INCREMENTAL LEARNING IN SEMANTIC SEGMENTATION

MLDL PROJECT – TA: FABIO CERMELLI

Vito Palmisano
s288859@studenti.polito.it

Valerio Zingarelli
s281586@studenti.polito.it

Daniele Falcetta
s289319@studenti.polito.it

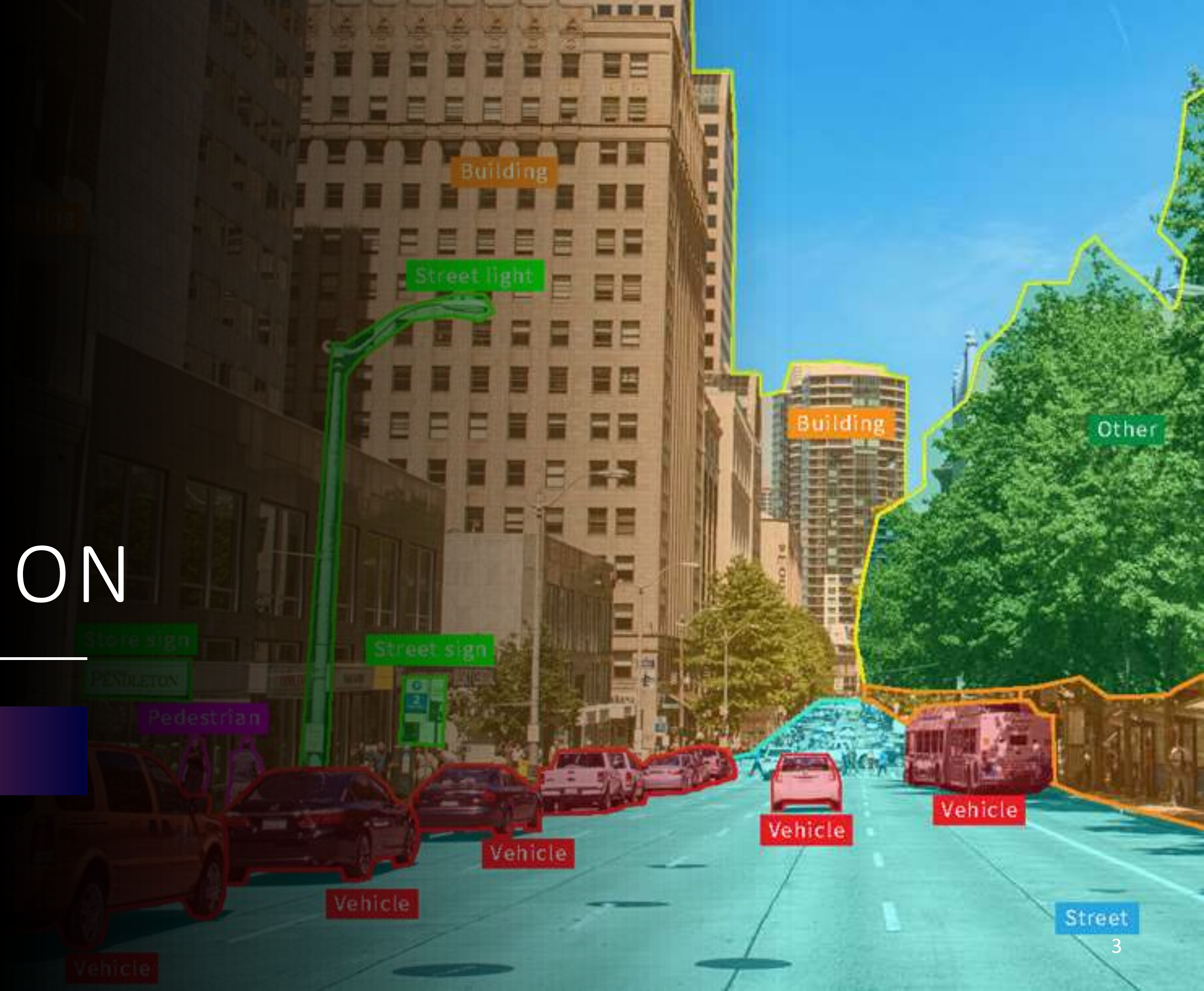
Politecnico di Torino

A.A. 2020/21

PROFESSOR: BARBARA CAPUTO

TABLE OF CONTENTS

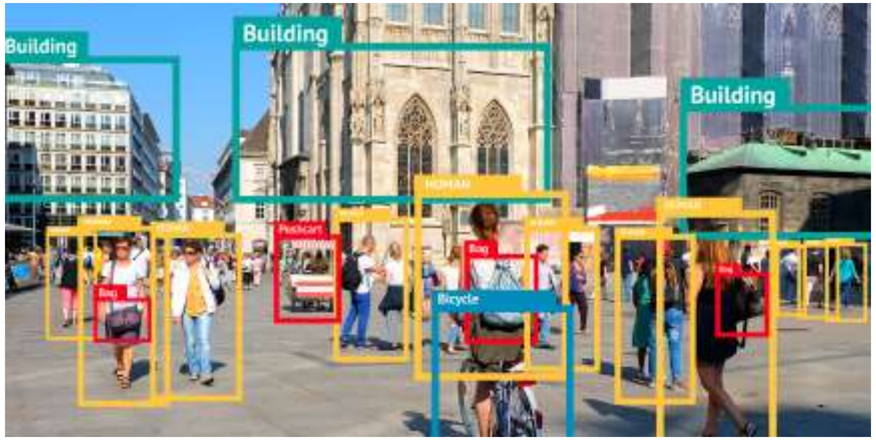
INTRODUCTION



INCREMENTAL LEARNING IN SEMANTIC SEGMENTATION

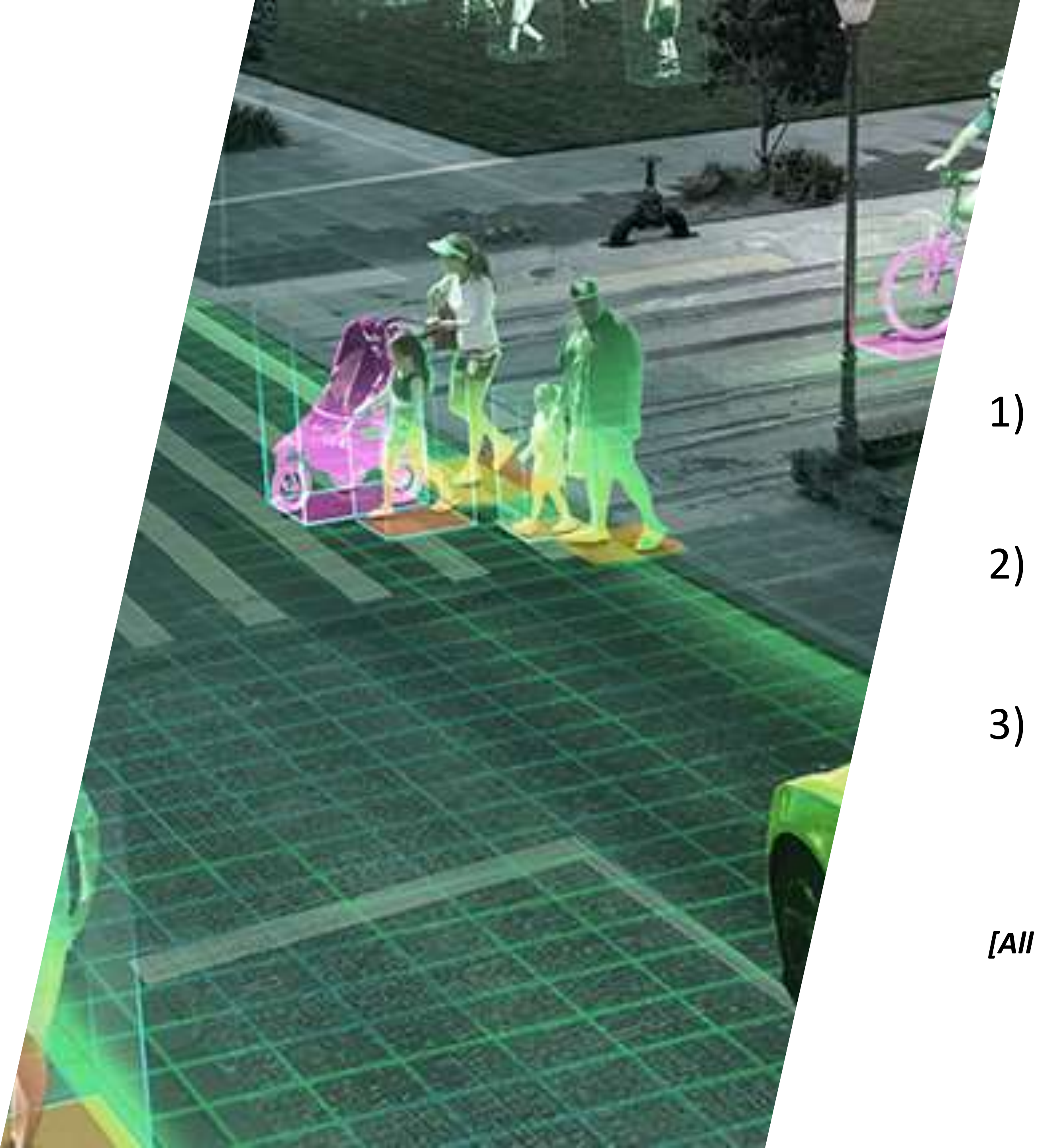
INTRODUCTION

(Offline) Semantic Segmentation



Object Detection

Incremental Semantic Segmentation



PROJECT’S STEPS

1) Implementation of BiSeNet Architecture in an **Offline Scenario**: Hyperparameters tuning

2) Testing of BiSeNet in an **Incremental Learning Setting**, following the MiB protocol

3) Our Contribution: Incremental Learning setting with **Weak Supervision** using pseudo-labels generated by SEAM

[All the settings are trained for 30 epochs]

BISENET:

BILATERAL SEGMENTATION NETWORK

BISENET

RELATED WORKS

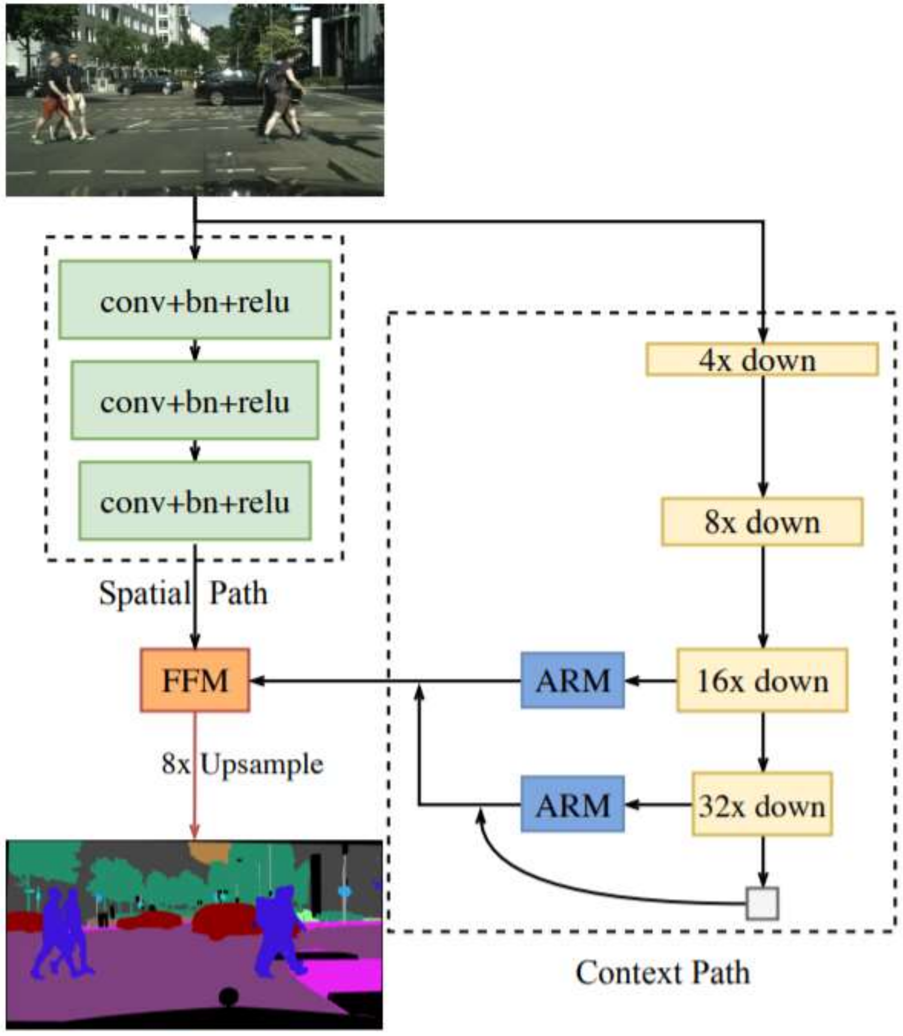
Offline Semantic Segmentation:
Good trade-off between Speed and Performances

Spatial Path:

1. Preserves spatial size of the original input image
2. Addresses the dropping of spatial information

Context Path:

1. Addresses the shrinkage of receptive fields
2. Utilizes lightweight model and an Attention Refinement Module (**ARM**)

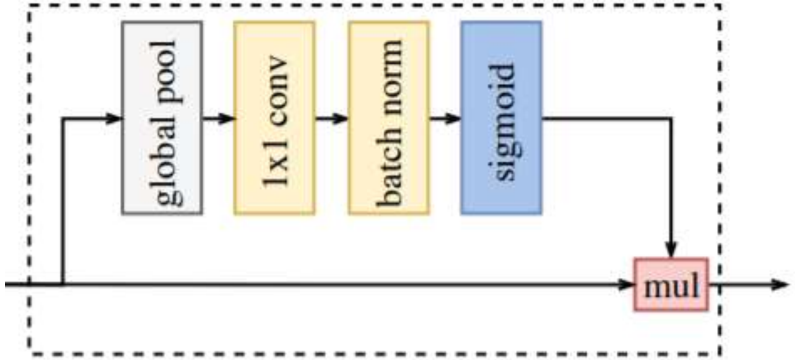


(a) Network Architecture

Changqian Yu, Jingbo Wang, Chao Peng, Changxin Gao, Gang Yu, Nong Sang:
BiSeNet: Bilateral Segmentation Network for Real-time Semantic Segmentation. *In ECCV 2018.*

BISENET

RELATED WORKS



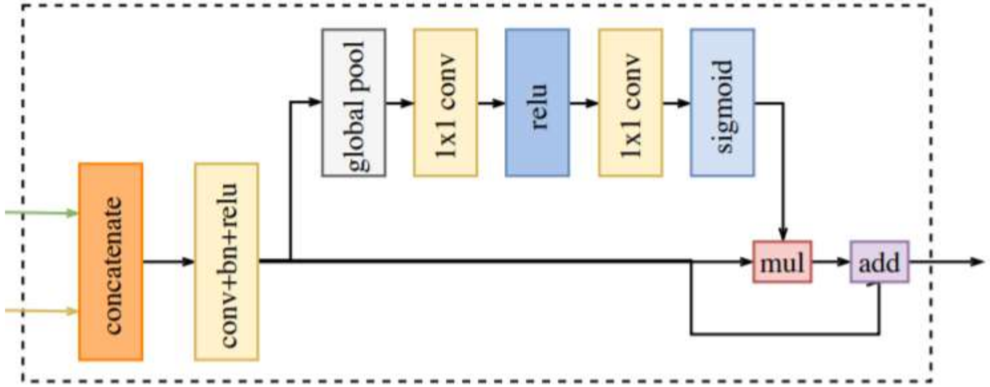
Attention Refinement Module (ARM)

Attention Refinement Module (ARM) :

1. Refines features of the last two stages of the ***Context Path***
2. Employs global average pooling to capture global context
3. Computes an attention vector to guide the feature learning

Feature Fusion Module (FFM):

1. Merges the two paths results
2. Scales the different features and re-weights them
3. Computes feature selection and combination



Feature Fusion Module (FFM)

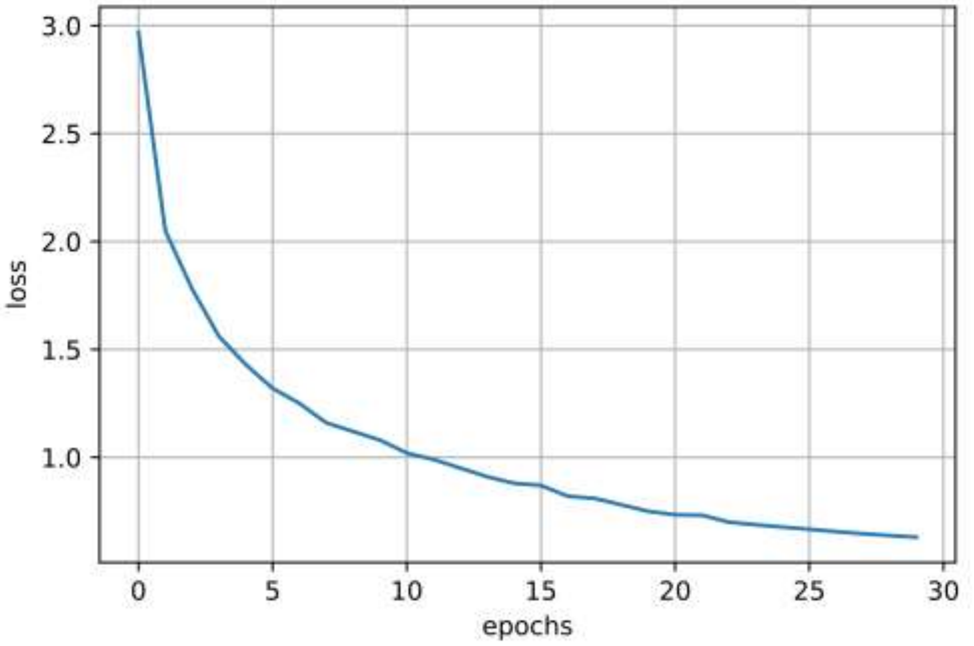
Changqian Yu, Jingbo Wang, Chao Peng, Changxin Gao, Gang Yu, Nong Sang:
BiSeNet: Bilateral Segmentation Network for Real-time Semantic Segmentation. In ECCV 2018.

BISENET

LOSS

Loss Function:

1. Combination of **Principal** Loss + **2 Auxiliary** Losses which supervise the output of the Context Path
2. All the loss functions are based on **Softmax** functions
3. In BiSeNet, *α = 1* and *k=3*



$$L(X, W) = l_p(X, W) + \alpha \sum_{i=2}^{k} l_i(X_i, W)$$

Changqian Yu, Jingbo Wang, Chao Peng, Changxin Gao, Gang Yu, Nong Sang:
BiSeNet: Bilateral Segmentation Network for Real-time Semantic Segmentation. In ECCV 2018.

BISENET: OFFLINE HYPERPARAMETERS TUNING

EXPERIMENTS AND RESULTS

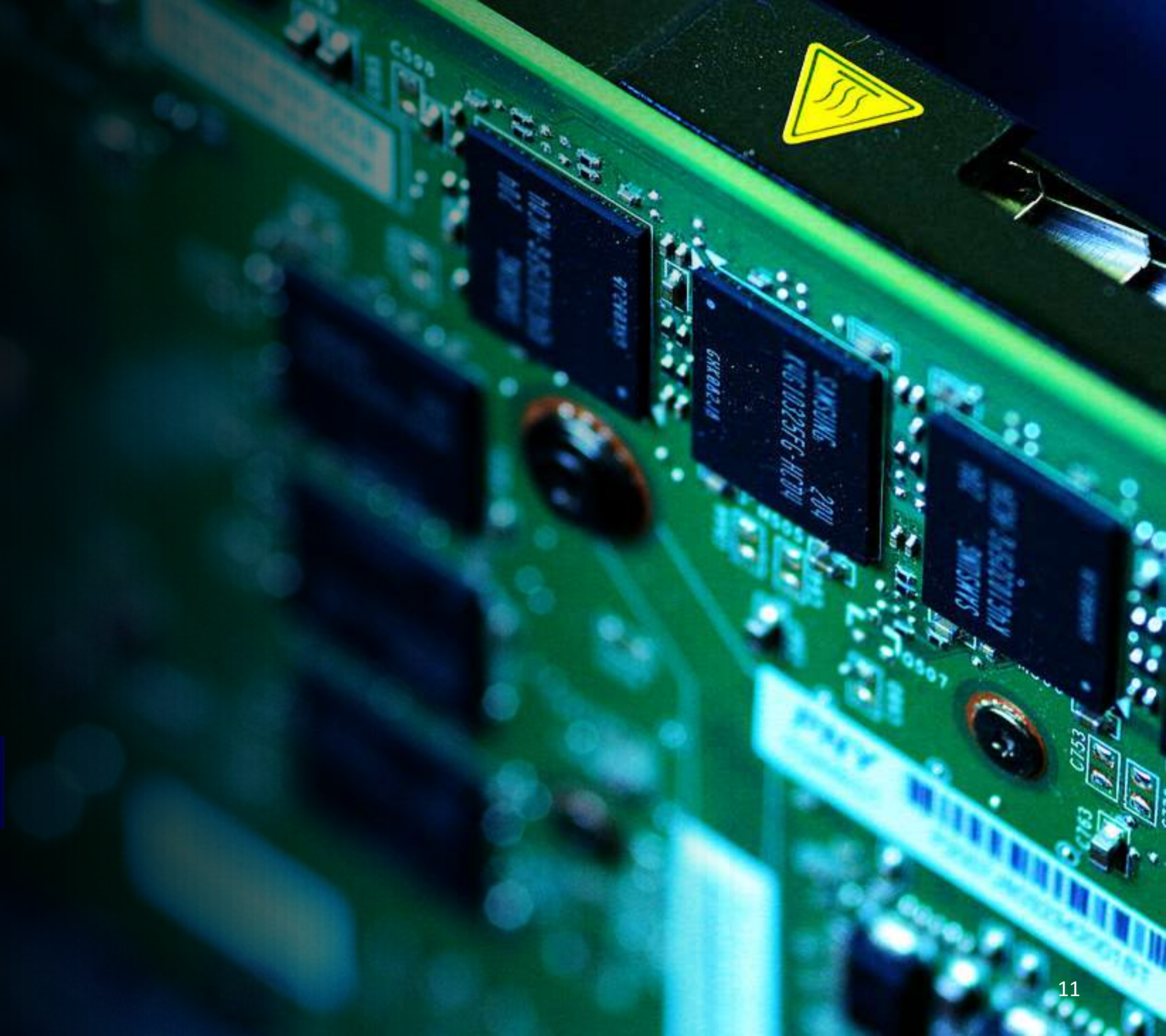
Backbone	Learning Rate	Batch Size	Precision	mIoU
ResNet18	0.001	16	88.0	53.3
ResNet18	0.002	32	88.9	54.2
ResNet50	0.001	16	90.6	64.0
ResNet101	0.001	16	90.1	62
ResNet101	**0.002**	**16**	**91.3**	**66.9**

CHOSEN MODEL

[In all the settings Data Augmentation is applied: Rotation, Vertical and Horizontal Flips, Crop and Color Jittering]

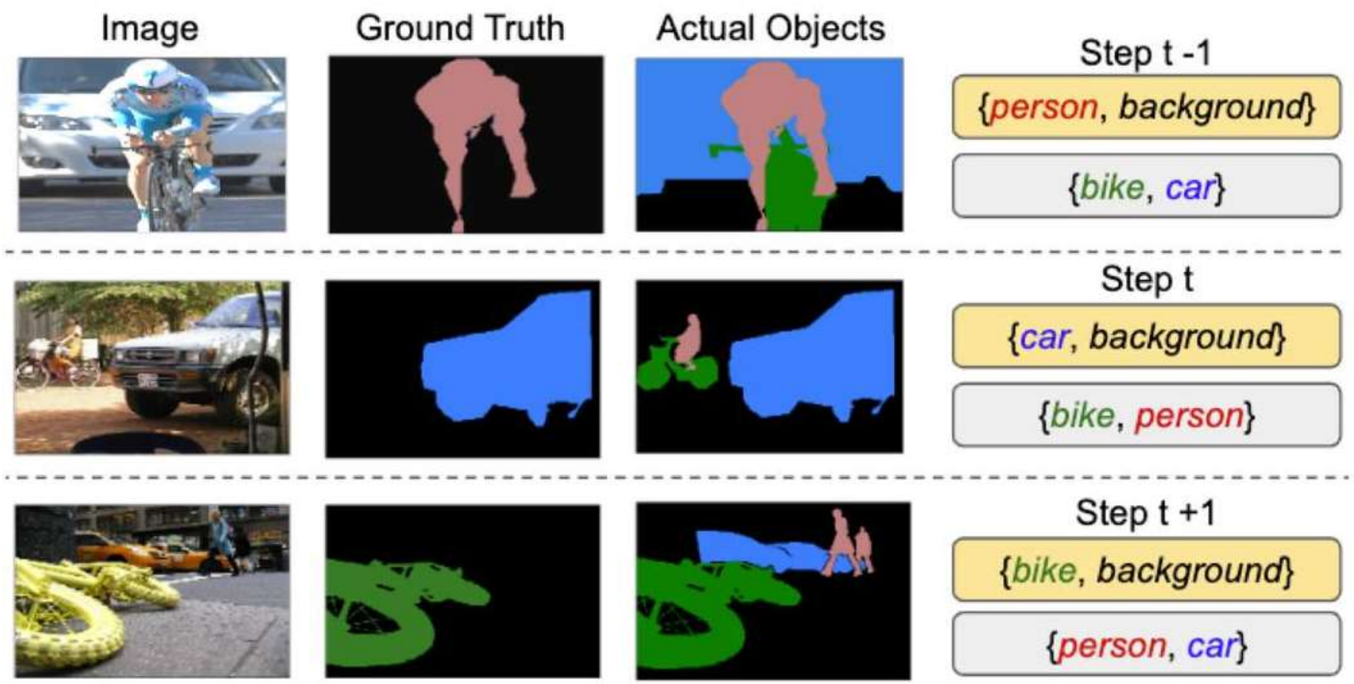
MIB:

MODELLING THE BACKGROUND



MIB

RELATED WORKS



Incremental Semantic Segmentation:

1. Address the shift of Background Class
2. Handles Catastrophic Forgetting
3. Revisiting the classical distillation framework by introducing two novel loss terms

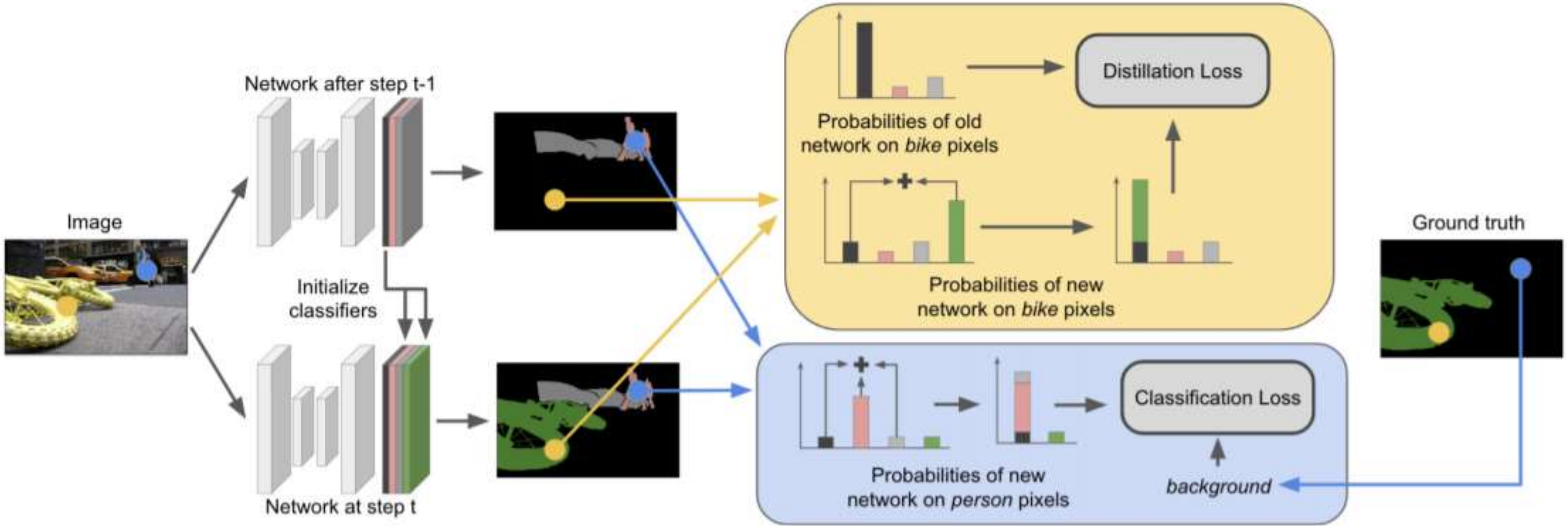
Fabio Cermelli, Massimiliano Mancini, Samuel Rota Bulò, Elisa Ricci, Barbara Caputo:
Modeling the Background for Incremental Learning in Semantic Segmentation. *In CVPR 2020.*

MIB

LOSS

$$L(\theta^t) = \frac{1}{|T^t|} \sum_{(x,y)\in T^t} \left(l_{ce}^{\theta^t}(x, y) + \lambda l_{kd}^{\theta^t}(x) \right)$$

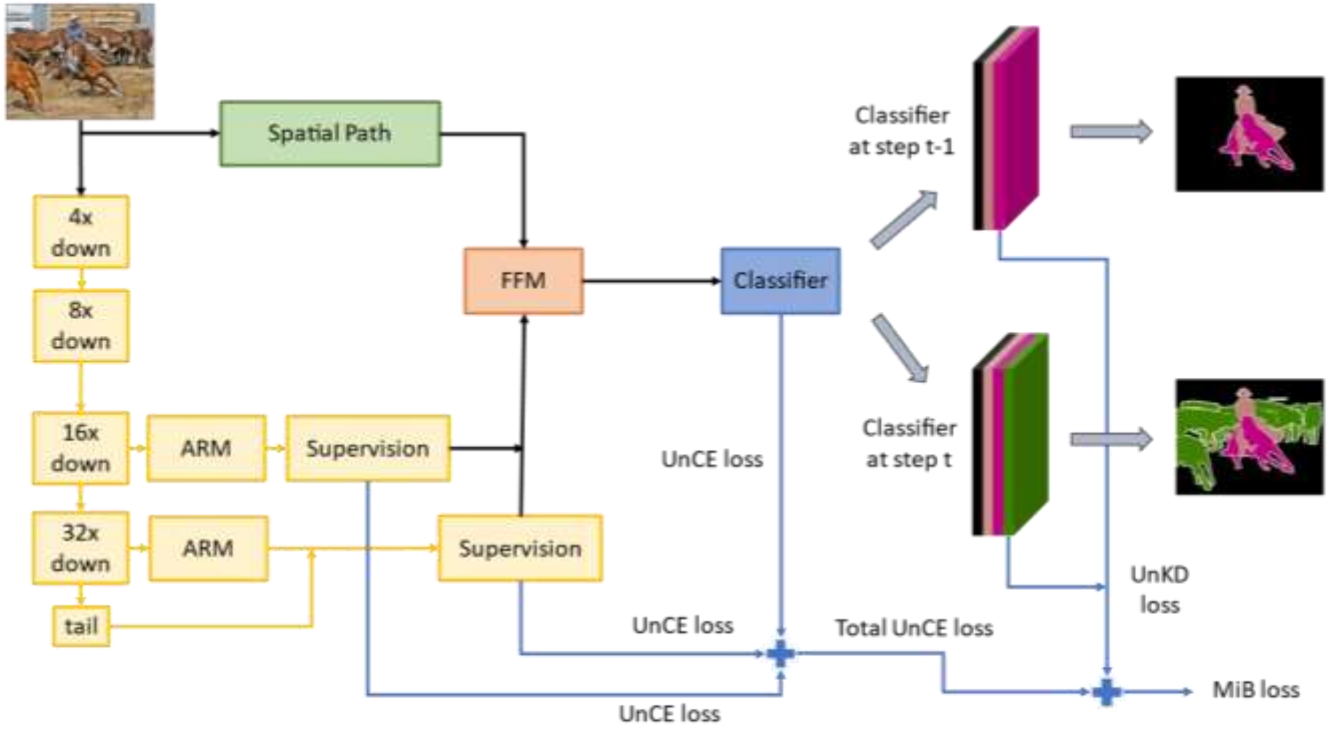
1. **Distillation Loss**: to preserve old knowledge
2. **Classification Loss**: to learn new classes



Fabio Cermelli, Massimiliano Mancini, Samuel Rota Bulò, Elisa Ricci, Barbara Caputo: ***Modeling the Background for Incremental Learning in Semantic Segmentation.*** *In CVPR 2020.*

BISENET IMPLEMENTATION INSIDE MIB

METHODS



Revisiting Losses:
Combination of MiB and BiSeNet to avoid compatibility issues

$$L(\theta^t) = \frac{1}{|T^t|} \sum_{(x,y)\in T^t} \left(L_{ce}^{\theta^t}(x,y) + \lambda L_{kd}^{\theta^t}(x) \right)$$

Distillation Loss → computed on the FFM ouput

MiB Revisited CE Loss → $L_{CE} = l_{FFM} + \sum_{i=2}^{k} l_{cx_i}$

1. Principal Loss computed on the FFM output
2. Two auxiliary Losses computed on the context path outputs

BISENET INTO MIB: INCREMENTAL SETTING

EXPERIMENTS AND RESULTS

mIoU	15-5			15-1		
	1-15	16-20	All	1-15	16-20	All
FT	5.2	32.5	12.0	1.3	4.1	2.0
LWF	51.6	63.2	47.7	5.3	7.8	5.9
ILT	59.1	35.8	53.3	3.4	6.3	4.1
MiB	63.4	40.9	58.2	16.1	7.9	14.2
Joint	76.6	69.9	74.9	76.6	67.9	74.9

21

Experiments are made using two different database splitting:

1. **15-5:** 15 classes for the first training step and 5 classes in the second incremental step

2. **15-1:** same 15 starting classes, then 1 single class is learnt in each of the 5 incremental steps

[FT= Fine-Tuning ---- LWF= Learning Without Forgetting ---- ILT= Incremental Learning Tehniques]

PROBLEM: WEAK SUPERVISION

OUR CONTRIBUTION

Image with Segmentations

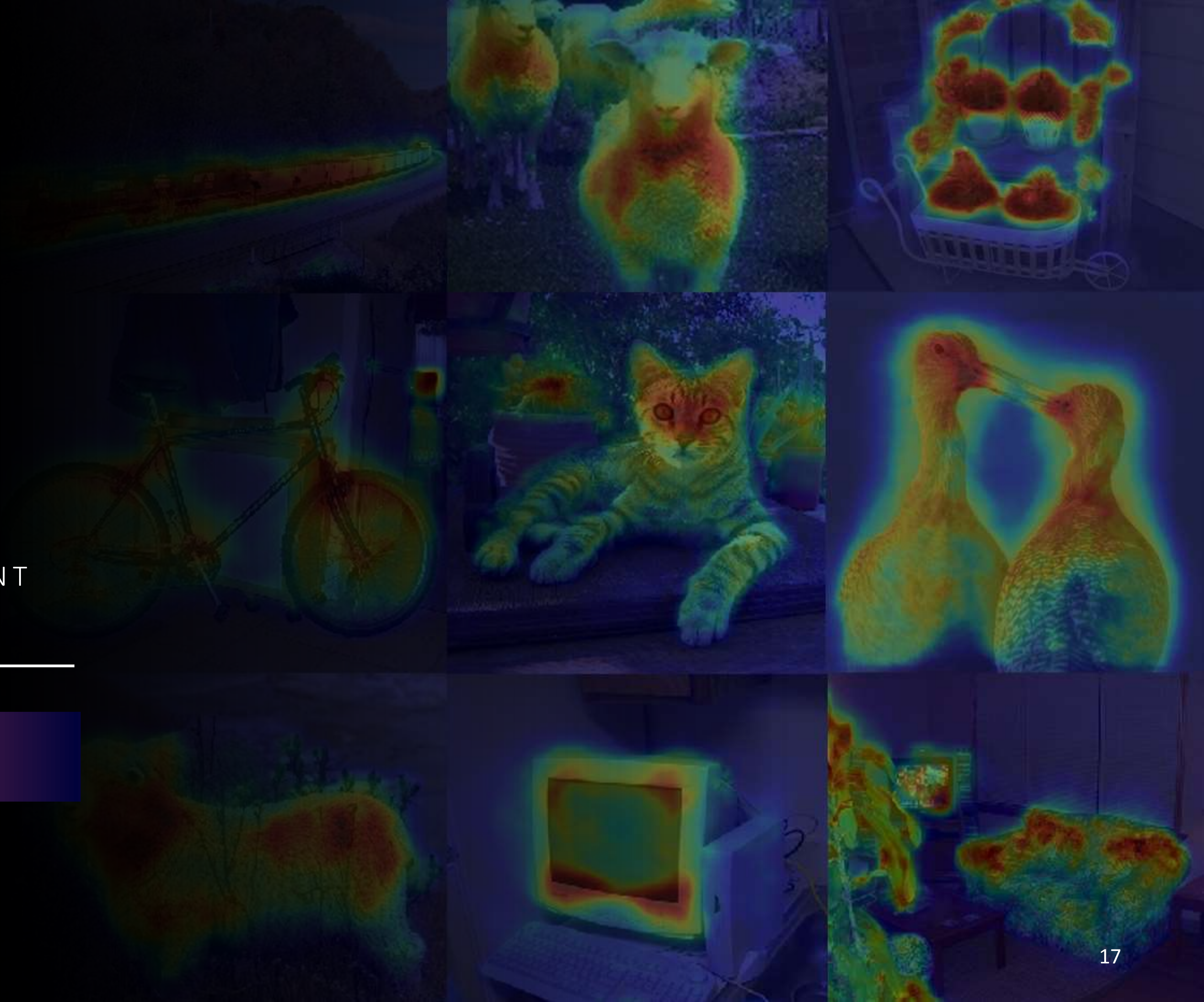
What if….?

Sky, Building, Road, Sidewalk, Fence, Vegetation, Pole, Sign, Car, Pedestrian, Cyclist

List of Classes

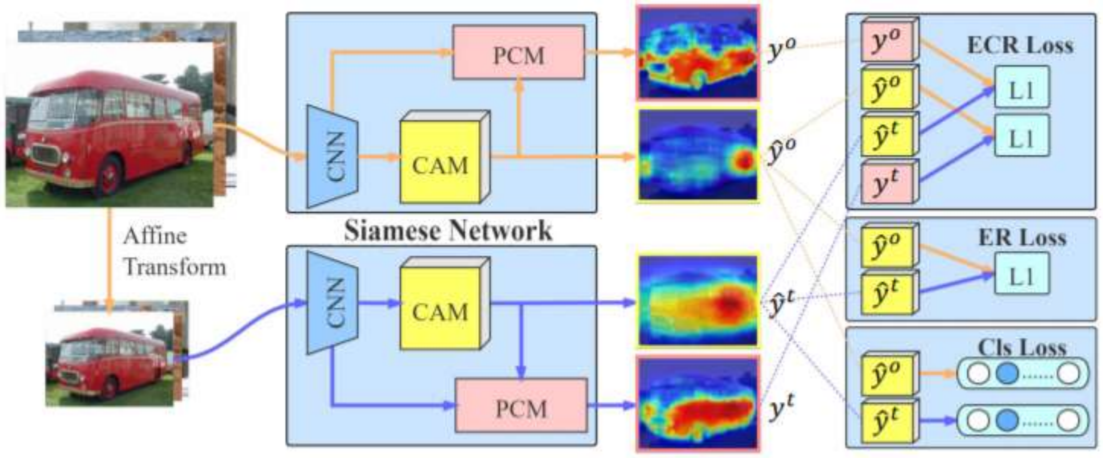
SEAM:

SELF-SUPERVISED EQUIVARIANT ATTENTION MECHANISM



SEAM

RELATED WORKS



Weak Supervision: Localize object from labels

Main **CAM problem**: over-activation in the zone of the image corresponding to the most significant part of the object

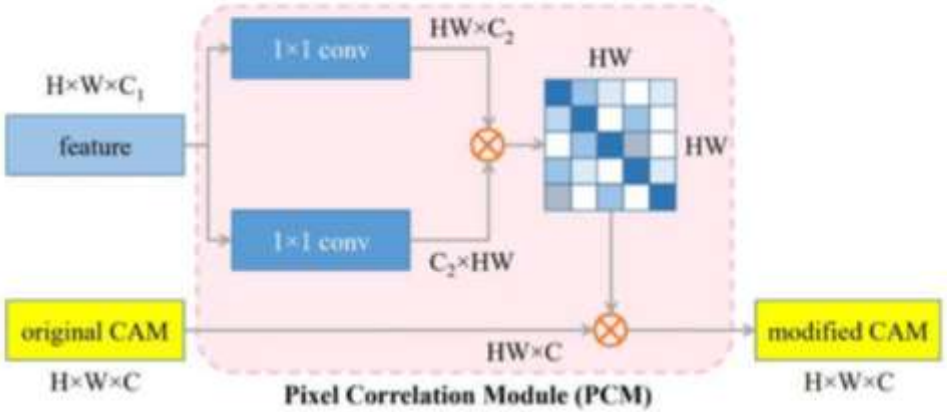
How SEAM try to solve it:

1. **Siamese Network** with an equivariant cross regularization (**ECR**) loss to efficiently couple the PCM and self-supervision
2. **Pixel Correlation Module (PCM)** to narrow the supervision gap between fully and weakly supervised semantic segmentation

Yude Wang, Jie Zhang, Meina Kan, Shiguang Shan, Xilin Chen:
Self-supervised Equivariant Attention Mechanism for Weakly Supervised Semantic Segmentation*. In CVPR 2020.*

SEAM

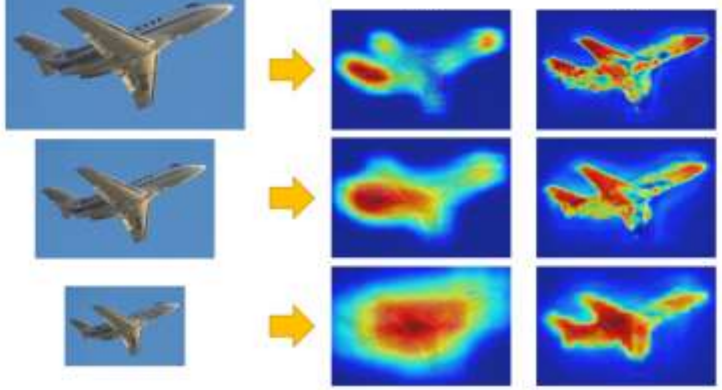
RELATED WORKS- LOSS



<u>Pixel Correlation Module (PCM):</u>

1. Takes features from an extractor
2. Two different 1x1 Convolutions
3. Merge the result with the ones produced by original CAM to obtain the new modified CAM

CAM *SEAM*



<u>Loss:</u>

1. **Classification Loss** to roughly localize objects
2. **ER Loss** to narrow the gap between pixel and image-level supevisions
3. **ECR Loss** to make predictions over various affine transformations

$$L = L_{ECR} + L_{cls} + L_{ER}$$

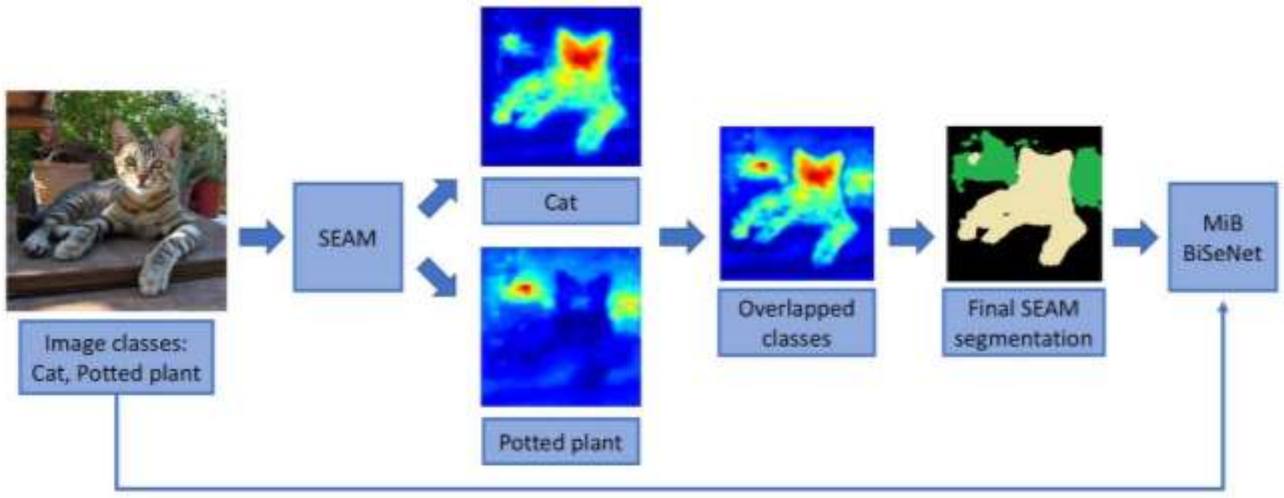
Yude Wang, Jie Zhang, Meina Kan, Shiguang Shan, Xilin Chen:
***Self-supervised Equivariant Attention Mechanism for Weakly Supervised Semantic Segmentation**. In CVPR 2020.*

SEAM PSEUDO-LABEL FOR THE INCREMENTAL STEP

METHODS

Weak Supervision

- ~~Annotated and classified Pixels~~
- List of classes for each image as label

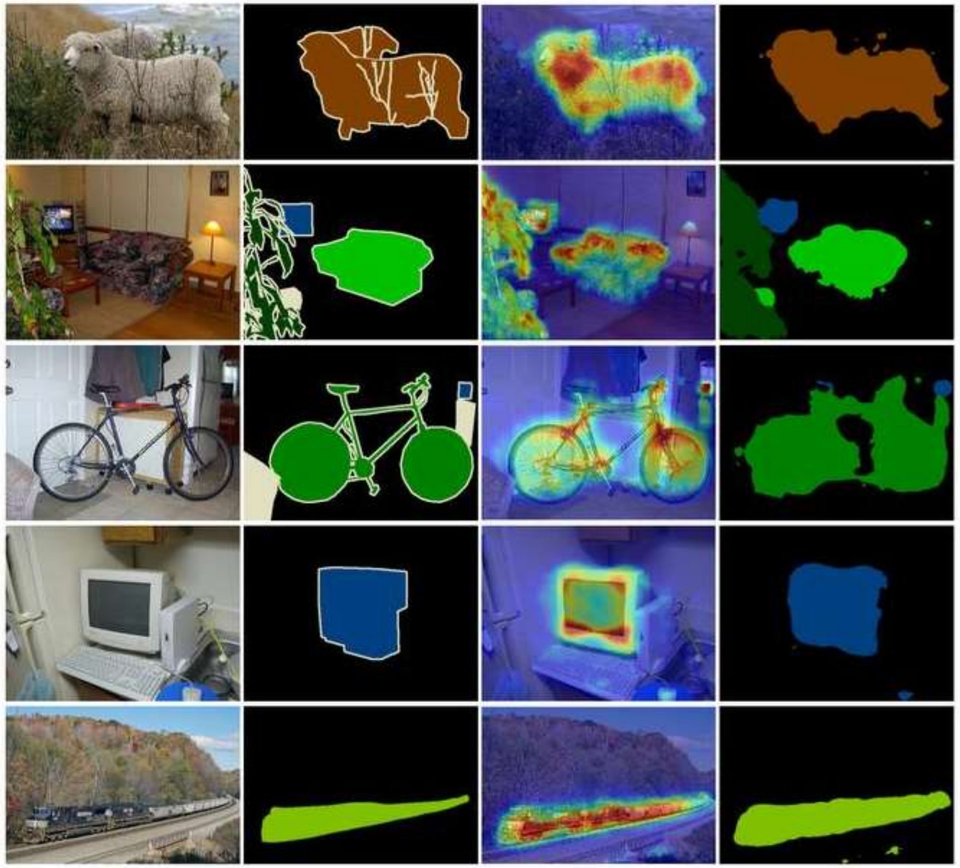


Generation of pseudo-label as ground truth for the incremental step through SEAM neural network

Incremental MiB step using SEAM segmentation labels

SEAM: WEAK SUPERVISION

EXPERIMENTS AND RESULTS



Original Image | *Original Segmentation* | *SEAM Probability Map* | *SEAM segmentation*

mIoU	15-5			15-1		
	1-15	16-20	All	1-15	16-20	All
FT-S	4.4	29.5	10.4	4.3	2.6	3.9
FT-VOC	5.2	32.5	12.0	1.3	4.1	2.0
MiB-S	62.5	34.4	56.1	15.5	1.3	11.9
MiB-VOC	63.4	40.9	58.2	16.1	7.9	14.2
Joint-S	55.4	52.0	54.6	55.4	52.0	54.6
Joint-VOC	76.6	69.9	74.9	76.6	67.9	74.9

SUMMARY AND CONCLUSIONS

CONCLUSION

SUMMARY

ANALYSIS OF THE MOST STUDIED SEMANTIC SEGMENTATION ISSUES

- **Offline Semantic Segmentation**: BiSeNet Architecture
- Implementation of BiSeNet in an **Incremental** Scenario
- **MiB** to handle catastrophic forgetting and the shift of the background problem

OUR CONTRIBUTION

- **Weak Supervision:** Development of a pipeline using **SEAM** to generate segmentation pseudo-labels
- Incremental train step in MiB using these **Pseudo-Label** as ground truth to learn new classes

THANK YOU FOR YOUR ATTENTION

PROFESSOR:
BARBARA CAPUTO

VITO PALMISANO S288859
VALERIO ZINGARELLI S281586
DANIELE FALCETTA S289319

TEACHING ASSISTANT:
FABIO CERMELLI

Machine Learning and Deep Learning

Politecnico di Torino
1859

A.A. 2020/21