Vehicle Re-Identification with the Space-Time Prior

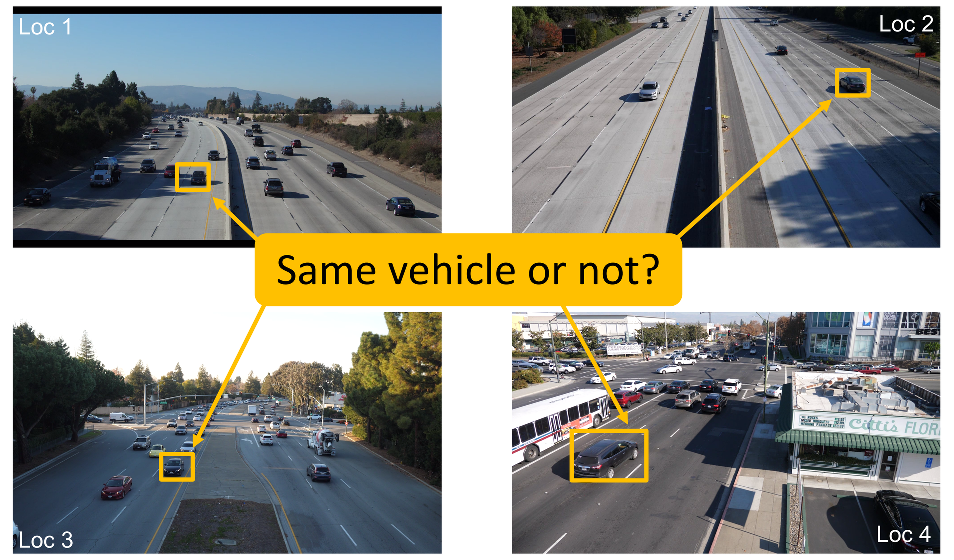
Chih-Wei Wu, Chih-Ting Liu, Chen-En Chiang, Wei-Chih Tu, Shao-Yi Chien
NTU IoX Center, National Taiwan University

National Taiwan University

NVIDIA AI CITY CHALLENGE

CVPR 2018 SALT LAKE CITY • JUNE 18-22

Track 3 Problem Statement



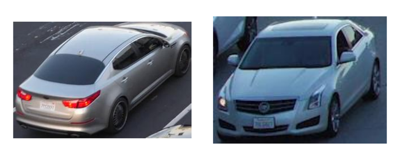
➢ **Vehicle Re-Identification:**
Find all vehicles that travel through all 4 locations within surveillance videos

Introduction

➢ **Challenges:**

(1) Lacks labeled dataset
→Only one vehicle Re-ID dataset (VeRi)

(2) Domain change


Silicon Valley


China

(3) Intra-Identity variance

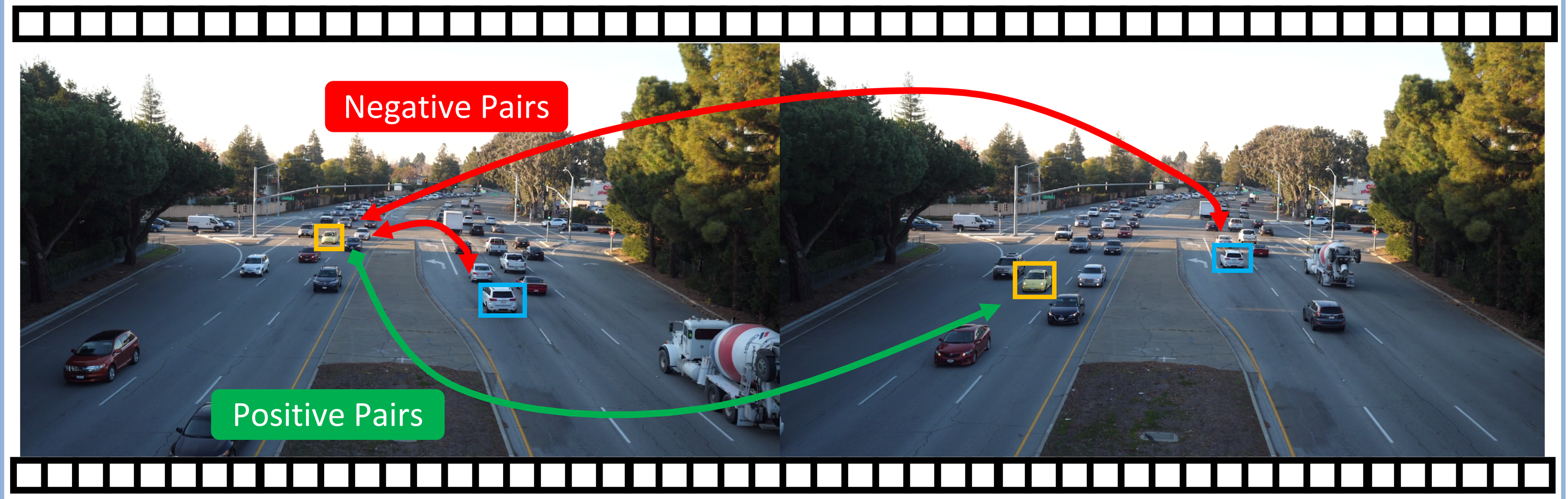

view change


illumination change

Main Contribution

- ➢ Propose **A**daptive **F**eature **L**earning (AFL) with **space-time prior**, for adapting **trained CNN** feature extractor to **target** domain
- ➢ Train a CNN feature extractor in a **multi-task** learning manner and utilize our **AFL** technique
- ➢ Develop an vehicle **detection**, **tracking** and **Re-ID** system

Adaptive Feature Learning (AFL)

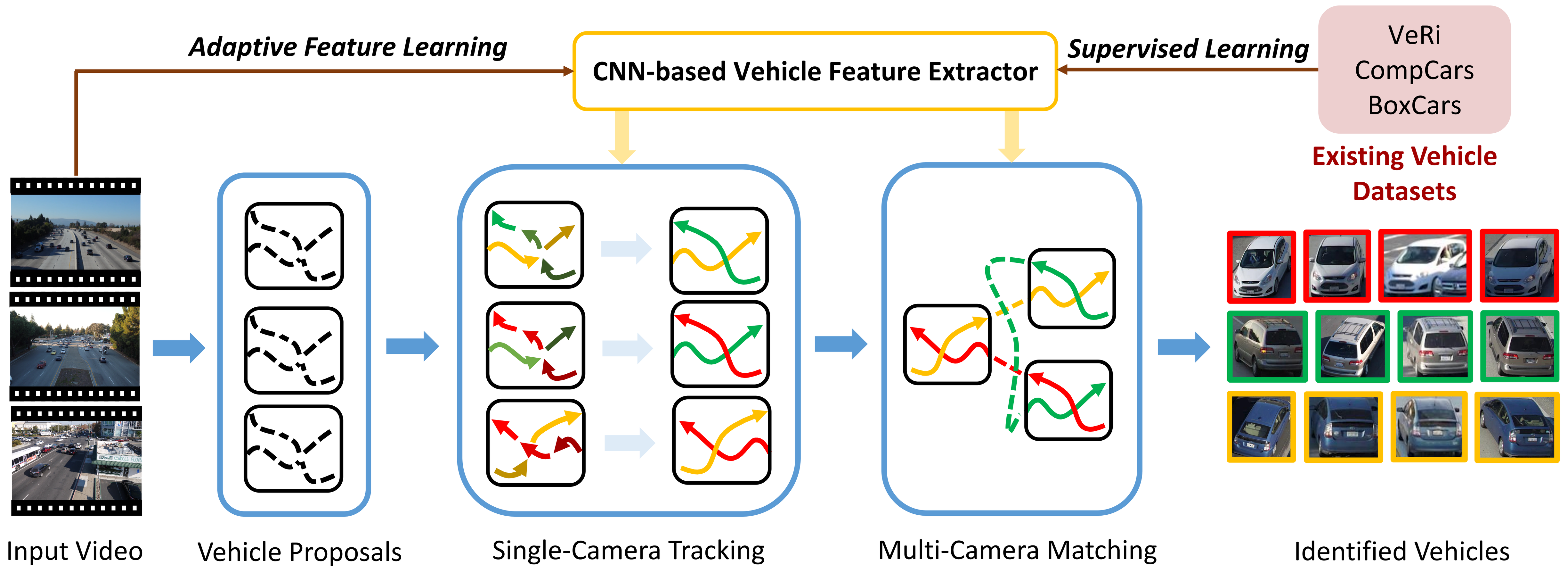


Frame t　　Frame t+1

➢ **The space-time prior within the videos:**
(1) One vehicle can not appear at multiple locations at the same time (Negative pairs)
(2) One vehicle should move continuously along the time (Positive pairs)

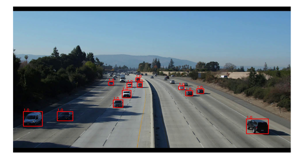
➢ **Adaptive Feature Learning (AFL):** Collect the **positive** and **negative** pairs automatically in the target dataset to adapt our trained CNN vehicle feature extractor

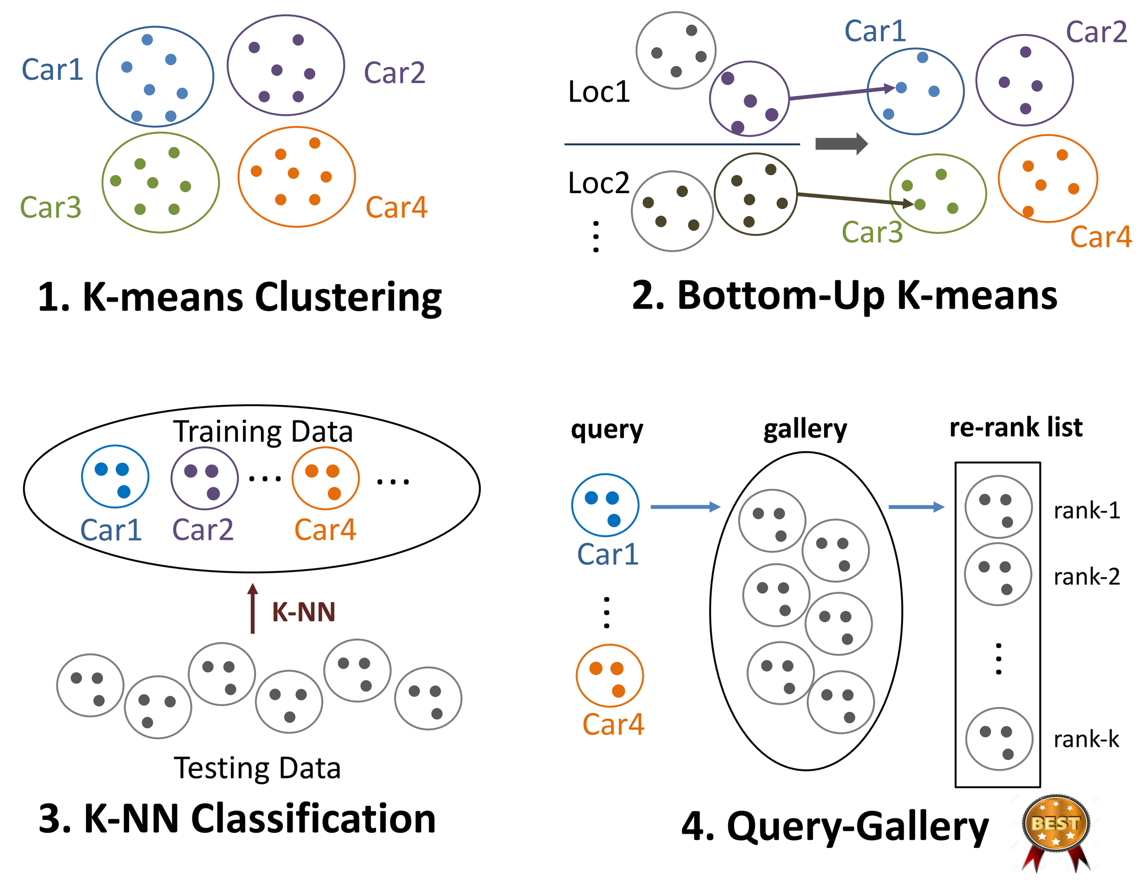
Vehicle Re-ID system



System Pipeline

- ➢ **Stage 1: Vehicle Proposals:**
  - ➢ Detectron [1] system
- ➢ **Stage 2: Single Camera Tracking:**
  - ➢ IOU tracker [2]
  - ➢ CNN Feature Matching
- ➢ **Stage 3: Multi-Camera Matching (track-based):**
  - ➢ Extract CNN features for each track
  - ➢ We try **4** methods, **Query-Gallery** performs best

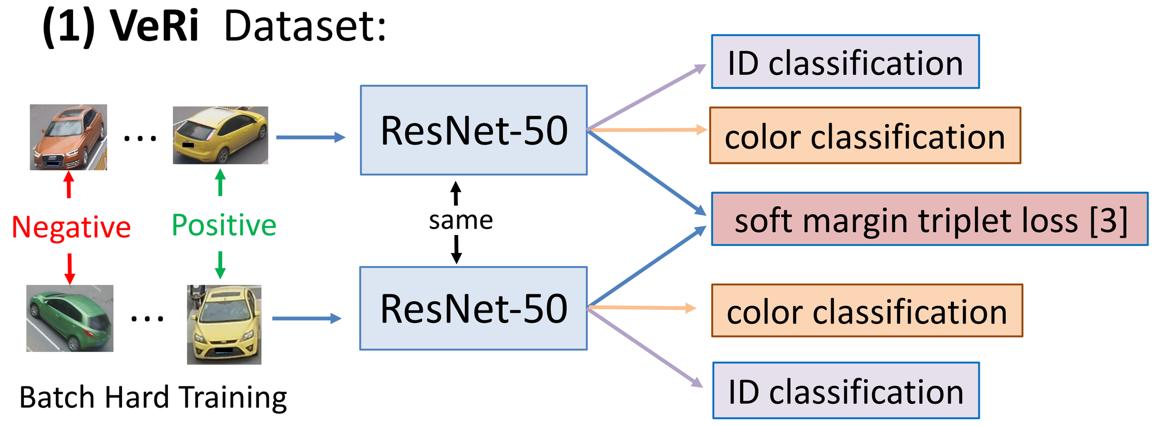




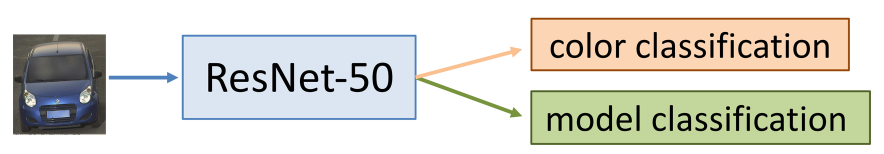
Multi-Task Vehicle Feature Extractor

- ➢ Use ResNet-50 as backbone feature extractor
- ➢ Use multiple tasks and datasets to jointly train

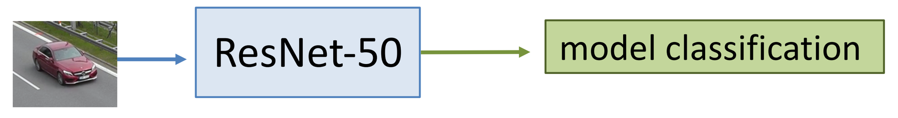
(1) VeRi Dataset:



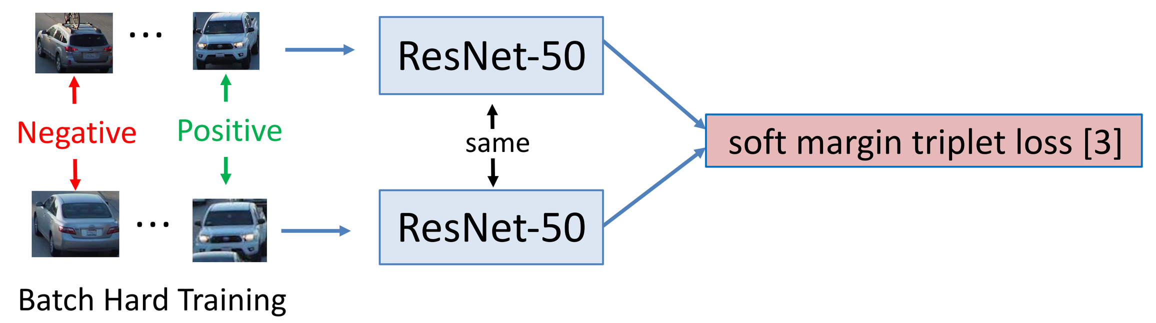
(2) CompCars Dataset:



(3) BoxCars Dataset:



(4) AIC Test Dataset: use **AFL** to adapt our CNN model



Experiment Results

(a) Effectiveness of AFL Technique

Train on **Market-1501** Human Re-ID
Test on **DukeMTMC-ReID** Humam Re-ID

Method	mAP(%)↑	Rank-1(%)↑
w/o AFL	13.46	25.99
w/ AFL	**14.20**	**28.50**

(b) Effectiveness of Multi-Task Training.

Test on **VeRi** vehicle Re-ID dataset
(* uses additional temporal information)

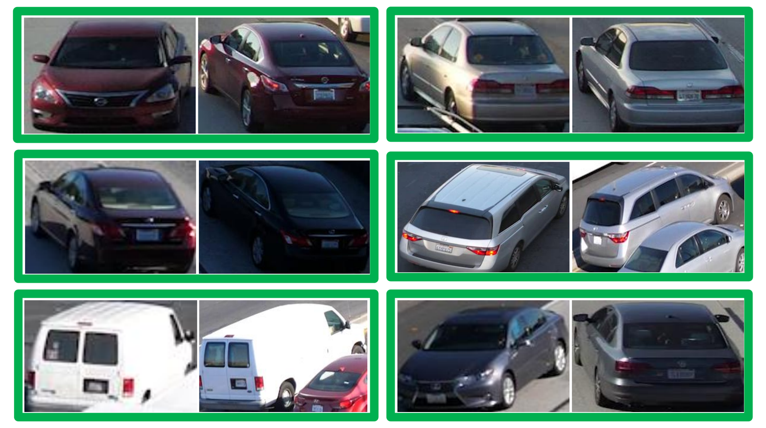
Method	mAP(%)↑	Rank-1(%)↑
SOTA CNN [4]	29.48	41.12
*SOTA [4]	**58.27**	83.49
Train on VeRi (**Ours**)	53.35	82.06
Train on all (**Ours**)	57.43	**86.29**

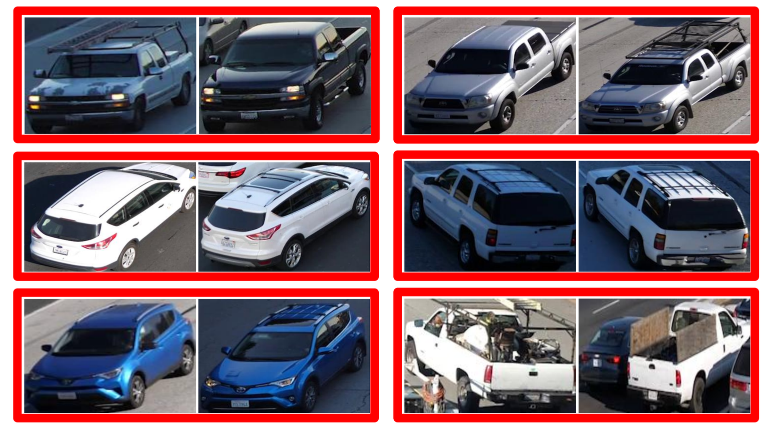
(c) Multi-Camera Matching Results

Test on **AIC** Track 3 dataset
TDR: Track Detection Rate
PR: Localization Precision
S3: 0.5*(TDR+PR)

Method	TDR↑	PR↑	S3↑
K-Means	0	0.0006	0.0003
Bottom-Up K-Means	0	0.0015	0.0007
K-NN	0.1429	**0.0020**	0.0725
Query-Gallery	**0.5714**	0.0007	**0.2861**

➢ Matching Results


Possibly Correct


Definitely Incorrect

(d) LeaderBoard of AIC 2018 Track3.

Team ID	S3↑
team48	0.7106
team37 (Ours!)	**0.2861**
team79	0.0785
team18	0.0074
team28	0.0026
team41	0.0024
team53	0.0002
team6	0.0001

[1] R. Girshick et al. "Detectron". https://github.com/facebookresearch/detectron, 2018. [2] E. Bochinski et al. "High-speed tracking-by-detection without using image information." AVSS, 2017. [3] A. Hermans et al. "In defense of the triplet loss for person re-identification." *arXiv*, 2017. [4] Y. Shen et al. "Learning deep neural networks for vehicle re-id with visual-spatio-temporal path proposals." CVPR, 2017.