南京理工大学
NANJING UNIVERSITY OF SCIENCE & TECHNOLOGY

ICPR 2020
25th INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION
Milan, Italy 10 | 15 January 2021

Learning a Dynamic High-Resolution Network for Multi-Scale Pedestrian Detection

Mengyuan Ding, Shanshan Zhang, Jian Yang

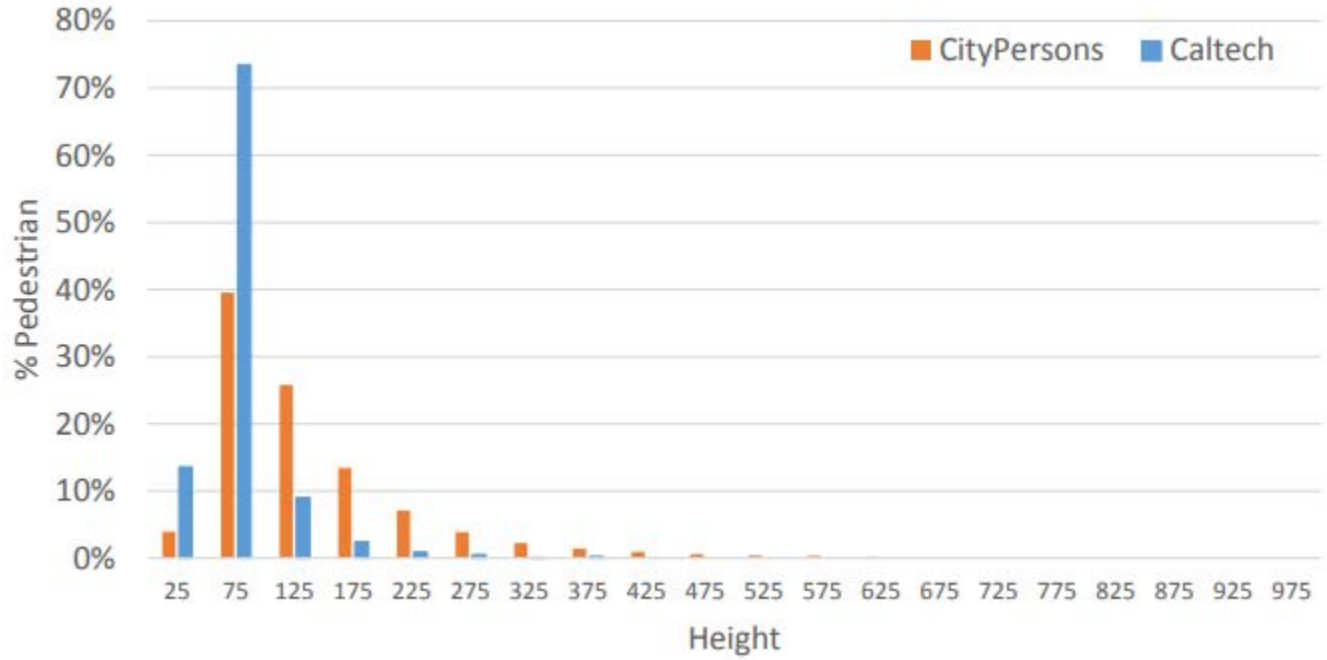
School Of Computer Science And Engineering,
Nanjing University Of Science And Technology, China

Contents

- Motivation
- The Proposed DHRNet
- Experiments

Backbone	MR^{-2}	MR_s^{-2}	MR_m^{-2}	MR_l^{-2}
ResNet50	14.79	21.76	7.11	**7.88**

Scale Variation In Pedestrian Detection

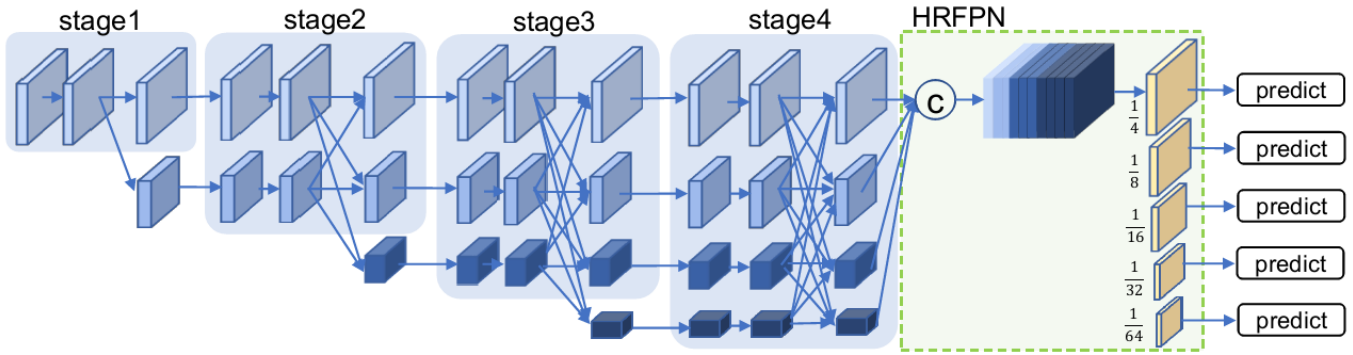


Height distributions of CityPersons and Caltech[1]

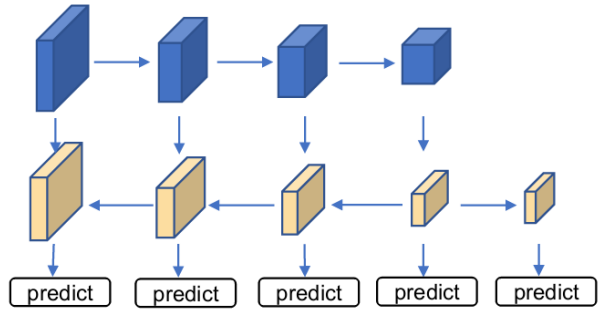
[1] S. Zhang, et al. “Citypersons: A diverse dataset for pedestrian detection.” CVPR, 2017

High-Resolution Network(HRNet)[1]

Backbone	MR^{-2}	MR_s^{-2}	MR_m^{-2}	MR_l^{-2}
ResNet50	14.79	21.76	7.11	**7.88**
HRNetV2p-W18	13.59	**16.41**	5.58	8.13
HRNetV2p-W32	**13.54**	18.51	**5.33**	7.94



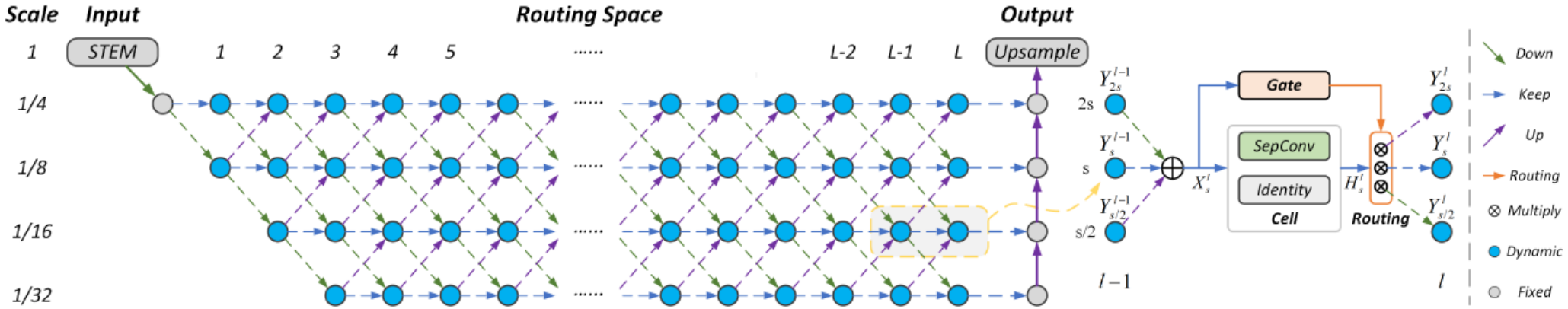
A simple example of HRNets



A simple example of Traditional CNNs

[1] K. Sun, et al. “High-resolution representations for labeling pixels and regions,” arXiv preprint arXiv:1904.04514, 2019.

Dynamic architectures



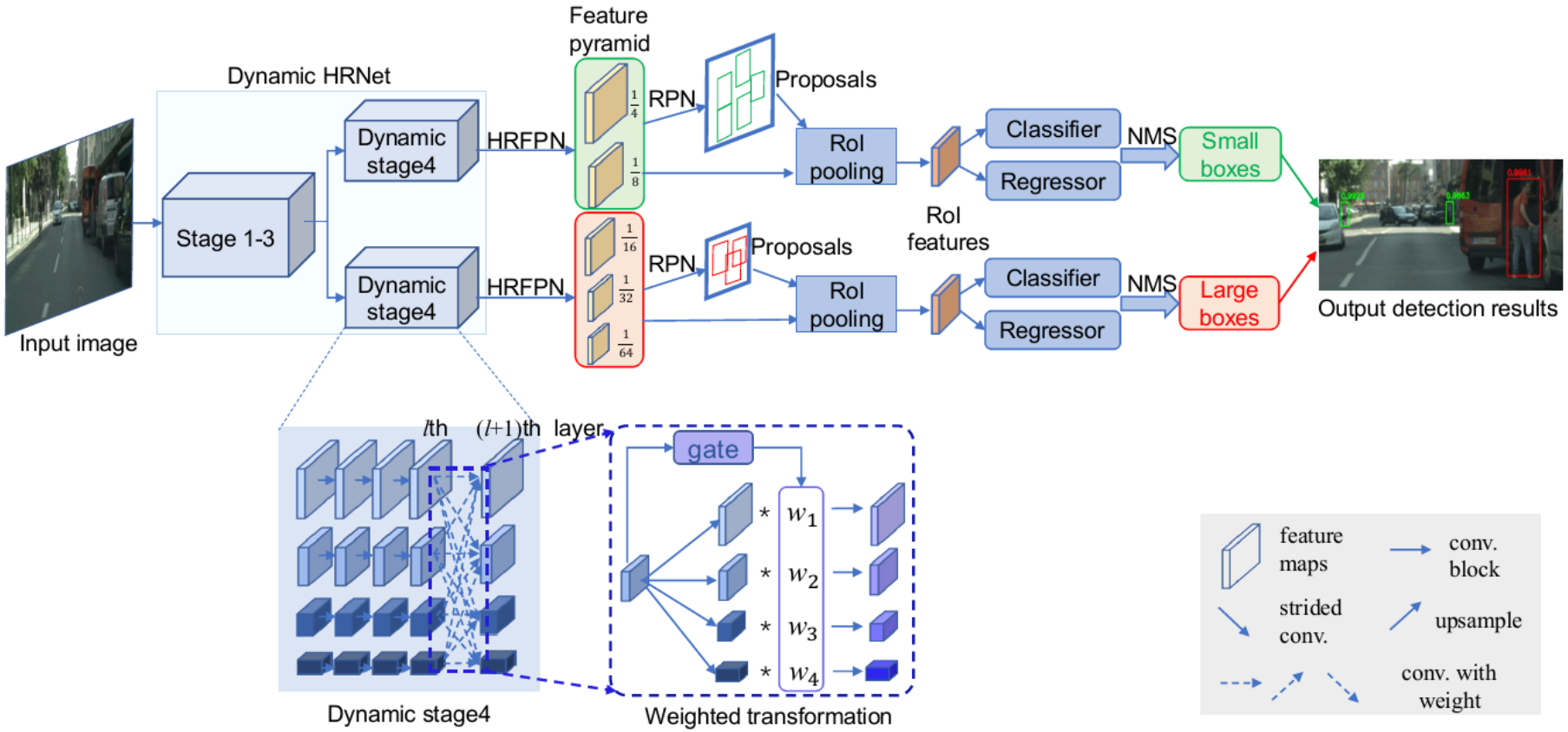
Self Attention — the soft conditional gate[1]

[1] Y. Li, et al. “Learning dynamic routing for semantic segmentation,” CVPR, 2020.

Contents

- Motivation
- The Proposed DHRNet
- Experiments

Faster RCNN with our proposed Dynamic HRNet (DHRNet)



Flow chart of our detector based on Faster RCNN and the proposed Dynamic HRNet (DHRNet)

Contents

- Motivation
- The Proposed DHRNet
- Experiments

Experiments

- Ablation studies

  - Stage for the soft gate module to add in

Stage	MR^{-2}	MR_s^{-2}	MR_m^{-2}	MR_l^{-2}
2	11.08	13.67	3.40	6.62
3	10.58	13.83	3.46	6.53
4	**10.40**	**13.43**	**2.69**	**6.21**
baseline	12.18	15.10	4.87	6.62

  - Number of branches

Branches	MR^{-2}	MR_s^{-2}	MR_m^{-2}	MR_l^{-2}
1	11.28	13.79	3.50	6.51
2	**10.40**	**13.43**	**2.69**	**6.21**
3	11.12	13.35	3.85	6.49
baseline	12.18	15.10	4.87	6.62

  - Activation function

Activation	MR^{-2}	MR_s^{-2}	MR_m^{-2}	MR_l^{-2}
Softmax	11.45	14.76	3.38	6.70
Max(0,Tanh)	10.43	**12.55**	3.28	6.46
Sigmoid	**10.40**	13.43	**2.69**	**6.21**

- Results

Method	Backbone	MR^{-2}	MR_s^{-2}	MR_m^{-2}	MR_l^{-2}
Adapted-FRCNN[9]	VGG-16	15.4	25.6	7.2	7.9
Adapted-FRCNN+Seg[9]	VGG-16	14.8	22.6	6.7	8.0
OR-CNN[29]	VGG-16	12.8	-	-	-
TLL+MRF[30]	ResNet-50	14.4	-	-	-
ALFNet[19]	ResNet-50	12.0	19.0	5.7	6.6
CSP(with offset)[20]	ResNet-50	**11.0**	16.0	**3.7**	**6.5**
FRCNN+FPN*	ResNet-50	12.77	19.57	5.45	7.03
FRCNN+HRFPN*	HRNet-W18	12.18	**15.10**	4.87	6.62
FRCNN+HRFPN	DHRNet-W18(ours)	**10.40**	**13.43**	**2.69**	**6.21**

- Visualization

HRNet DHRNet

Visualization of detection results on CityPersons validation set
(FPPI=0.1; ground truth: green; detection results: red)

ICPR 2020
25th INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION
Milan, Italy 10 | 15 January 2021

THANK YOU !