

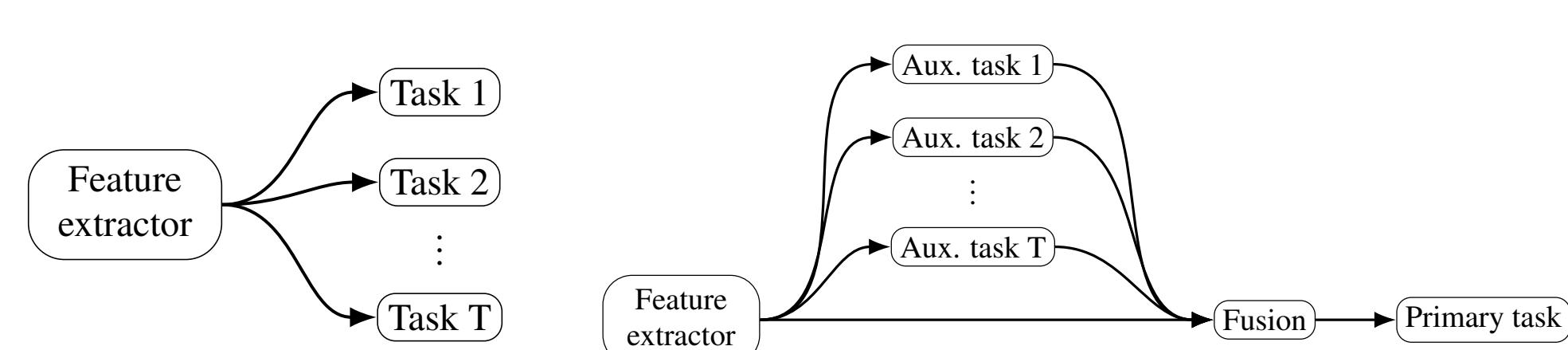
# REVISITING MULTI-TASK LEARNING WITH ROCK: A DEEP RESIDUAL AUXILIARY BLOCK FOR VISUAL DETECTION

Taylor MORDAN, Nicolas THOME, Gilles HENAFF, Matthieu CORD

## CONTEXT: MULTIPLE TASKS

- Solution to data starving: Transfer Learning
- sequential transfer: Fine Tuning [1]
- parallel transfer: Multi-Task Learning [2, 3]

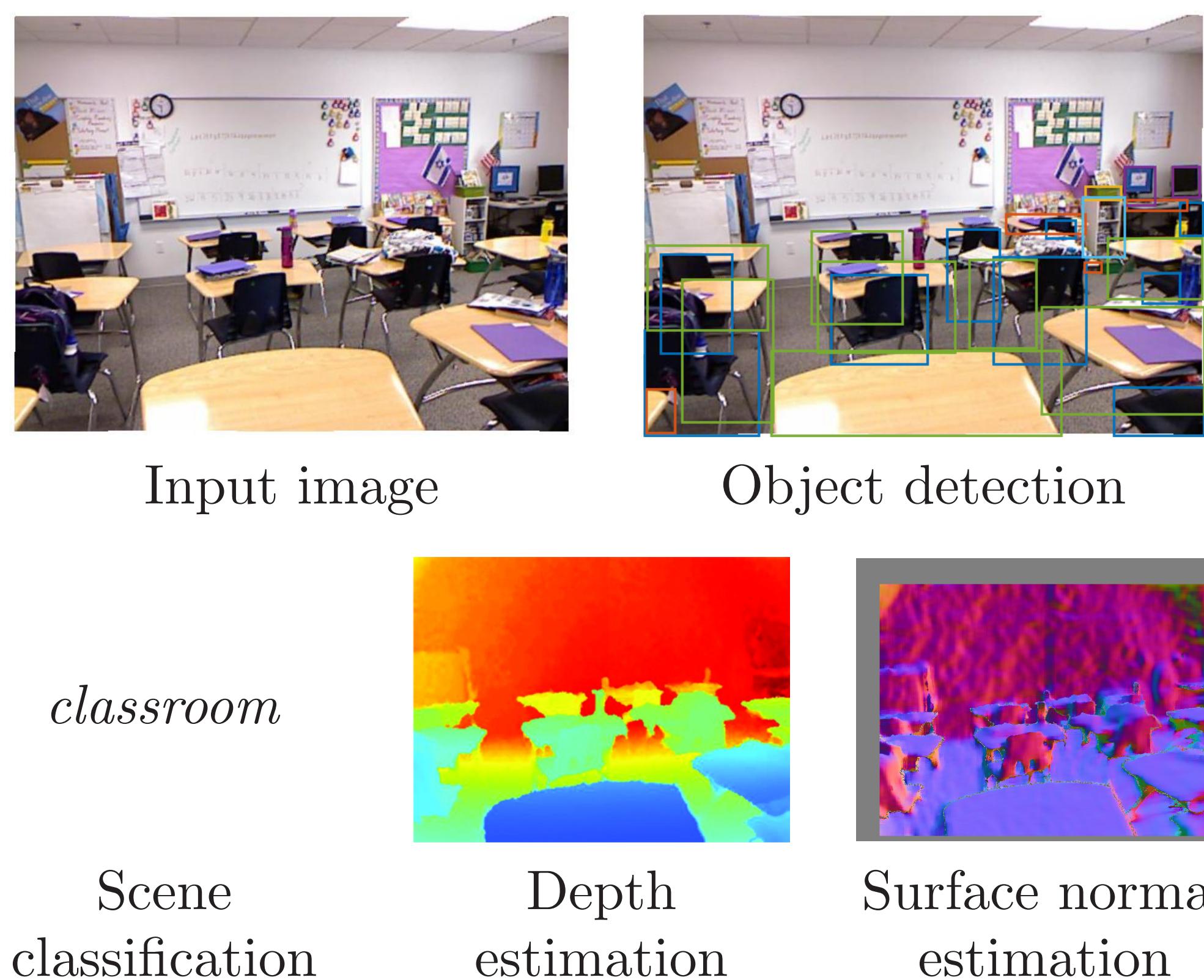
### Primary MTL



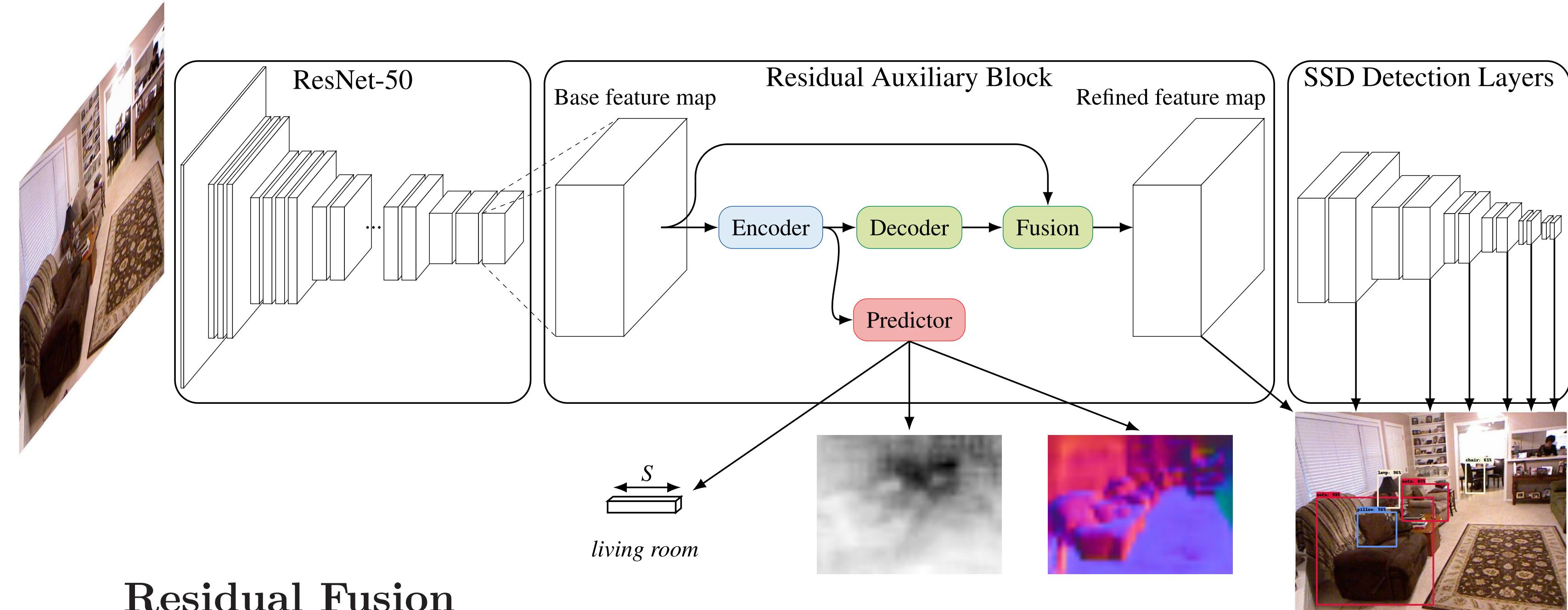
Flat MTL      Primary MTL

- Primary task (focus)  $\neq$  Auxiliary tasks (help)
- Related to privileged information (LUPI) [4, 5]

### Multi-Modal Object Detection



## ROCK: RESIDUAL AUXILIARY BLOCK

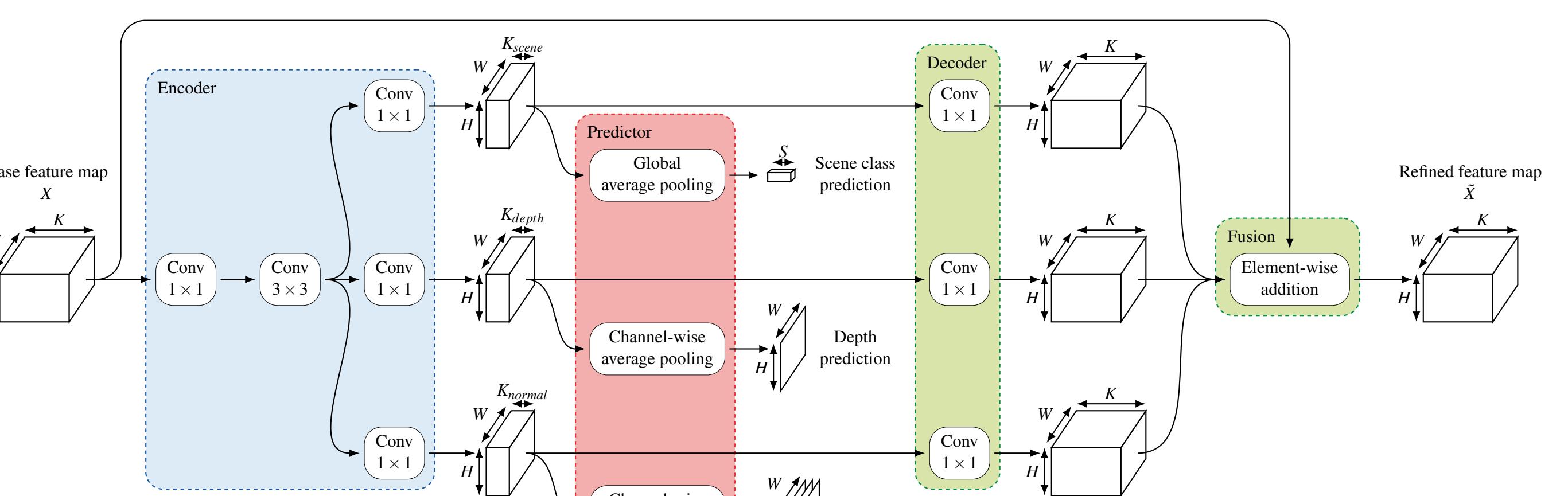


### Residual Fusion

- Break symmetry between tasks
- Explicit influence of auxiliary tasks

$$\tilde{X} = X + \sum_{i=1}^T Dec_t_i(Enc_t(X))$$

- Other fusions possible



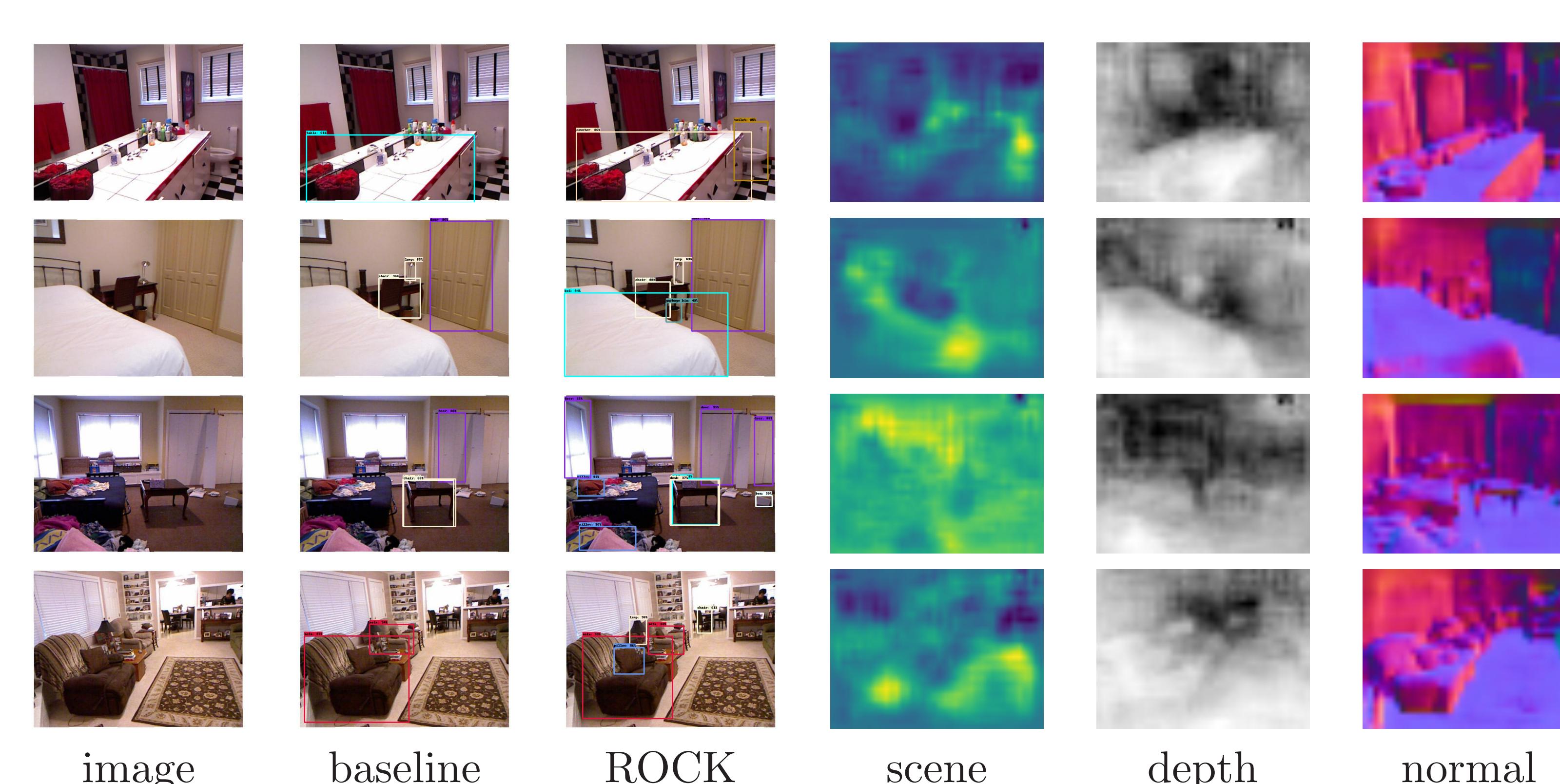
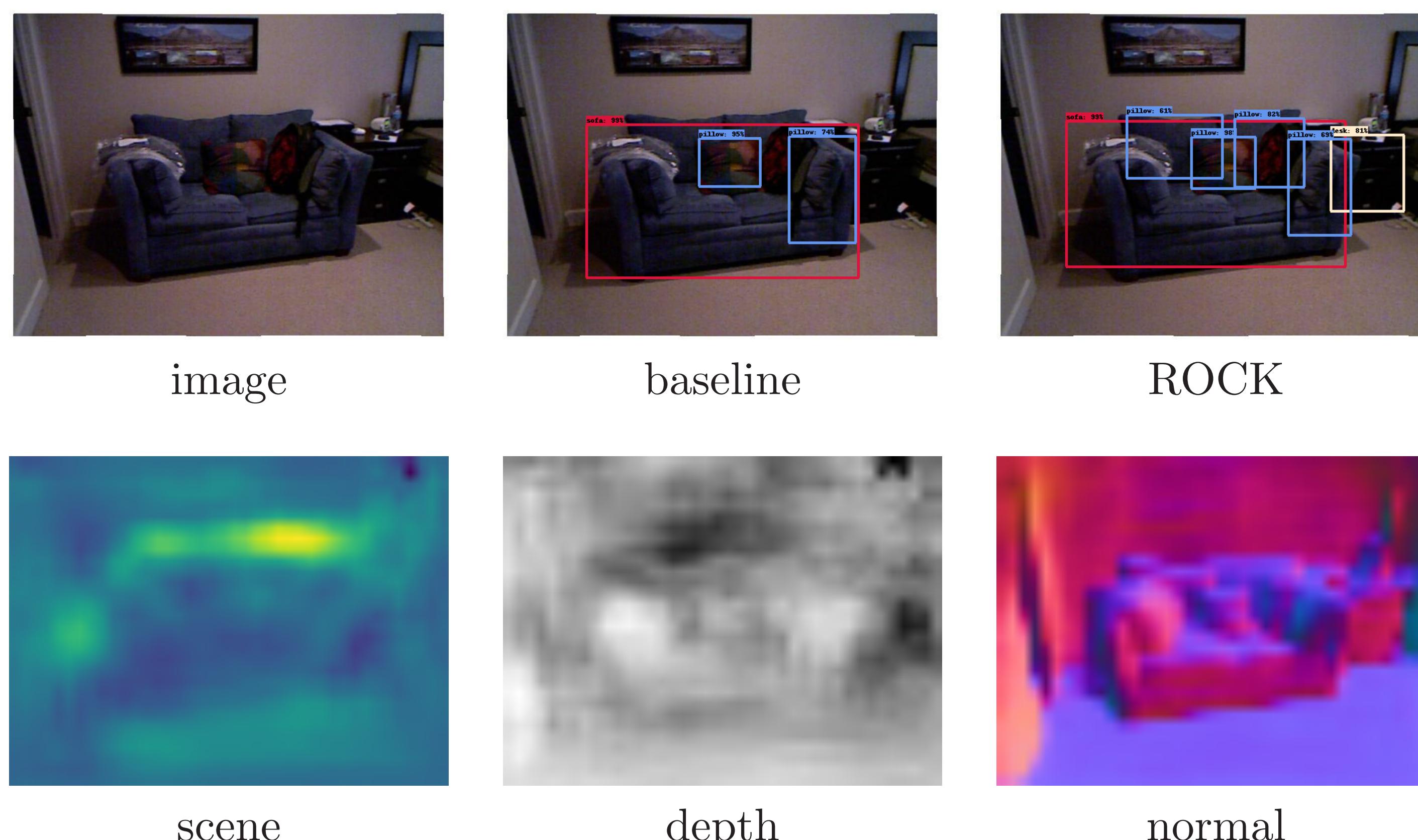
### Generic Block

- Easily inserted into any model
- Encode any task
- Low overhead (w.r.t. SSD [6]): 17% params, 7% time

### Perspectives

- Domain Adaptation
- Test-time auxiliary supervision

## VISUALIZATIONS OF PREDICTIONS



## ANALYSIS AND RESULTS

- Experiments on NYUv2 dataset [7]

### Ablation Study

Model	mAP@ 0.5	mAP@ 0.75	mAP@ [0.5:0.95]
SSD [6]	31.2	15.8	16.2
+ flat MTL	34.3	16.0	17.4
+ intensive pool.	35.7	16.2	17.4
+ fusion	37.6	17.1	18.5

### Multi-Modal Object Detection

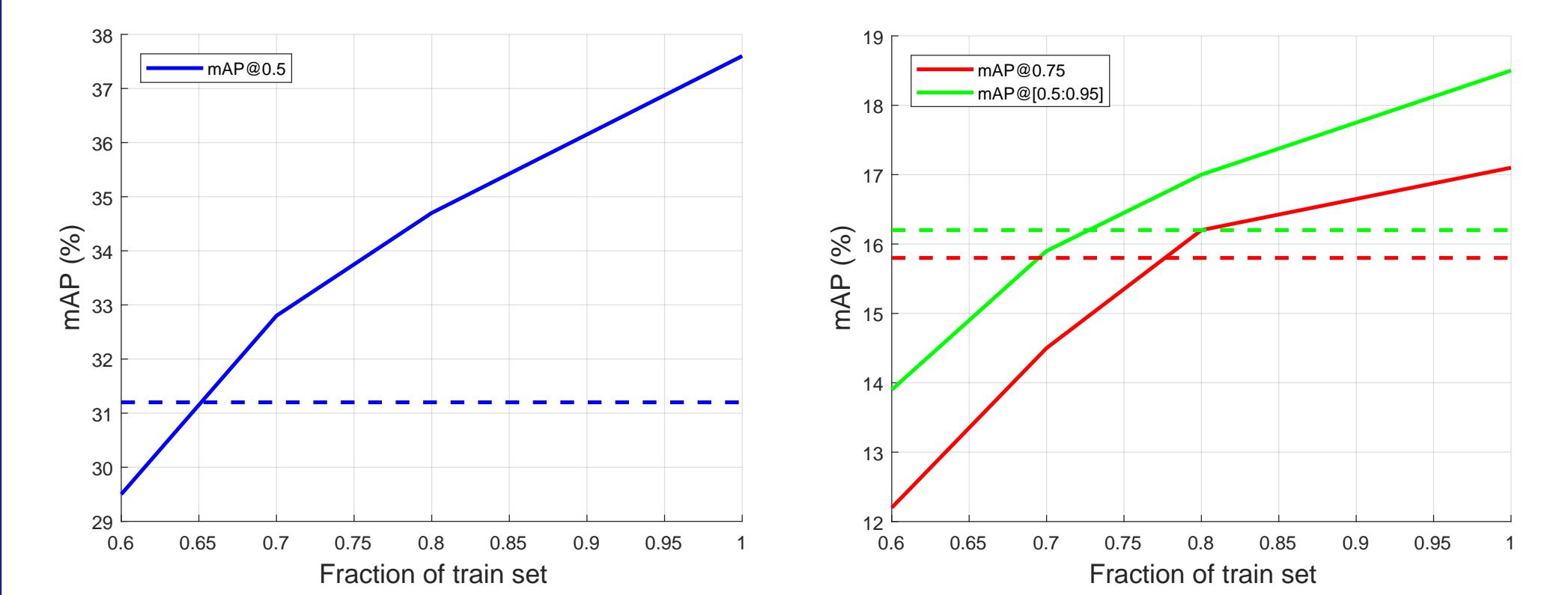
- Auxiliary information: depth only

Model	Aux.	Sup.	mAP
Modality Hallucination [5]	D		34.0
ROCK	D		37.1
RGB-D R-CNN [8]		D*	35.5
RGB-D R-CNN [9]		D* + SYN	37.3

- Auxiliary information: all available

Model	Aux.	Sup.	mAP
Pose CNN [10]	D*	N* + SYN	38.8
RGB-Geo R-CNN [8]	D*	N* C*	39.3
ROCK	D	N S	39.8
ROCK	D	N S + SYN	46.8

### Effectiveness of Additional Supervision



- Auxiliary supervision  $\sim 30\%$  images here

## REFERENCES

- [1] Azizpour *et al.*, Factors of Transferability, *TPAMI* (2016)
- [2] Caruana, Multitask Learning, *Machine Learning* (1997)
- [3] Kokkinos, UberNet, *CVPR* (2017)
- [4] Vapnik and Vashist, LUPI, *Neural Networks* (2009)
- [5] Hoffman *et al.*, Modality Hallucination, *CVPR* (2016)
- [6] Liu *et al.*, SSD, *ECCV* (2016)
- [7] Silberman *et al.*, NYUv2, *ECCV* (2012)
- [8] Wang and Siddiqi, Differential Geometry, *CVPR* (2016)
- [9] Gupta *et al.*, RGB-D Features, *ECCV* (2014)
- [10] Gupta *et al.*, RGB-D Pose, *arXiv* (2015)