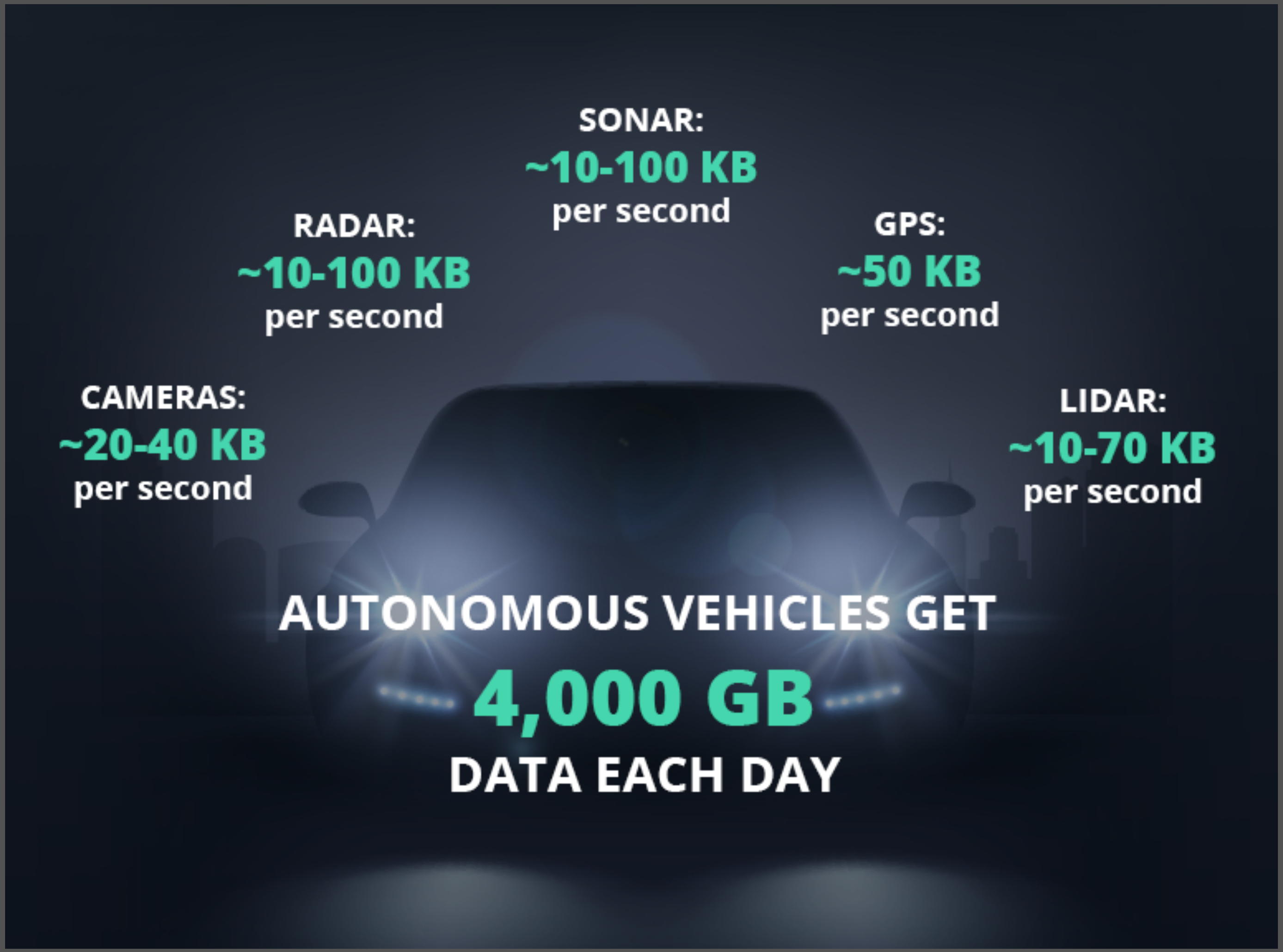
2021 ICCV OCTOBER 11-17 VIRTUAL

TMCOSS: Thresholded Multi-Criteria Online Subset Selection for Data-Efficient Autonomous Driving

Soumi Das[1], Harikrishna Patibandla[1], Suparna Bhattacharya[2], Kshounis Bera[3], Niloy Ganguly[1], Sourangshu Bhattacharya[1]

[1]Indian Institute of Technology, Kharagpur,
[2]Hewlett Packard Labs, Hewlett Packard Enterprise
[3]Hewlett Packard Enterprise

Autonomous Driving



Collection of data using various types of sensors like camera, lidar, radar is leading to a high rise in driving data.

Aim

- We study the problem of data efficient training of autonomous driving systems.

- Training using many frames on straight road sections may not be necessary. Frames at the turns turn out to be useful.

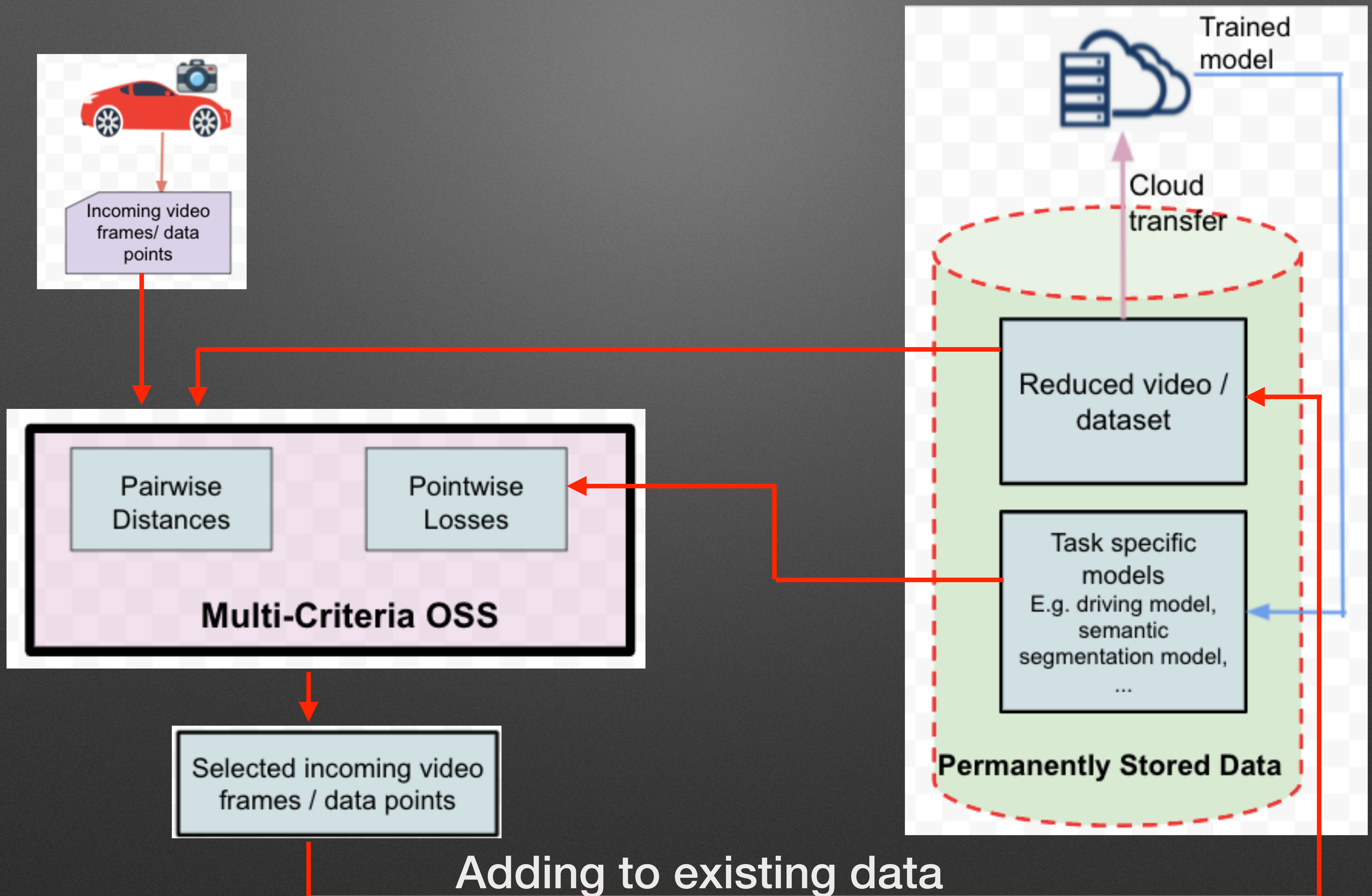
REDUNDANT

INFORMATIVE

Method	Train One-Turn	Test One-Turn
Uniform Skip	3/10	5/10

In the context of edge device deployment, multi-criteria online subset selection (OSS) framework can be useful in selecting informative frames, essential for an end-task.

Subset selection on Edge devices



High Level Idea

- Given a compression ratio, find out representatives which have the least dissimilarity with the left-out elements besides having the highest task-specific loss.

- Least dissimilarity ensures having a diverse set of elements.

- Highest task-specific loss ensures having situational tasks needed to be learnt more by the model.

TMCOSS

Adopts a facility location objective involving multiple criteria

$$\min_{z^o_{ij}, z^n_{ij}} \mathcal{G}(z^o_{ij}, z^n_{ij}) s.t. \sum_{j=1}^{|R_t|} z^o_{i,j} + \sum_{j=1}^{m} z^n_{i,j} = 1; z^n_{i,j}, z^o_{i,j} \in [0,1]; \sum_{j=1}^{m} \|[z^n_{1,j} \dots z^n_{m,j}]\|_p \le frac * m$$

Objective function

Constraint 1

Constraint 2

Compression Ratio

$z^o_{ij} = 1$ Denotes j from existing set o is a representative of element i from incoming set n

$z^n_{ij} = 1$ Denotes j from incoming set n is a representative of element i from incoming set n

$$\mathcal{G}(z^o_{ij}, z^n_{ij}) = \rho\Big(\sum_{i=1}^{m}\sum_{j=1}^{|R_t|} z^o_{ij} d^o_{ij}(t) + \sum_{i,j=1}^{m} z^n_{ij} d^n_{ij}(t)\Big) - (1-\rho)\Big(\sum_{j=1}^{|R_t|} S^o_j * L^o_j + \sum_{j=1}^{m} S^n_j * L^n_j\Big), where, S^o_j = \frac{1}{\epsilon}\min(\epsilon, \sum_{i=1}^{m} z^o_{ij}), S^n_j = \frac{1}{\epsilon}\min(\epsilon, \sum_{i=1}^{m} z^n_{ij})$$

Dissimilarity

Task specific Loss

Representative power of element j thresholded by ϵ

Justification for thresholding

Theorem 1 *Let z^o_{ij} and z^n_{ij} be the optimal solution for formulation 1. A new frame $j \in X_{t+1}$ is selected as a representative frame for at least one incoming frame $i \in X_{t+1}$, i.e. $z^n_{ij} = 1$, only if BOTH these conditions hold:*

- *For some incoming frame $i \in X_{t+1}$, $Q^n_{ij} < Q^n_{ij'}$, for all $j' \in X_{t+1}$ and $j' \neq j$*
- *For some incoming frame $i \in X_{t+1}$, $Q^n_{ij} < \frac{\sum_{i'=1}^{m} z^o_{i',k} Q^o_{i'k} + \lambda \|[z^n_{1,j} \dots z^n_{m,j}]\|_p}{\|\mathbf{z}^n_j\|_1}$*

where $k = argmin_j \sum_{i=1}^{m} z^o_{i,j} Q^o_{i,j}$, and $\|\mathbf{z}^n_j\|_1 = \sum_{i'=1}^{m} z^n_{i'j}$

$\rho = 0$

Corollary 1.1 *Let z^o_{ij} and z^n_{ij} be the optimal solution for formulation 1. A new frame $j \in X_{t+1}$ is selected as a representative frame for at least one incoming frame $i \in X_{t+1}$, i.e. $z^n_{ij} = 1$, only if BOTH these conditions hold:*

- *$L^n_j > L^n_{j'}$ for all $j' \in X_{t+1}$ and $j' \neq j$*
- *$L^n_j > \frac{\sum_{i=1}^{m} z^o_{i,k} L^o_k - \lambda \|[z^n_{1,j} \dots z^n_{m,j}]\|_p}{\|\mathbf{z}^n_j\|_1}$*

where $k = argmin_j \sum_{i=1}^{m} z^o_{i,j} Q^o_{i,j}$, and $\|\mathbf{z}^n_j\|_1 = \sum_{i'=1}^{m} z^n_{i'j}$

Multi-criteria OSS (MCOSS)[1]

$$Q^n_{ij} = \rho d^n_{ij} - (1-\rho) L^n_j ; Q^o_{ij} = \rho d^o_{ij} - (1-\rho) L^o_j$$

$$\min_{z^o_{ij}, z^n_{ij}} \sum_{i=1}^{m} \sum_{j=1}^{|R_t|} z^o_{ij} Q^o_{ij} + \sum_{i,j=1}^{m} z^n_{ij} Q^n_{ij} + \lambda \sum_{j=1}^{m} \|[z^n_{1,j} \dots z^n_{m,j}]\|_p$$

$$s.t. \sum_{j=1}^{|R_t|} z^o_{i,j} + \sum_{j=1}^{m} z^n_{i,j} = 1, \ \forall i \in X_{t+1} z^n_{i,j}, z^o_{i,j} \in [0,1], \ \forall i,j$$

1. Soumi Das, Sayan Mondal, Ashwin Bhoyar, Madhumita Bharde, Niloy Ganguly, Suparna Bhattacharya, Sourangshu Bhattacharya, "Multi-criteria onlineframe-subset selection for autonomous vehicle videos." *Pattern Recognition Letters* 133 (2020): 349-355.

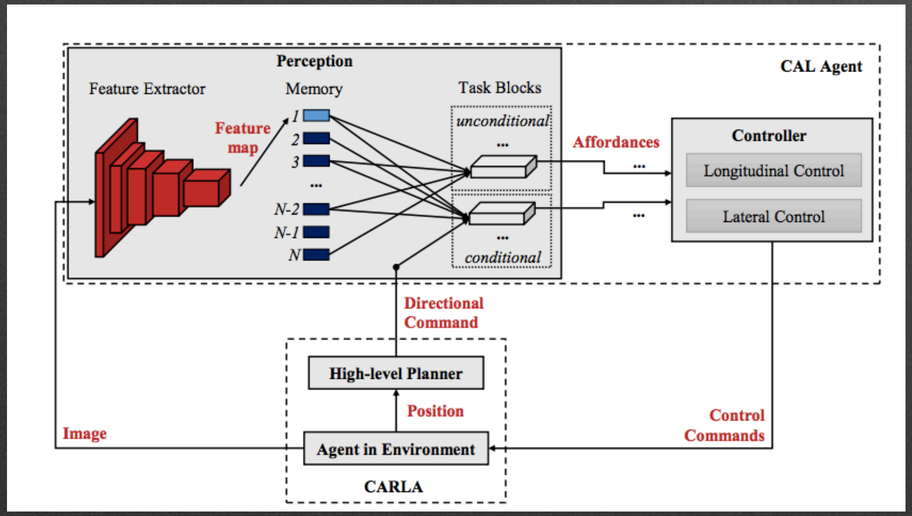
Experiments

Datasets and Models:

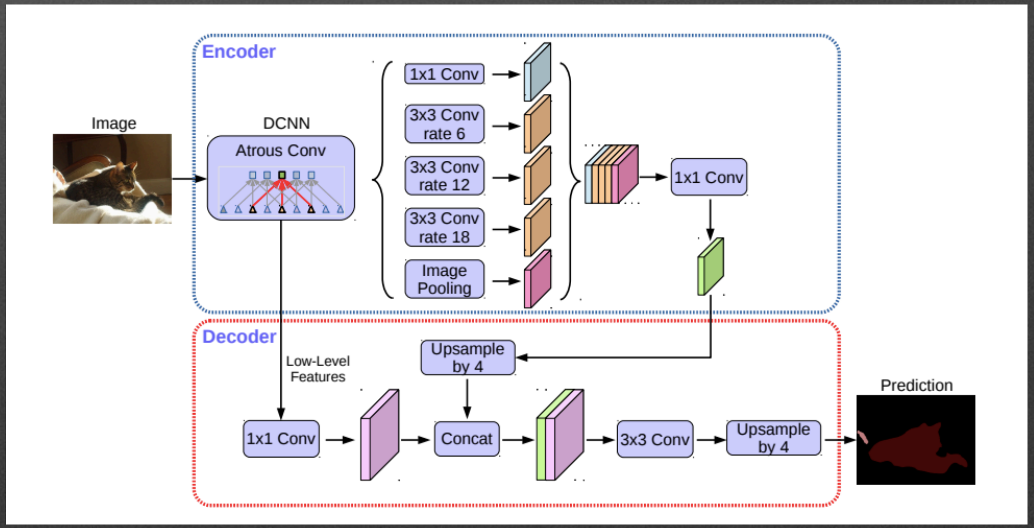
CARLA

Cityscapes

BDD Drivable Area



Conditional Affordance Learning [1](CAL)



[2]DeepLabV3+

1.Sauer, Axel, Nikolay Savinov, and Andreas Geiger. "Conditional affordance learning for driving in urban environments." *Conference on Robot Learning*. PMLR, 2018.
2.Chen, Liang-Chieh, et al. "Encoder-decoder with atrous separable convolution for semantic image segmentation." *ECCV*. 2018.

Episode Completion

- Relative angle is essential for steer values especially during turns.
- We introduce variants of TMCOSS based on point wise criteria:

TMCOSS - TL (Total Loss) and TMCOSS-BL (Bucket specific Relative Angle Loss)

100:20 compression ratio

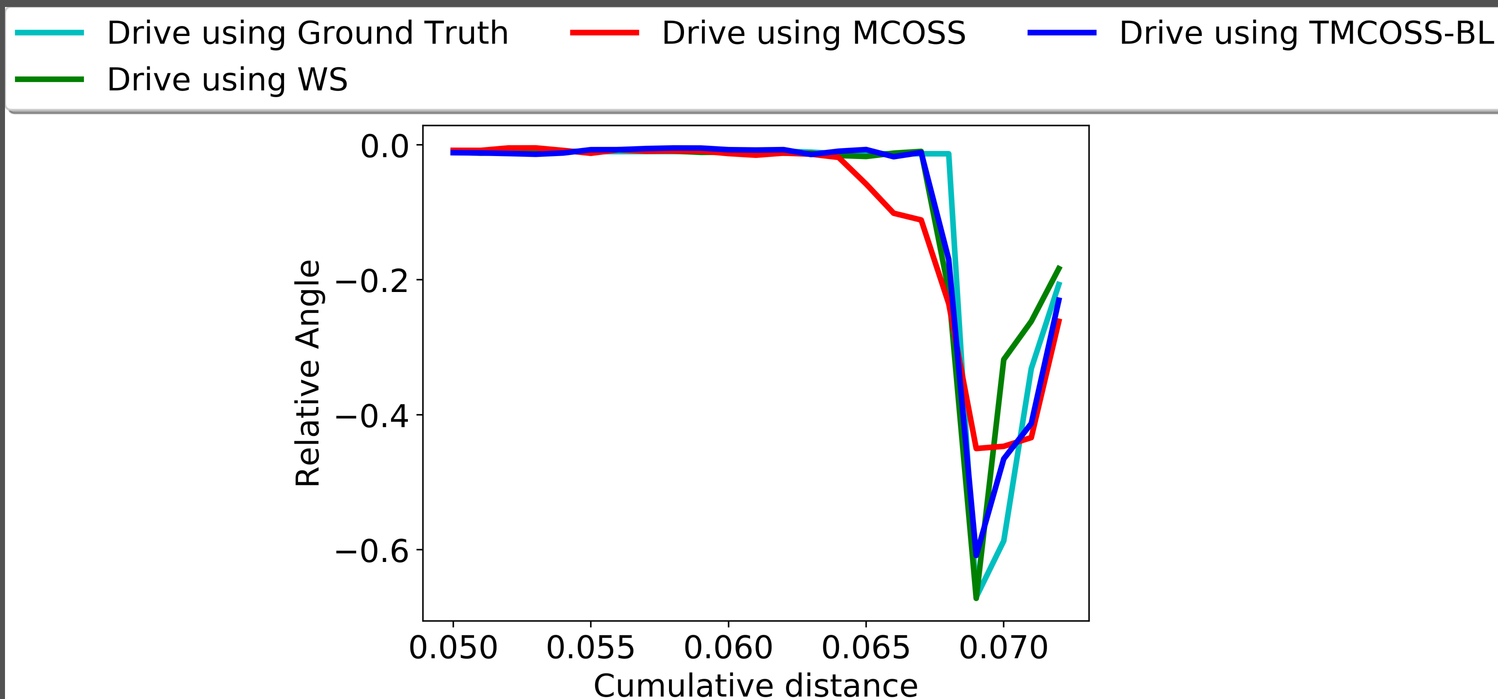
Method	Train One-Turn	Test One-Turn
Whole Set	10	10
MCOSS[1]	8	7
TMCOSS-TL	8	9
TMCOSS-BL	10	10

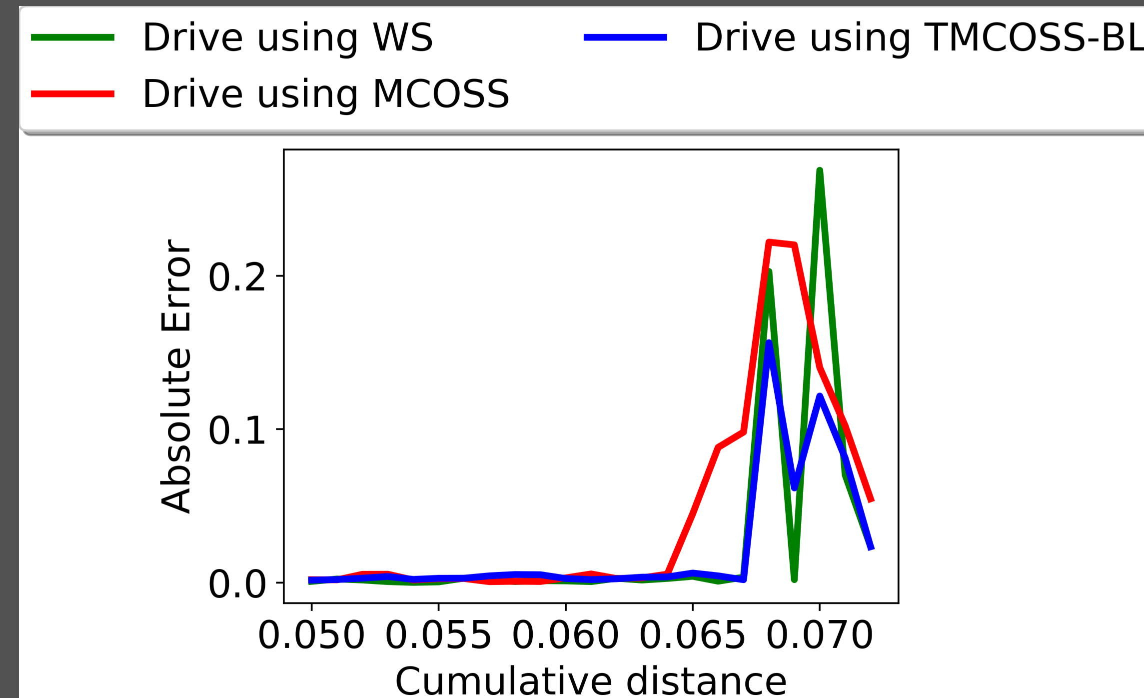
100:7 compression ratio

MCOSS[1]	5	4
TMCOSS-TL	7	9
TMCOSS-BL	8	9

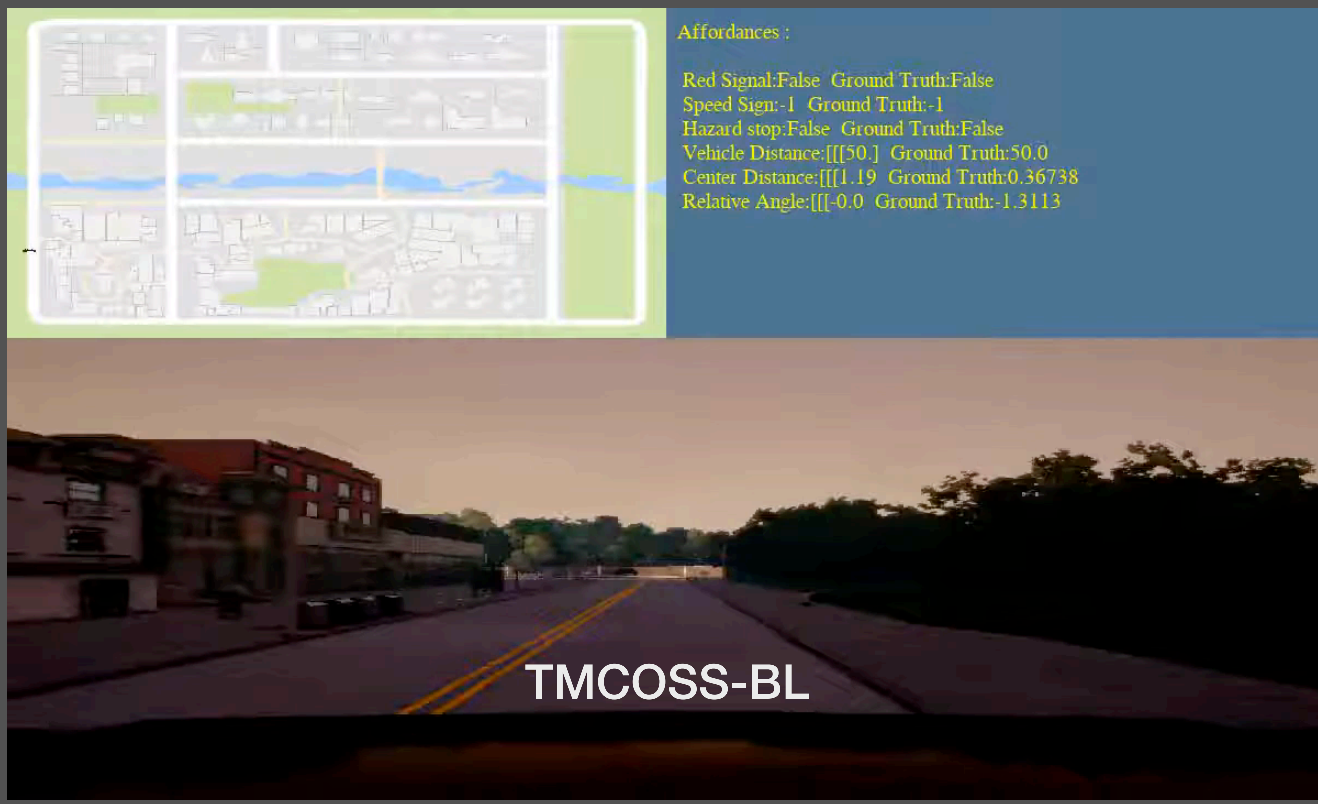
1. Soumi Das, Sayan Mondal, Ashwin Bhoyar, Madhumita Bharde, Niloy Ganguly, Suparna Bhattacharya, Sourangshu Bhattacharya, "Multi-criteria onlineframe-subset selection for autonomous vehicle videos." *Pattern Recognition Letters* 133 (2020): 349-355.

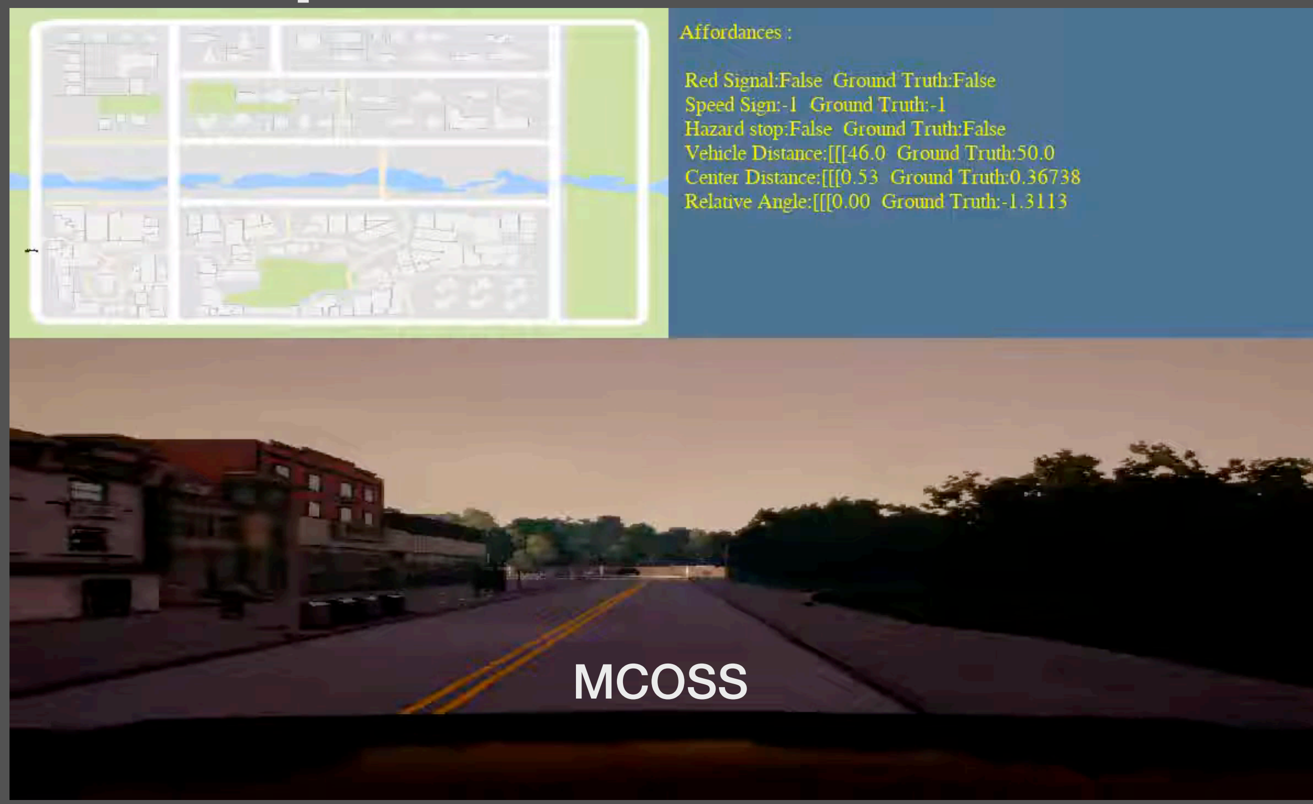
Anecdotal Example





Simulation of an episode





Conclusion

- We proposed TMCOSS, a thresholded multi-criteria convex optimisation based OSS method, and also a submodular variant termed as SubMCOSS.
- We study its effectiveness on driving data in terms of episode completion and on real world data (BDD, Cityscapes) in terms of semantic segmentation.
- We show that that TMCOSS is successful in maintaining an accuracy close to that of whole set even after dropping 80% frames.

Thank you for listening!!

Code: https://github.com/SoumiDas/TMCOSS