Data exploitation: detection and segmentation with partially annotated data

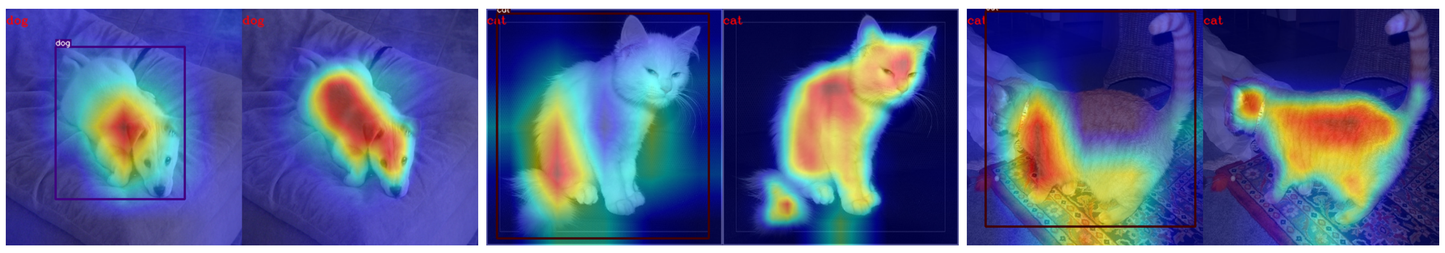
Hoàng-Ân Lê and Minh-Tan Pham

IRISA Vannes, Université Bretagne Sud

1 Motivation

Multi-task learning: learns interrelationship between tasks

- requires annotations of all tasks for each training example
- partially annotated data: each example is annotated for a single task;
- more data **but** hard to learn cross-task features



2 Experimental Setup

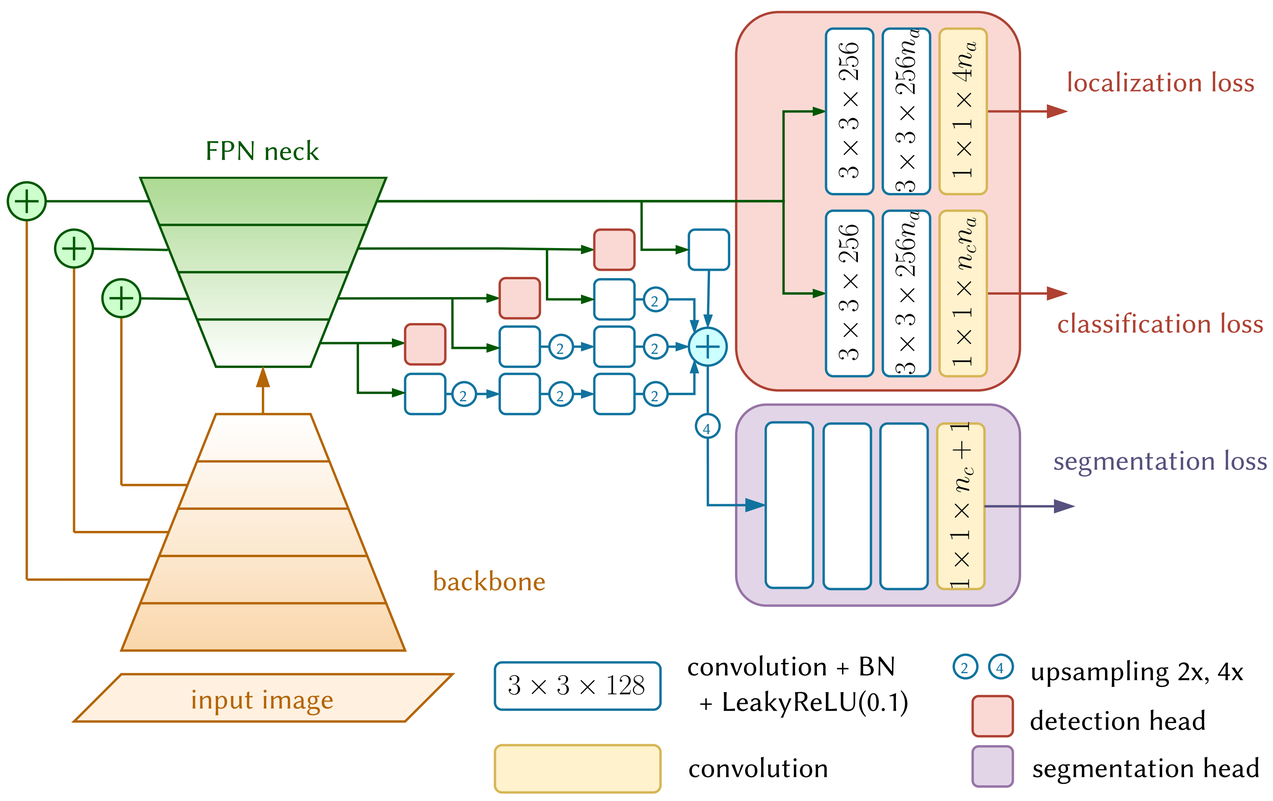
Datasets: Pascal VOC + augVOC: 20 categories

Object detection	Semantic segmentation
▶ ◖ = 7,558 images	▶ ◗ = 7,656 images
▶ ◔ = 3,776 images	▶ ◔ = 3,825 images

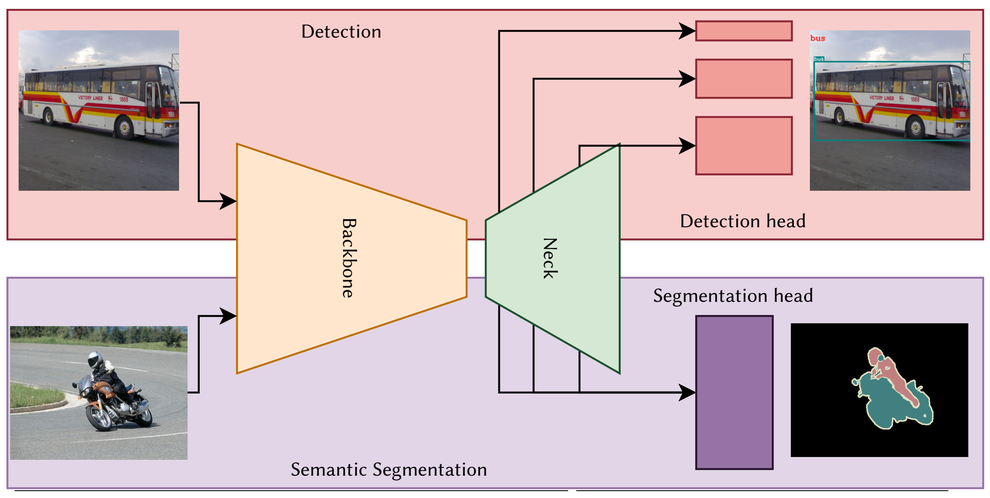
Metrics: mAP@[.5:.95]

Network: RetinaNet = ResNet family + FPN family
Segmentation head: FPN panoptic [1]
Knowledge distillation: feature imitation using MSE

3 Architecture

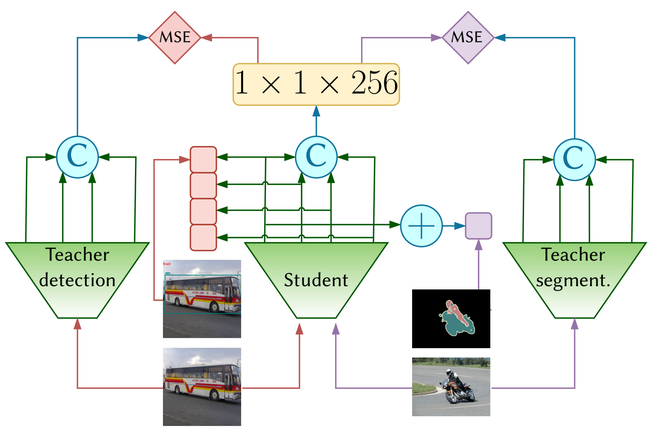


4 Multi-task partial annotation



Training		Detection			Segmentation	
		RN18+FPN	RN50+PAFPN		RN18+FPN	RN50+PAFPN
Single task	◖	42.81	50.22	◗	64.55	72.32
	◔	38.10	43.73	◔	63.02	68.96
Multi-task	◐	44.78	**52.10**		**67.57**	**73.66**
	◑	40.89	47.27		65.39	73.17
	◕	**44.99**	51.43		66.16	73.03

6 Multi-task learning with Knowledge Distillation



Training		Detection	Segmentation
Single task	◖	42.907	◗ 65.291
+ KD	◖	44.982	◗ 67.375
Multi-task	◐	45.678	67.310
+ 1mse	◐	45.989	69.126
+ 0mse	◐	47.337	**70.056**
+ 2mse	◐	**47.611**	69.911

5 Deviation

Different target categories: segmentation on 4 classes

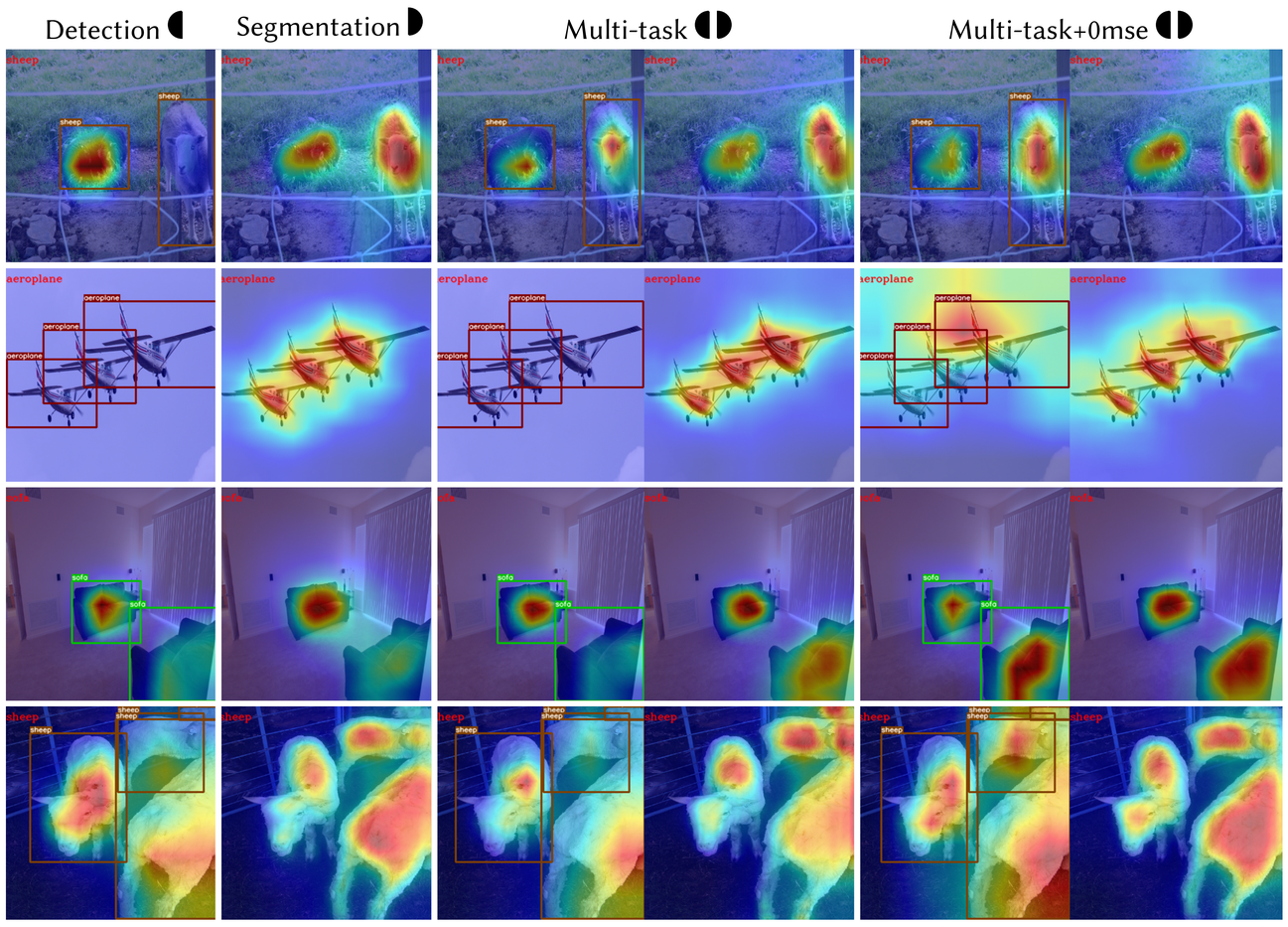
- transportation, animals, furniture, person

	Detection (20 classes)		Segmentation (4 classes)	
	RN18+FPN	RN50+PAFPN	RN18+FPN	RN50+PAFPN
Single task	42.81	50.22	78.47	81.82
Multi-task	**44.48**	**50.38**	**79.32**	**81.89**

Out-of-domain inputs: segmentation on Cityscapes

	Detection (VOC)		Segmentation (Cityscapes)	
	RN18+FPN	RN50+FPN	RN18+FPN	RN50+FPN
Single task	**38.688**	**44.683**	**71.389**	**72.398**
Multitask	37.531	39.910	69.481	70.247

7. Qualitative Results



Conclusions

Combining training data for different tasks helps

- cross-task optimization by self-training improves further

Deviation of tasks is to be studied

- different target categories
- different domains of inputs

Acknowledgement

This work was supported by the SAD 2021 ROMMEO project (ID 21007759) and the ANR AI chair OTTOPIA project (ANR-20-CHIA-0030).

References

[1] A. Kirillov, K. He, R. Girshick, C. Rother, and P. Dollar, "Panoptic Segmentation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.