



RESULTS

1. Important Object Track Identification

Model	Conditioning	Precision	Recall	F1-Score
BiGRU	-	35.0	86.1	49.8
Class Conditioned BiGRU	At t = 0	34.8	86.5	49.6
Class Conditioned BiGRU	At t = T	34.9	85.6	49.6
Class Conditioned BiGRU	**At all t**	**35.8**	**86.8**	**50.7**

- **Object's Class** information conditioned **throughout** its trajectory improves performance.

2. Important Object Explanation Prediction

Input Features	Congestion	Confrontation	Avoid Congestion	Overtaking	Crossing	Interfering Being	Cut-in	Avoid On-road Being	Avg. F1-Score	Wt. Avg. F1-Score
TOI-Aligned	**72.2**	48.6	**53.3**	**53.9**	37.5	17.4	7.3	0.0	36.3	54.8
TOI-Aligned + Context	71.0	**53.6**	47.2	48.6	**44.7**	**22.4**	**34.7**	**25.0**	**43.4**	**55.8**

- **Video Features (Context)** along with **per-Track Features** gives best *average* and *overall* performances.

ACKNOWLEDGMENT

The project is funded by the iHubData and Mobility at IIIT Hyderabad. We thank the data collection and annotation team for their effort.

ICRA2024 YOKOHAMA | JAPAN

Scan for Code, Dataset download and related resources on the Project Page.

idd-x.github.io

INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY HYDERABAD

HUBDATA IIIT HYDERABAD

CVIT