Multi-Class Lane Semantic Segmentation using Efficient Convolutional Networks

Shao-Yuan Lo, Hsueh-Ming Hang, Sheng-Wei Chan, Jing-Jhih Lin

National Chiao Tung University

MMSP 2019 Oral

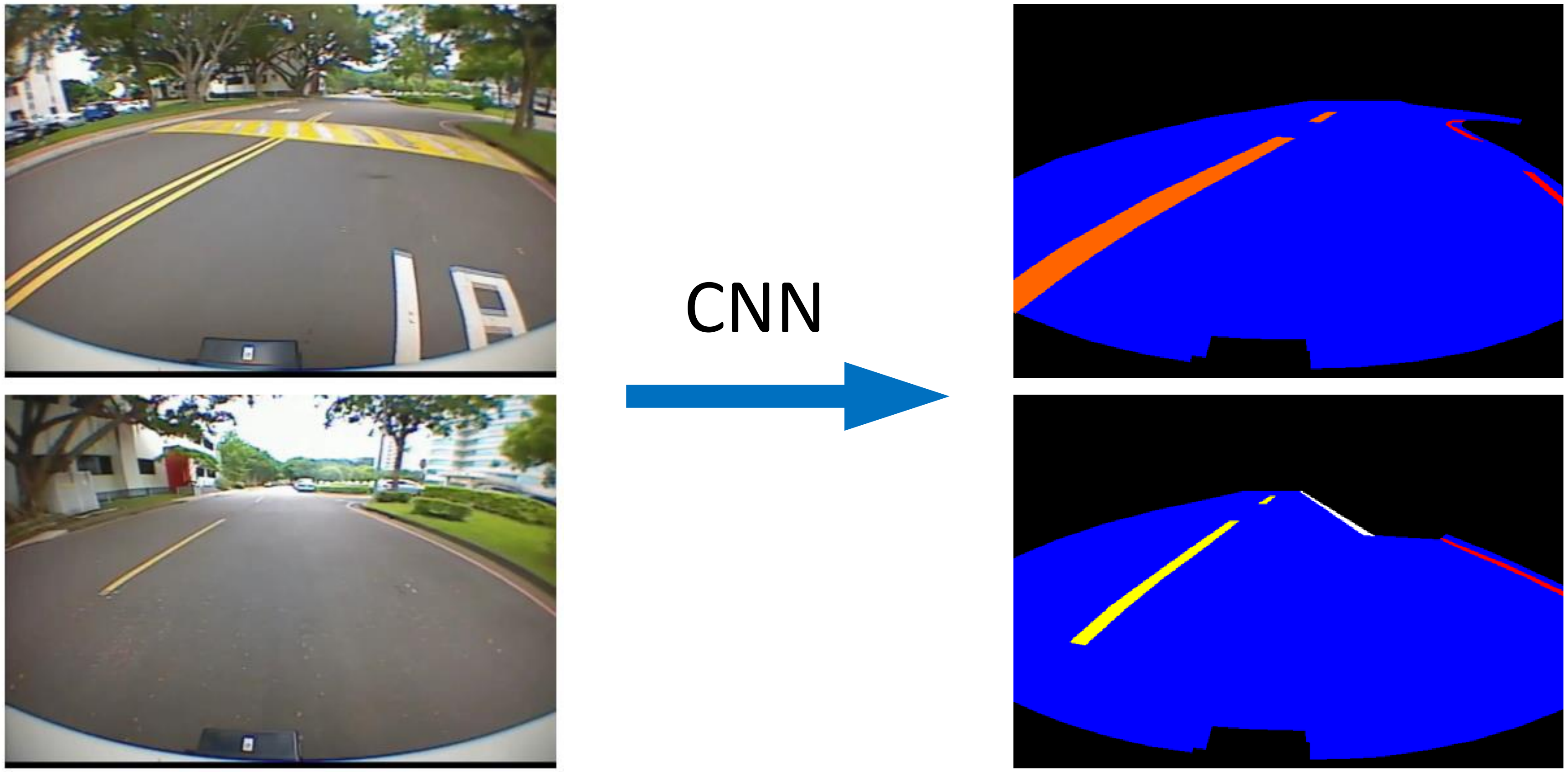
September 27, 2019

Outline

- Introduction
- Backbone Network: EDANet
- Methods
  - Feature Size Selection (FSS)
  - Degressive Dilation Block (DD Block)
- Experiments
  - Dataset / Implementation Details
  - Feature Size Selection (FSS)
  - Degressive Dilation Block (DD Block)
- Conclusion

Introduction

Introduction



Input:
RGB image

Output:
Semantic segmentation

Introduction

Self-driving applications

Efficiency

Accuracy
as high as possible

Inference Time
real-time

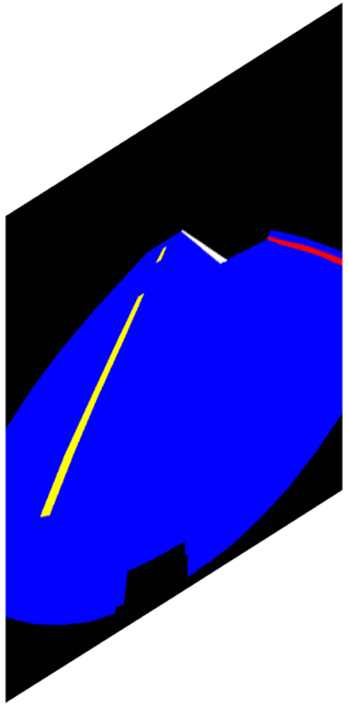
Model Size
low memory consumption

Introduction

The proposed
FSS or DD Block

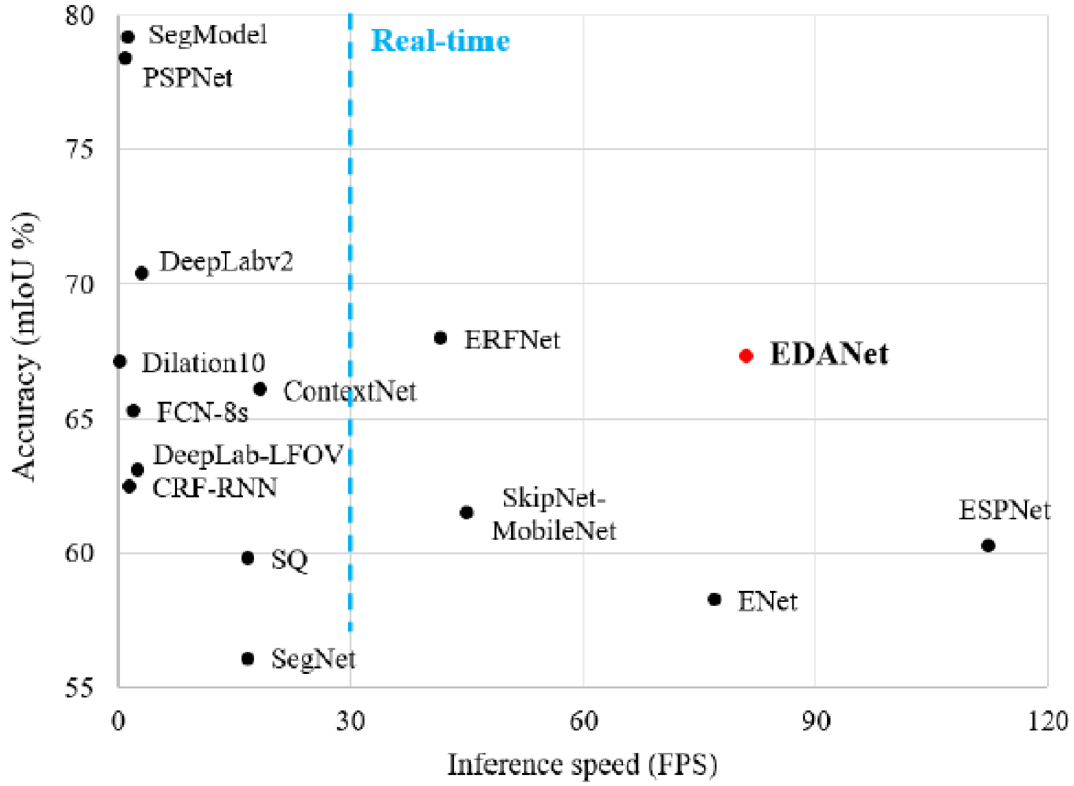
Existing semantic
segmentation
network (EDANet)

[Lo et al.]



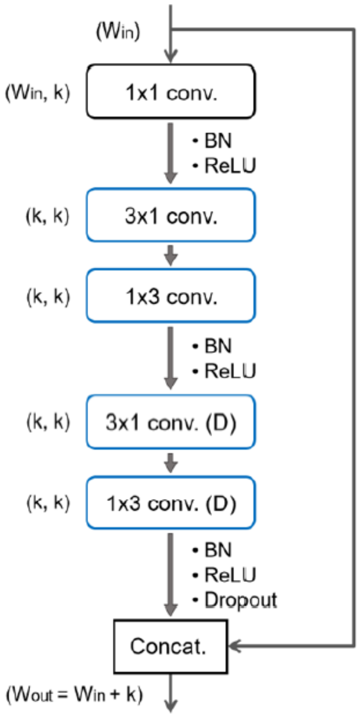
Backbone Network

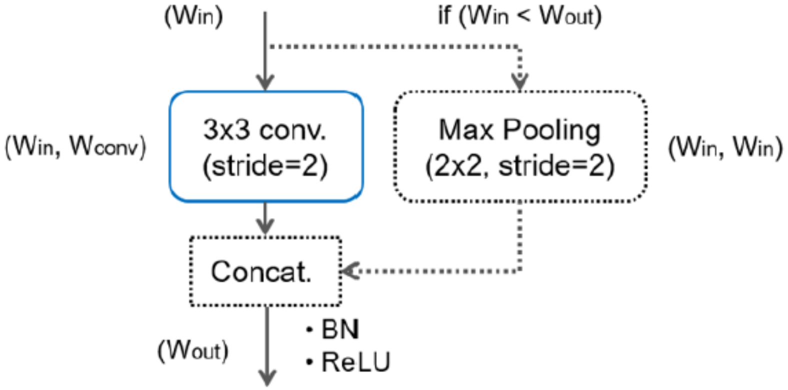
Backbone Network: EDANet

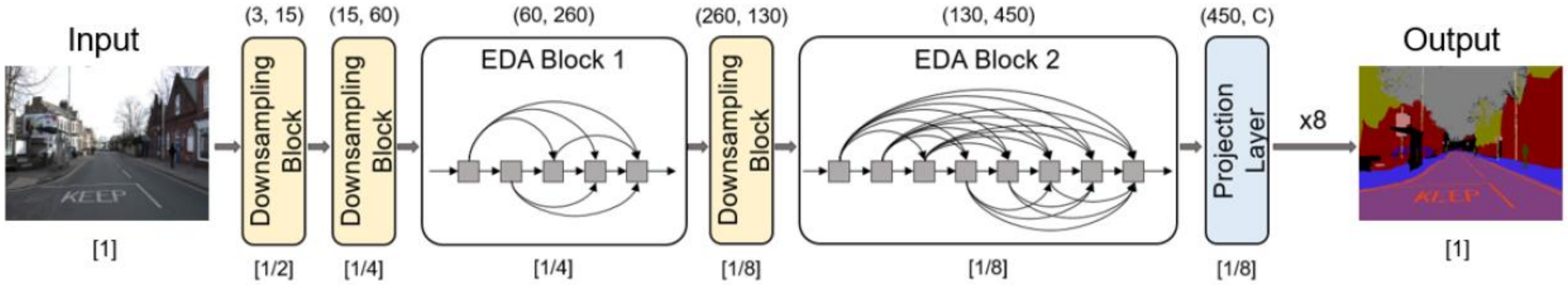


*S.-Y. Lo, H.-M. Hang, S.-W. Chan, and J.-J. Lin, "Efficient dense modules of asymmetric convolution for real-time semantic segmentation," arXiv preprint arXiv: 1809.06323, 2018.

Backbone Network: EDANet







Methods

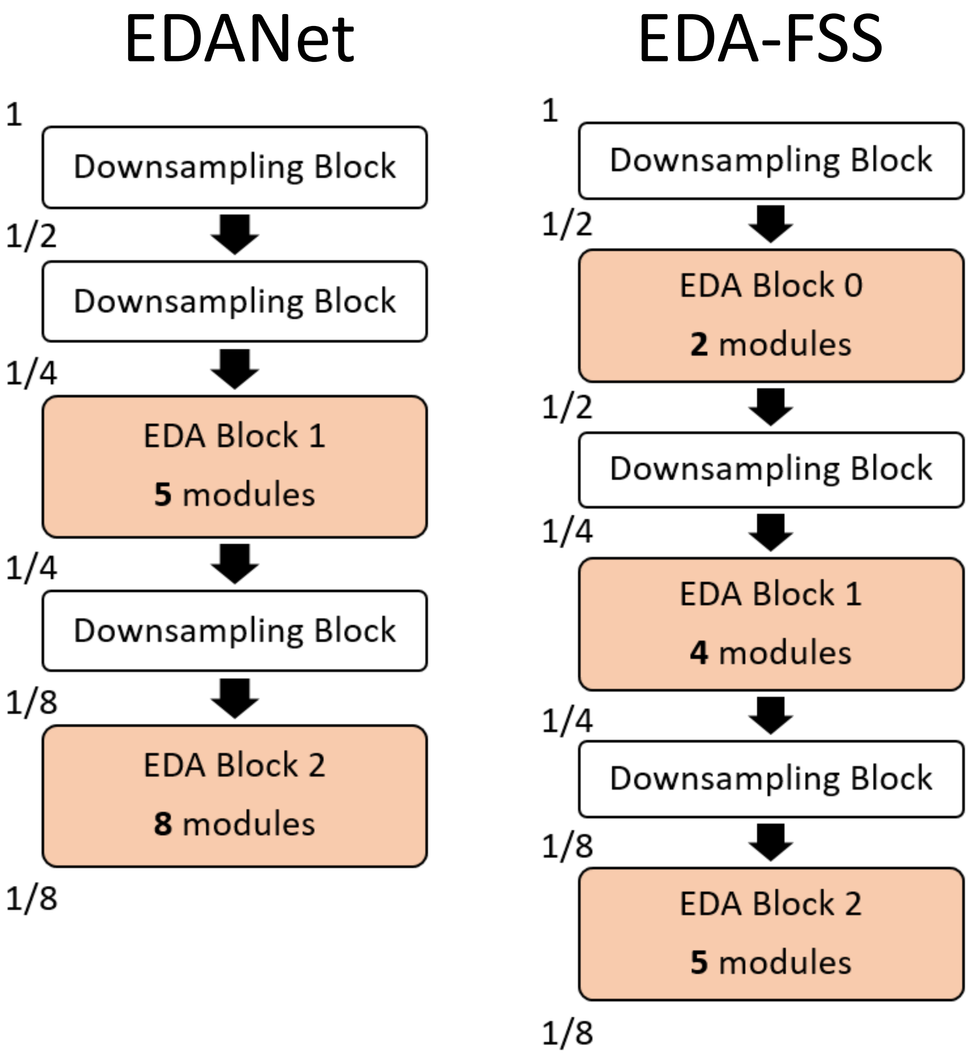
Methods

- Lane is a small object.
- Network modifications for small objects.

- Feature Size Selection (FSS)
- Degressive Dilation Block (DD Block)

Feature Size Selection

- Add an additional EDA Block 0 to extract features on larger feature maps, and thus small objects or detailed boundaries can be detected.
- To maintain similar computational complexity for high inference speed and fair comparison, we reduce the number of EDA modules in EDA Block 1 and EDA Block 2



Degressive Dilation Block

- Dilated convolution is not good at gathering essential local features because of the sparsity of convolutional kernels, which causes inaccurate recognition of small and thin objects.
- Inspired by LFE module [Hamaguchi et al.], we propose DD Block, which consists of 4 EDA modules with the **degressive dilation rates**.

Network	EDA Block 2								DD Block			
EDANet	2	2	4	4	8	8	16	16	-	-	-	-
EDA-DDB	2	4	8	16	-	-	-	-	8	4	2	1

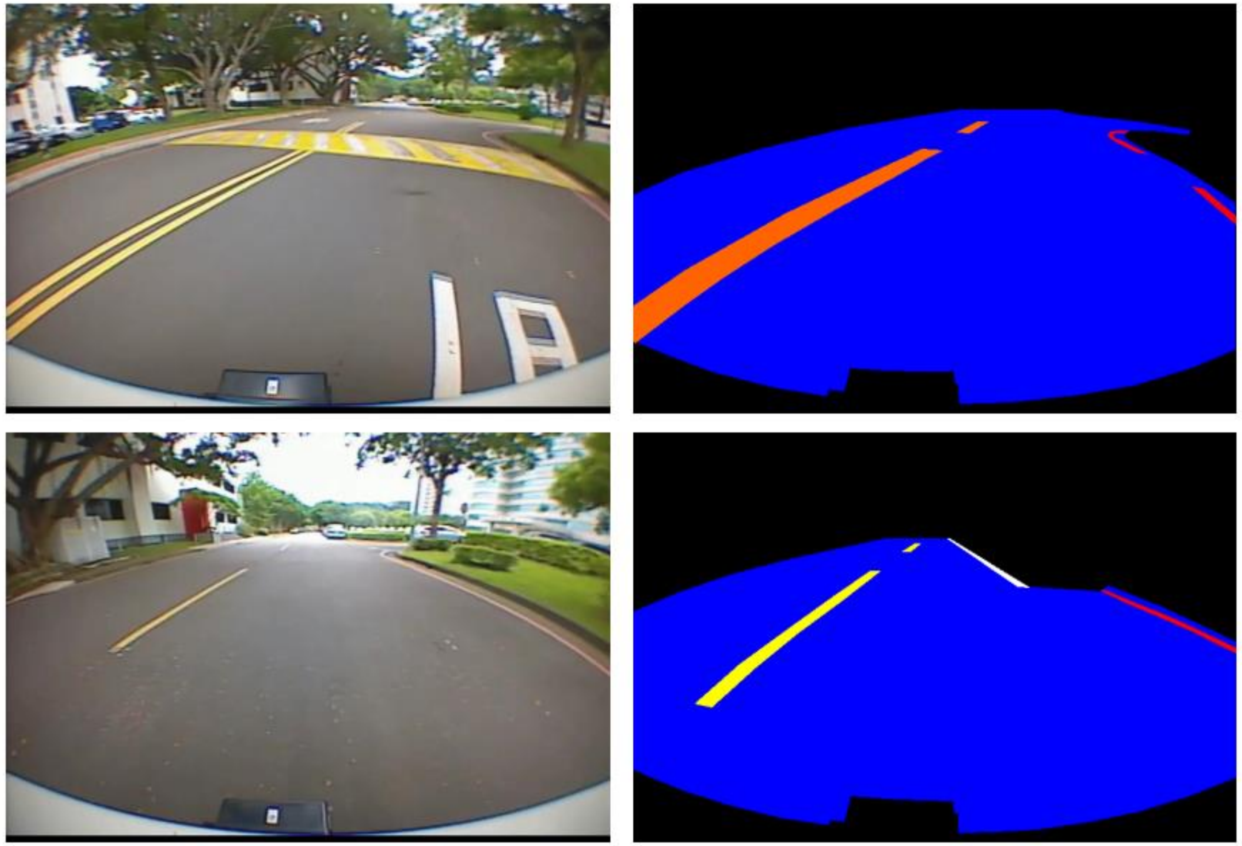
Experiments

ITRI Dataset

- Class: 6
- Double solid yellow, single dashed yellow, single solid red, single solid white

- Training data: 2192
- Test data: 567

- Resolution: 480 x 720



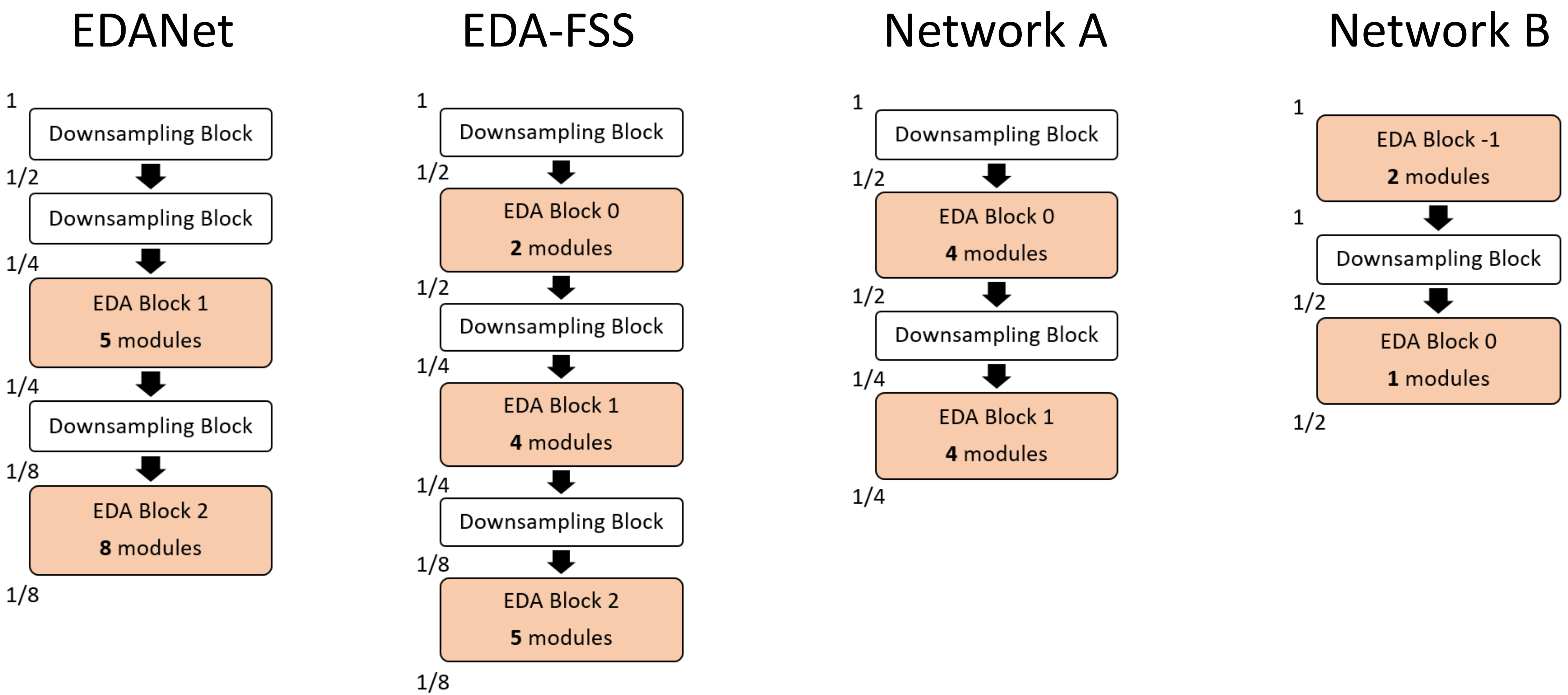
ITRI

Implementation Details

- Optimizer: Adam
- Batch size: 16
- Weight decay: 0.0001
- Poly learning rate policy: $lr' = lr \times (1 - iter/\max\_iter)^{power}$
- Initial learning rate: 0.0005
- Power: 0.9

- Data augmentation: random horizontal flip and translation of 0-2 pixels on both axes.

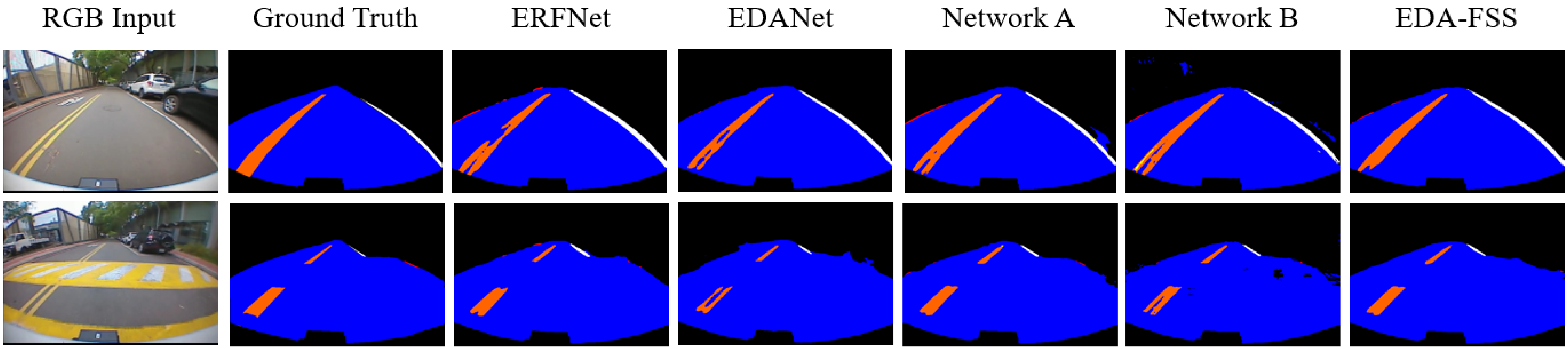
Experiment on FSS



Experiment on FSS

Network	DS-Y	SD-Y	SS-R	SS-W	Road	mIoU	Run time
ERFNet [Romera et al.]	78.6	82.4	39.1	46.1	96.9	73.3	24ms
EDANet	85.2	69.4	38.4	61.4	97.0	74.7	9.2ms
Network A	85.1	72.1	37.7	57.1	96.9	74.2	12ms
Network B	81.8	65.9	41.7	57.1	96.2	73.1	18ms
❤ EDA-FSS	84.5	74.4	43.0	54.8	96.9	**75.0**	**8.7ms**

Experiment on FSS



Experiment on DD Block

Network	EDA Block 2								DD Block			
EDANet	2	2	4	4	8	8	16	16	-	-	-	-
EDA-w/o-di	1	1	1	1	1	1	1	1	-	-	-	-
EDA-DDB-L	2	2	4	4	8	8	16	16	8	4	2	1
EDA-Large-1	2	2	4	4	8	8	16	16	1	1	1	1
EDA-Large-16	2	2	4	4	8	8	16	16	16	16	16	16
EDA-DDB	2	4	8	16	-	-	-	-	8	4	2	1

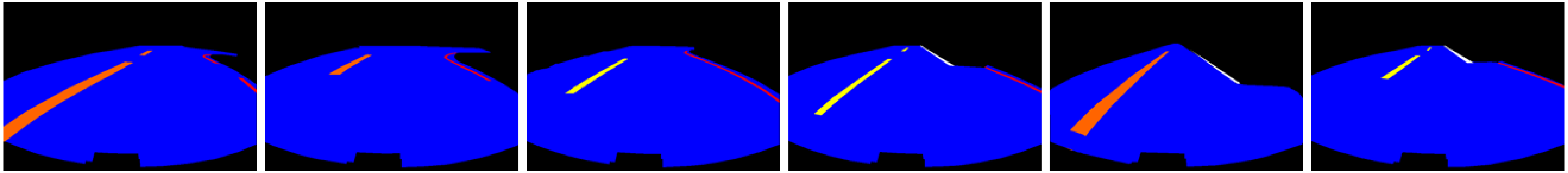
Experiment on DD Block

Network	DS-Y	SD-Y	SS-R	SS-W	Road	mIoU	Run time
EDANet	85.2	69.4	38.4	61.4	97.0	74.7	**9.2ms**
EDA-w/o-di	78.8	70.9	40.2	60.9	96.7	74.0	**9.2ms**
EDA-DDB-L	86.3	71.9	40.9	59.1	97.1	75.3	12ms
EDA-Large-1	82.7	68.7	37.4	58.7	96.6	73.4	12ms
EDA-Large-16	81.1	71.4	44.0	62.5	96.8	75.4	12ms
♥ EDA-DDB	86.6	76.4	43.6	57.6	97.1	**75.9**	**9.2ms**

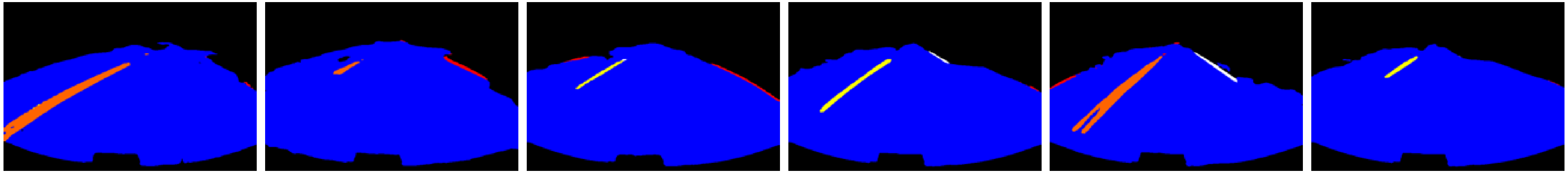
Experiment on DD Block

RGB input

Ground truth



EDA-DBB



Video Demo

Conclusion

Conclusion

- We propose two techniques, FSS and DD Block, for multi-class lane semantic segmentation.
- Using larger feature sizes can acquire more localization information for small object segmentation, but keeping a good balance between the network depth and inference speed is critical.
- Using degressive dilation rates achieves improvement by more fine-grained spatial information.
- Our methods are able to run at real-time on high-resolution inputs, so they are feasible for real self-driving cars.