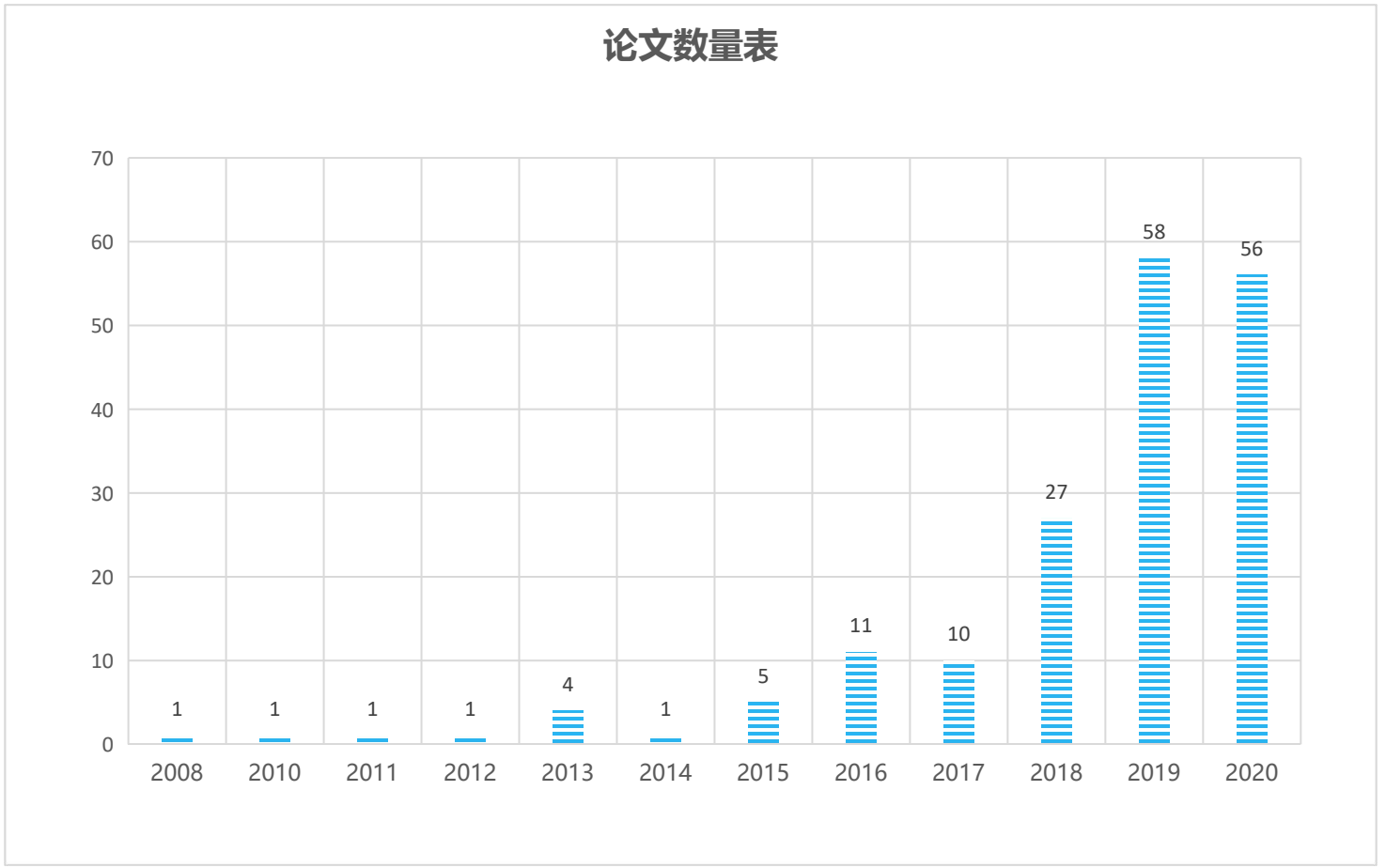
華東師範大學
EAST CHINA NORMAL UNIVERSITY

人群计数概述

汇 报 人：吴 兴 蛟

人群计数（Crowd counting ）



人群计数（Crowd counting ）

问题3

人群计数（Crowd counting ）

aims： count the number of people in **a crowded scene**

Applications

- **Safety monitoring**
- **Disaster management**
- **Design of public spaces**
- **Intelligence gathering and analysis**
- **Virtual environments**
- **Forensic search**

人群计数（Crowd counting ）

$$\mathrm{MAE}=\frac{1}{\mathrm{N}}\sum_{i=1}^{N}\left|y_i-y_i^{'}\right|$$

平均绝对误差（Mean Absolute Deviation）

$$\mathrm{MSE}=\frac{1}{\mathrm{N}}\sum_{i=1}^{N}\sqrt{\left|y_i-y_i^{'}\right|^2}$$

均方误差（Mean Square Error）

数据集

数据集

Dataset	Year	No.ofimages	Resolution	Min	Ave	Max	Total count
UCSD	2008	2000	158x238	11	25	46	49885
Mall	2012	2000	320x240	13	-	53	62325
UCF_CC_50	2013	50	Varied	94	1279	4543	63974
WorldExpo'10	2015	3980	576x720	1	50	253	199923
ShanghaiTech Part A	2016	482	Varied	33	501	3139	241677
ShanghaiTech Part B	2016	716	768x1024	9	123	578	88488
UCF-QNRF	2018	1535	2013x2902	-	815	12865	1251642

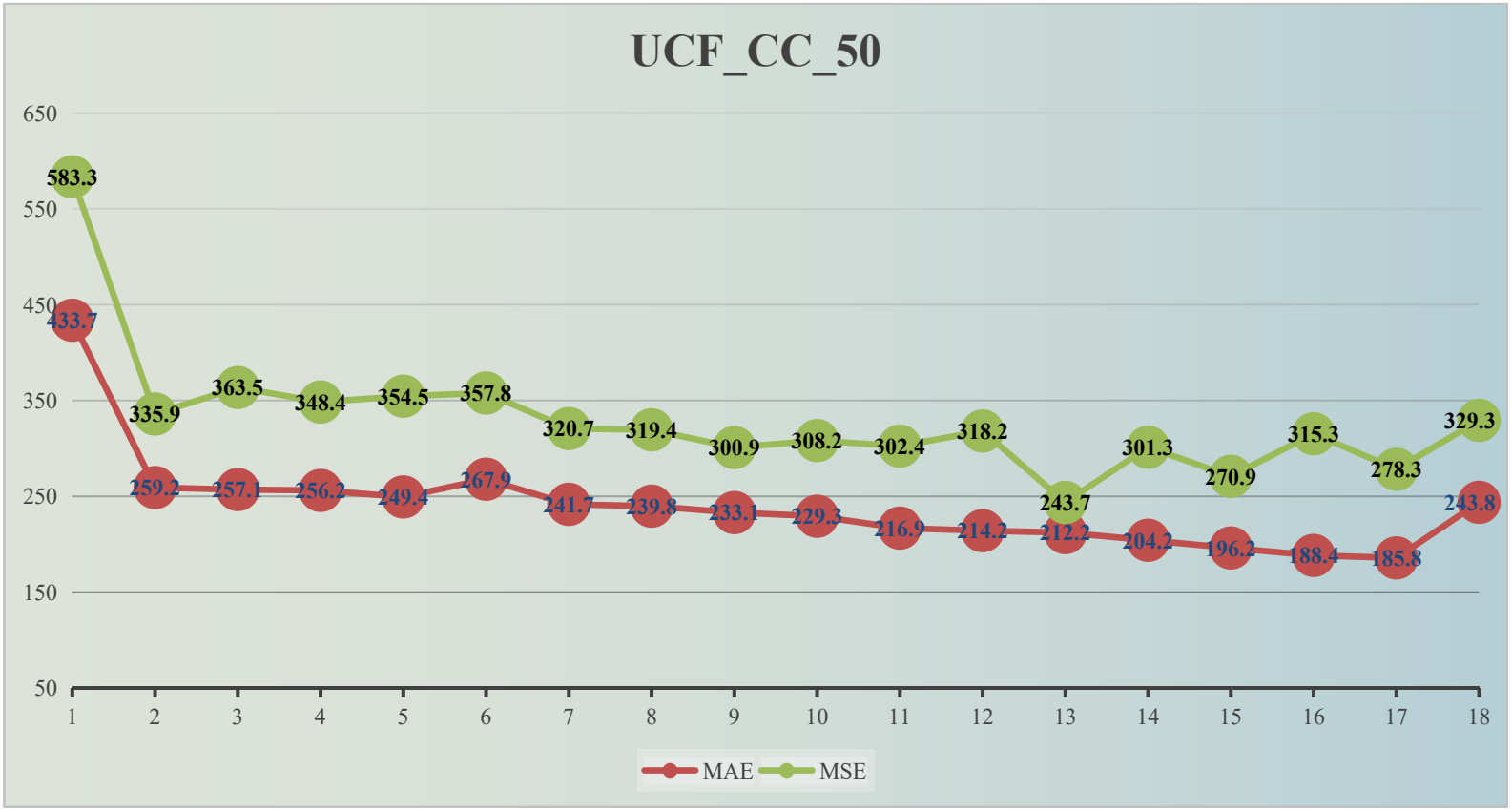
UCSD

Year-Conference/Journal	Method	MAE	MSE
2015--CVPR	Zhang 2015	1.6	3.31
2016--ECCV	Hydra-CNN	1.65	-
2016--ECCV	CNN-Boosting	1.1	-
2016--CVPR	MCNN	1.07	1.35
2017--ICCV	ConvLSTM-nt	1.73	3.52
2017--CVPR	Switching CNN	1.62	2.1
2017--ICCV	ConvLSTM	1.3	1.79
2017--ICCV	Bidirectional ConvLSTM	1.13	1.43
2018--CVPR	CSRNet	1.16	1.47
2018--CVPR	ACSCP	1.04	1.35
2018--ECCV	SANet	1.02	1.29
2018--TIP	BSAD	1	1.4
2019--WACV	SPN	1.03	1.32
2019--ICCV	SPANet+SANet	1	1.28
2019--CVPR	ADCrowdNet(DME)	0.98	1.25
2019--BMVC	E3D	0.93	1.17
2019--CVPR	PACNN	0.89	1.18
2019--TIP	PaDNet	0.85	1.06

MALL

Year-Conference/Journal	Method	MAE	MSE
2012--BMVC	Chen 2012	3.15	15.7
2016--ECCV	CNN-Boosting	2.01	-
2017--ICCV	ConvLSTM-nt	2.53	11.2
2017--ICCV	ConvLSTM	2.24	8.5
2017--ICCV	Bidirectional ConvLSTM	2.1	7.6
2018--CVPR	DecideNet	1.52	1.9
2018--IJCAI	DRSAN	1.72	2.1
2019--BMVC	E3D	1.64	2.13
2019--WACV	SAAN	1.28	1.68

UCF_CC_50

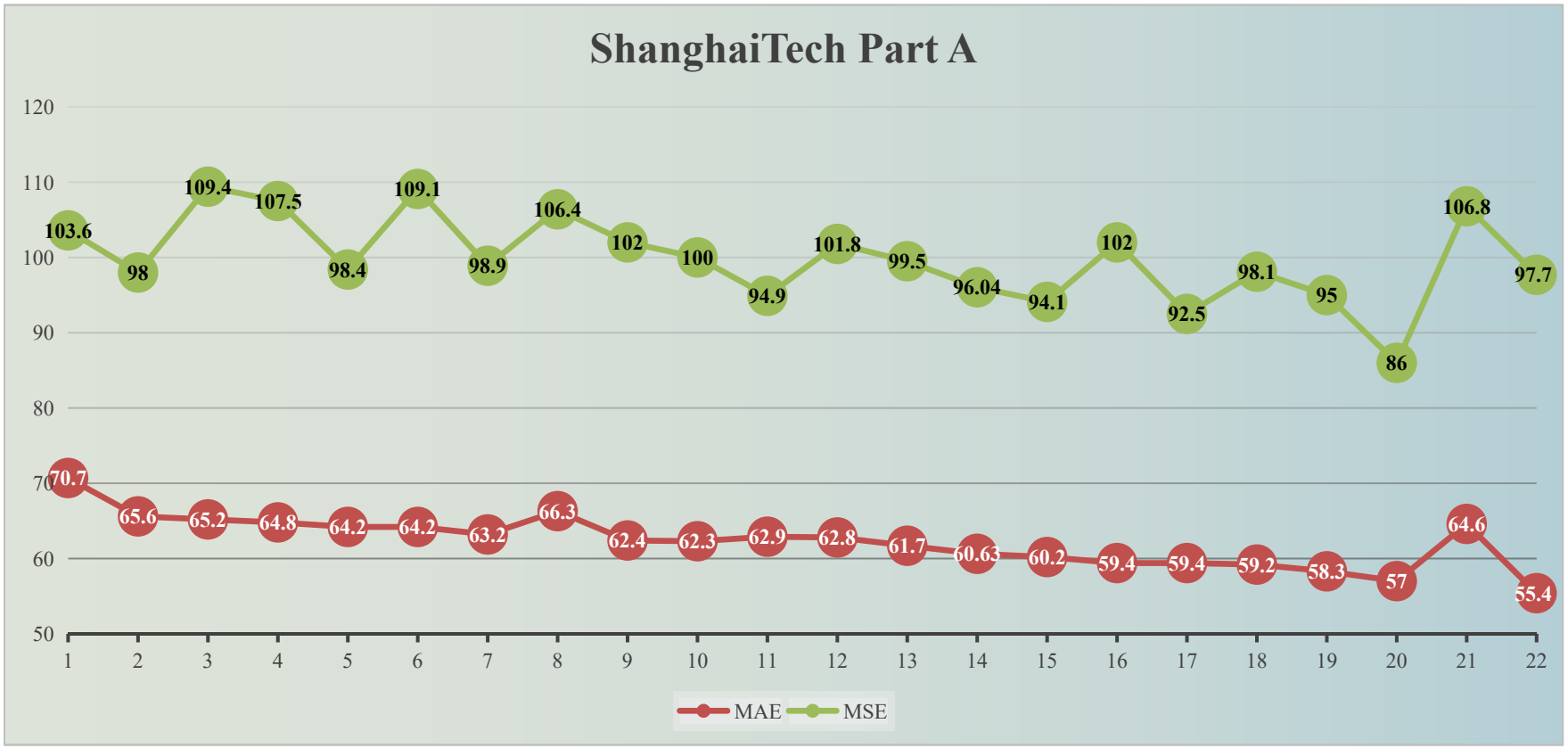


1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
GWTA-CCNN(2019AAAI)	SPN(2019WACV)	ADCrowdNet(DME)(2019CVPR)	HA-CCN(2019TIP)	TEDnet(2019CVPR)	PACNN(2019CVPR)	PACNN+CSRNet(2019CVPR)	RANet(2019ICCV)	MBTTBF-SCFB(2019ICCV)	BL(2019ICCV)	DSSINet(2019ICCV)	SFCN(2019CVPR)	CAN(2019CVPR)	S-DCNet(2019ICCV)	ASD(2019ICASSP)	SPN+L2SM(2019ICCV)	PaDNet(2019TIP)	DUBNet(2020AAAI)

WorldExpo'10

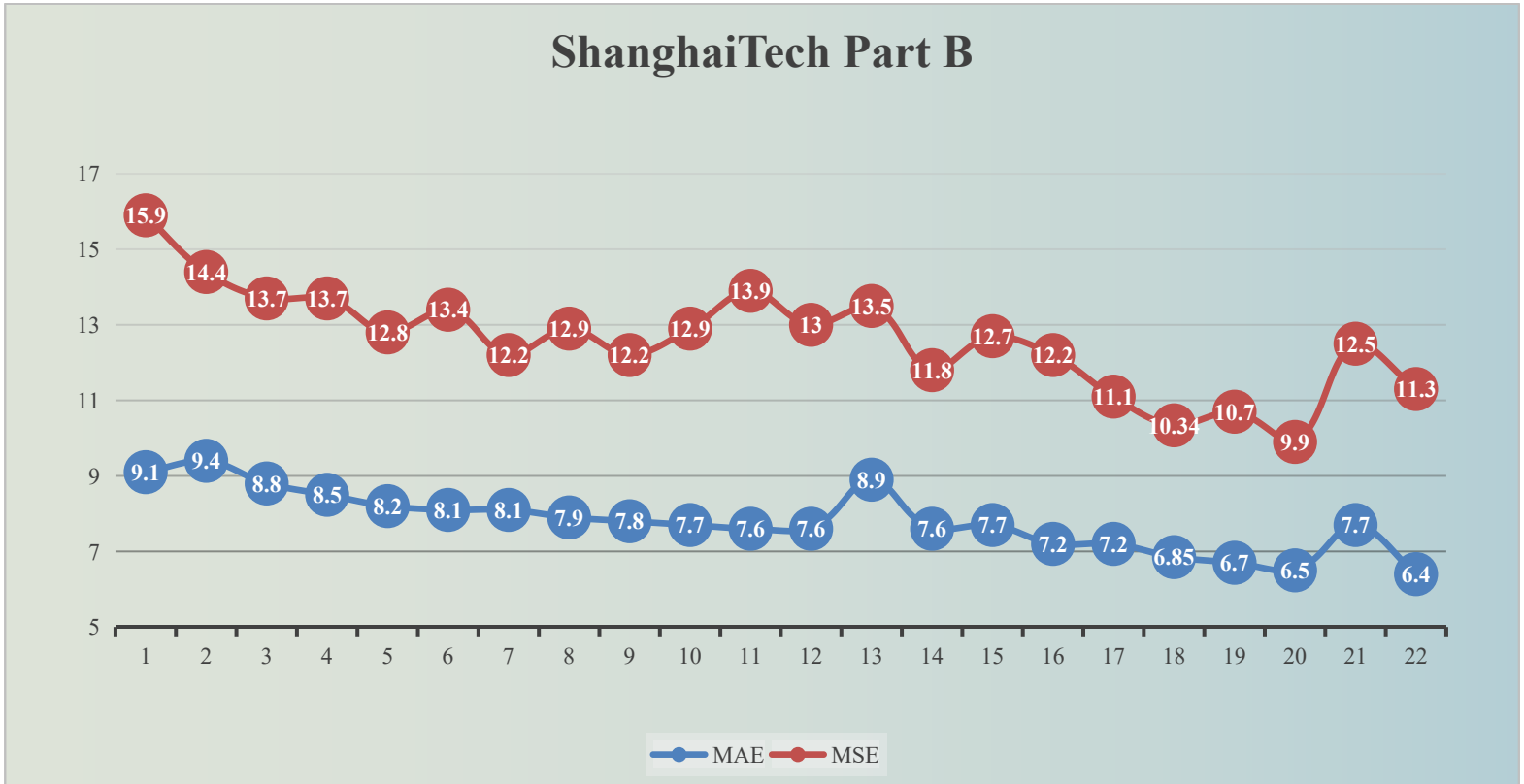
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
TDF-CNN(2018AAAI)	IG-CNN(2018CVPR)	BSAD (2018TIP)	ic-CNN(2018ECCV)	DecideNet(2018CVPR)	D-ConvNet-v1(2018CVPR)	CSRNet (2018CVPR)	SaCNN (2018WACV)	SANet (2018ECCV)	DRSAN (2018IJCAI)	ACSCP (2018CVPR)	PGCNet(2019ICCV)	TEDnet (2019CVPR)	PACNN(2019CVPR)	ADCrowdNet(2019CVPR)	CAN(2019CVPR)	CAN(ECAN)(2019CVPR)	DSSINet(2019ICCV)

ShanghaiTech Part A



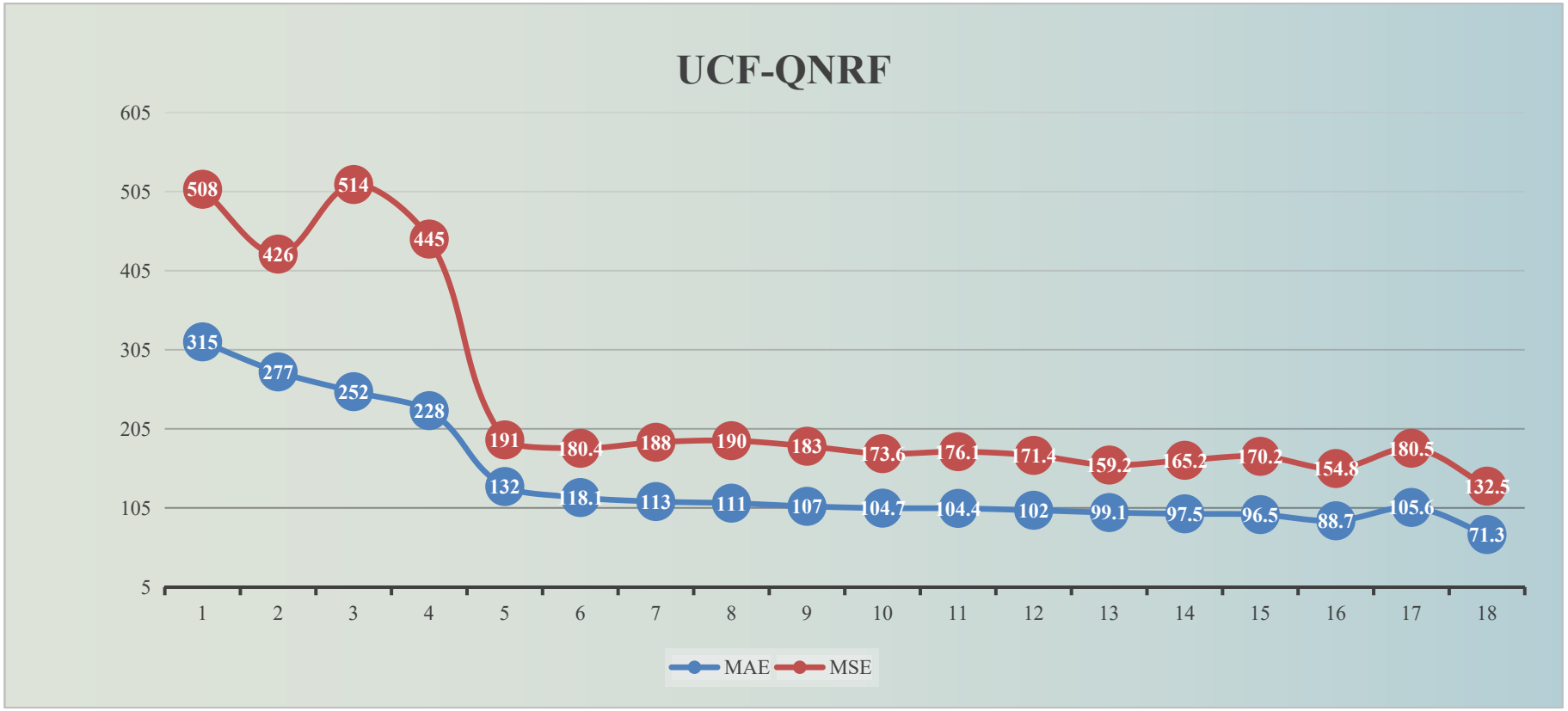
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
GSP(2019CVPRW)	ASD(2019ICASSP)	CFF(2019ICCV)	SFCN (2019 CVPR)	SPN+L2SM(2019ICCV)	TEDnet(2019CVPR)	ADCrowdNet (2019CVPR)	PACNN(2019CVPR)	PACNN+CSRNet(2019CVPR)	CAN (2019CVPR)	HA-CCN (2019 TIP)	BL(2019ICCV)	SPN (2019WACV)	DSSINet(2019ICCV)	MBTT BF-SCFB (2019ICCV)	RANet(2019ICCV)	SPANet+SANet(2019ICCV)	PaDNet(2019TIP)	S-DCNet(2019ICCV)	PGCNet(2019ICCV)	DUBNet(2020AAAI)	ADSCNet(2020CVPR)

ShanghaiTech Part B



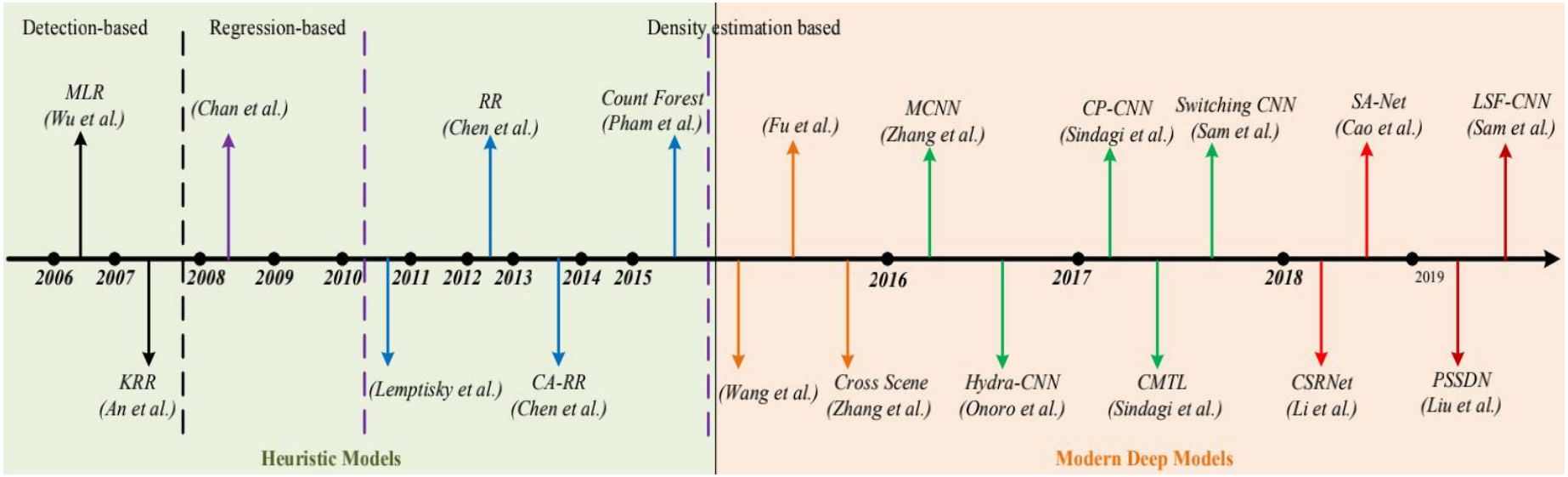
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
GSP(2019CVPRW)	SPN(2019WACV)	PGCNet(2019ICCV)	ASD(2019ICASSP)	TEDnet(2019CVPR)	HA-CCN(2019TIP)	PaDNet(2019TIP)	RANet(2019ICCV)	CAN(2019CVPR)	ADCrowdNet(2019CVPR)	ADCrowdNet(2019CVPR)	SFCN(2019CVPR)	PACNN(2019CVPR)	PACNN+CSRNet(2019CVPR)	BL(2019ICCV)	CFF(2019ICCV)	SPN+L2SM(2019ICCV)	DSSINet(2019ICCV)	S-DCNet(2019ICCV)	SPANet+SANet(2019ICCV)	DUBNet(2020AAAI)	ADSCNet(2020CVPR)

UCF-QNRF



1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Idrees 2013CL(2013CVPR)	MCNNCL(2016CVPR)	CMTLCL(2017AVSS)	Switching CNNCL(2017CVPR)	CL(2018ECCV)	HA-CCN(2019TIP)	TEDnet (2019CVPR)	RANet (2019ICCV)	CAN(2019CVPR)	SPN+L2SM(2019ICCV)	S-DCNet (2019ICCV)	SFCN (2019 CVPR)	DSSINet(2019ICCV)	MBTTBF-SCFB(2019ICCV)	PaDNet(2019 TIP)	BL(2019ICCV)	DUBNet(2020 AAAI)	ADSCNet(2020CVPR)

人群计数（Crowd counting ）



人群计数（Crowd counting ）

Detection-based approaches

Regression-based approaches

Density estimation-based approaches

sliding window detector

parts-based and shape-based detectors

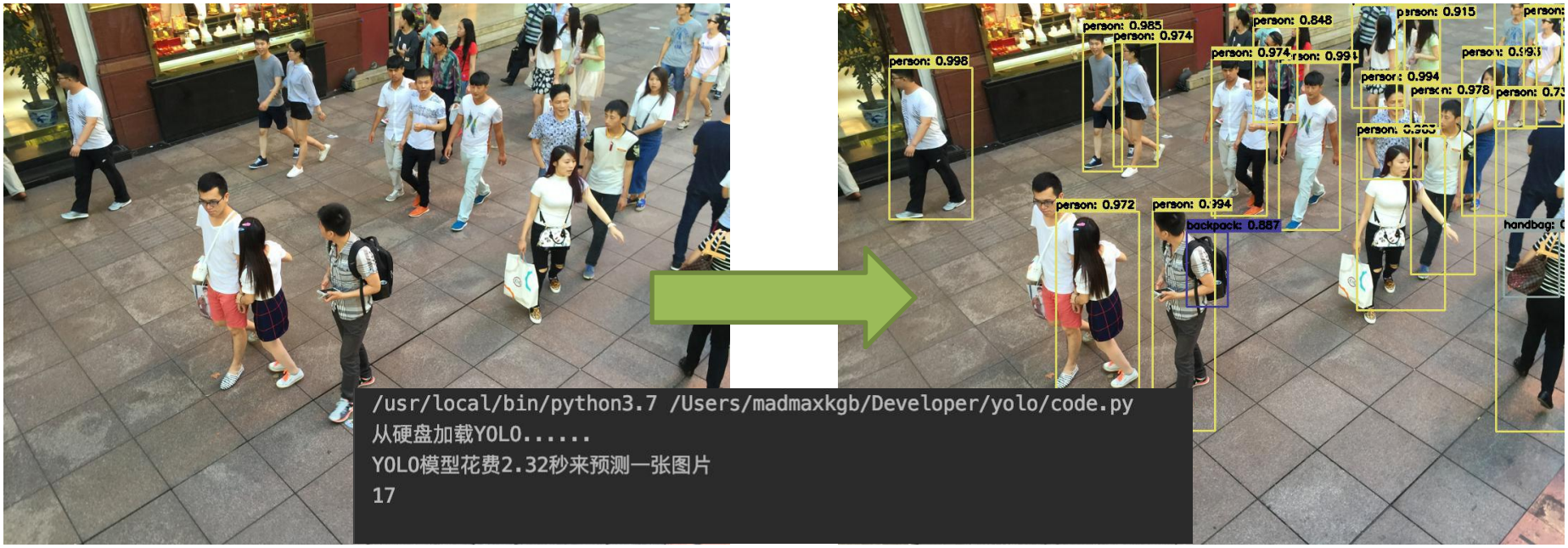
ignored important spatial information

人群计数（Crowd counting ）

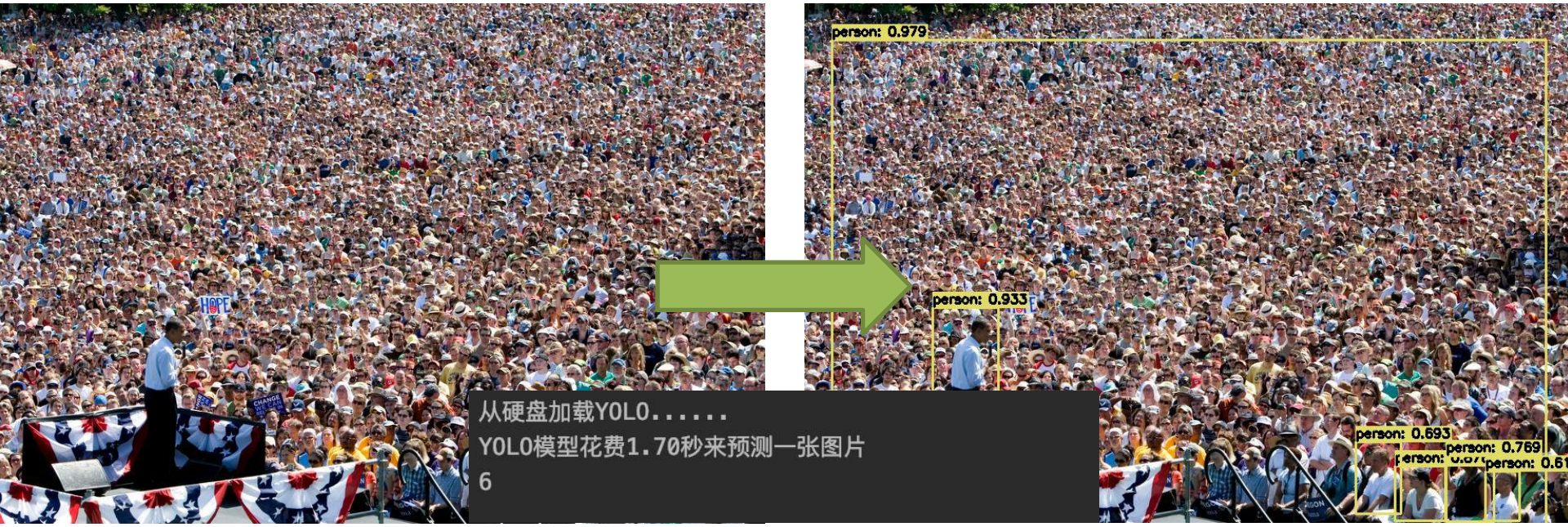
Detection-based approaches

- 利用不同尺寸的滑动窗口框住图中的某一部分作为候选区域；
- 提取候选区域相关的视觉特征。比如人脸检测常用的Harr特征；行人检测和普通目标检测常用的HOG特征等；
- 利用分类器进行识别，比如常用的SVM模型。

基于检测的方法


person: 0.998
person: 0.985
person: 0.974
person: 0.848
person: 0.974
person: 0.915
person: 0.994
person: 0.978
person: 0.972
backpack: 0.887
/usr/local/bin/python3.7 /Users/madmaxkgb/Developer/yolo/code.py
从硬盘加载YOLO.....
YOLO模型花费2.32秒来预测一张图片
17

基于检测的方法

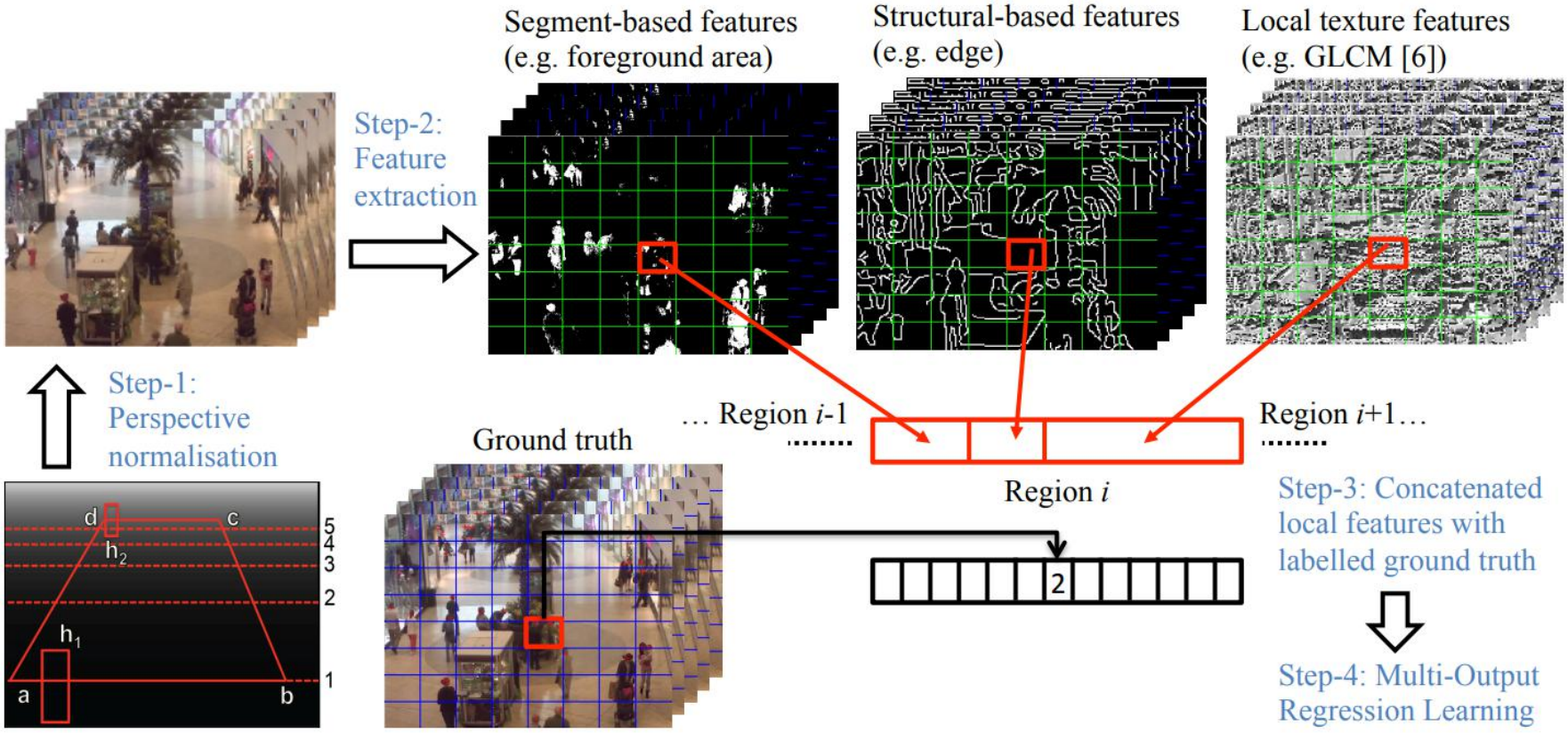

person: 0.979
person: 0.933
person: 0.693
person: 0.769
从硬盘加载YOLO.....
YOLO模型花费1.70秒来预测一张图片
6

人群计数（Crowd counting ）

Regression-based approaches

- 把输入图片缩放到**固定大小**；
- 使用卷积网络；
- 对模型置信度设置阈值，得到目标位置与类别。

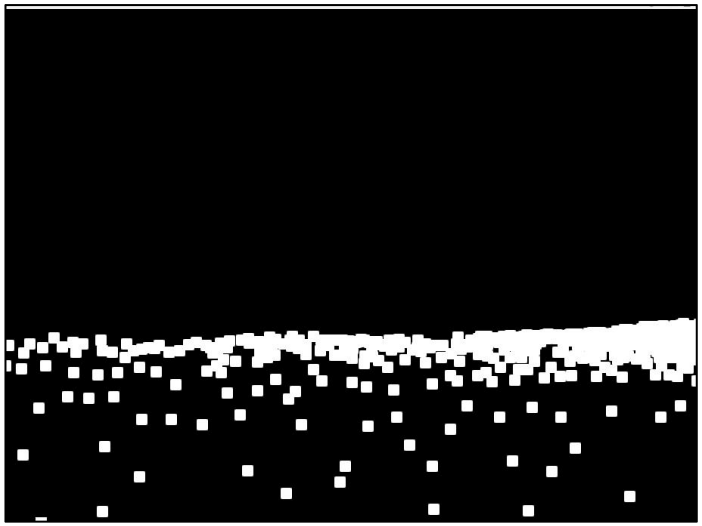
人群计数（Crowd counting ）



Chen, Ke, et al. "Feature mining for localised crowd counting." BMVC. Vol. 1. No. 2. 2012.

人群计数（Crowd counting ）

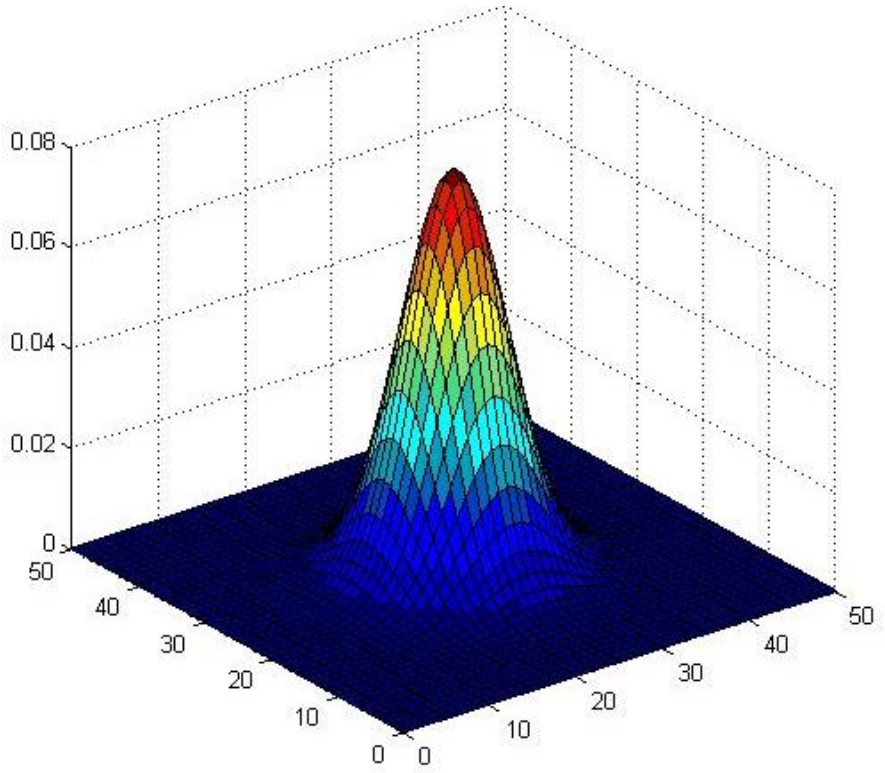


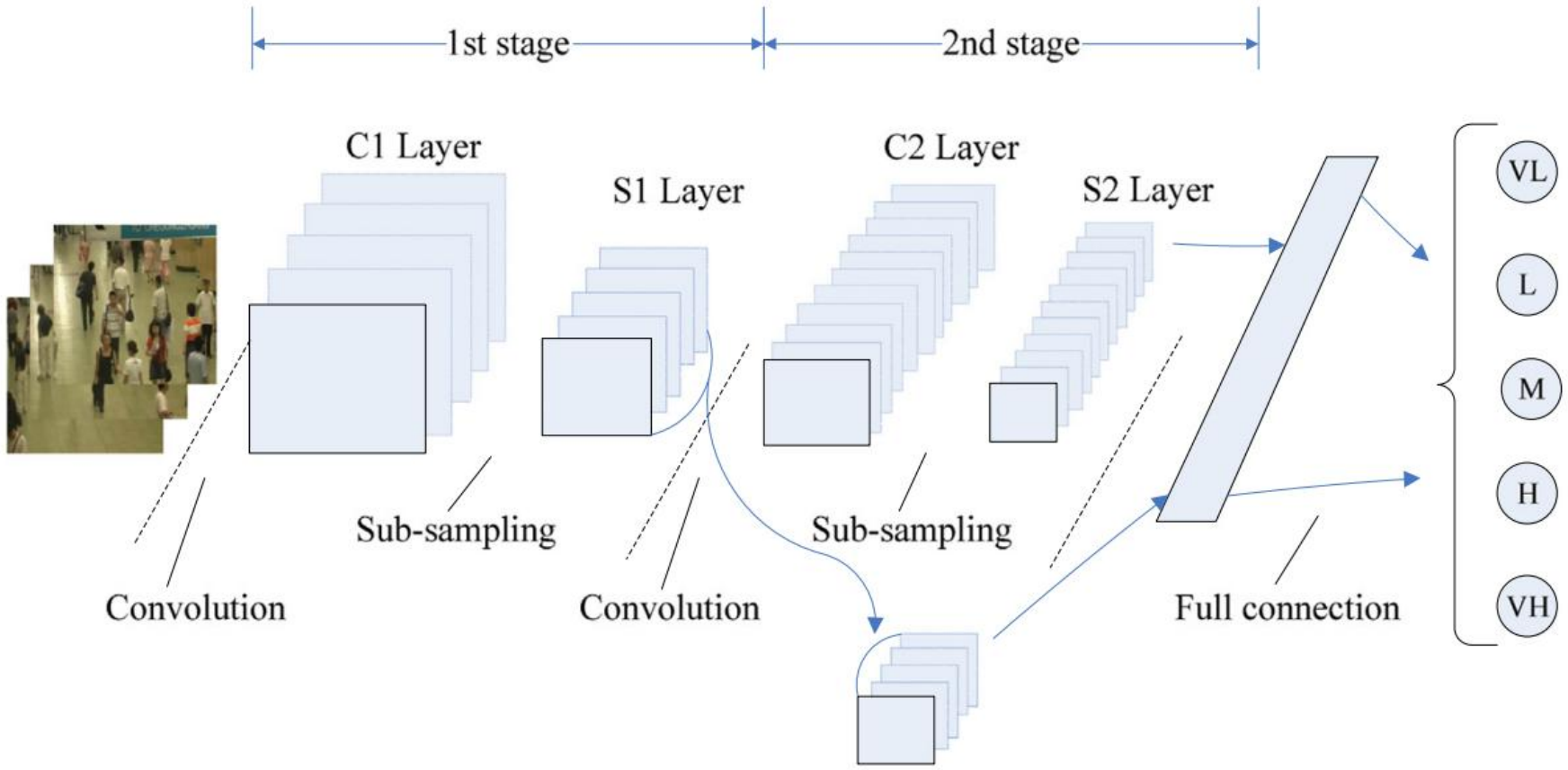
人群计数（Crowd counting ）

H = fspecial('Gaussian',[f_sz, f_sz],sigma);

$$G(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}$$

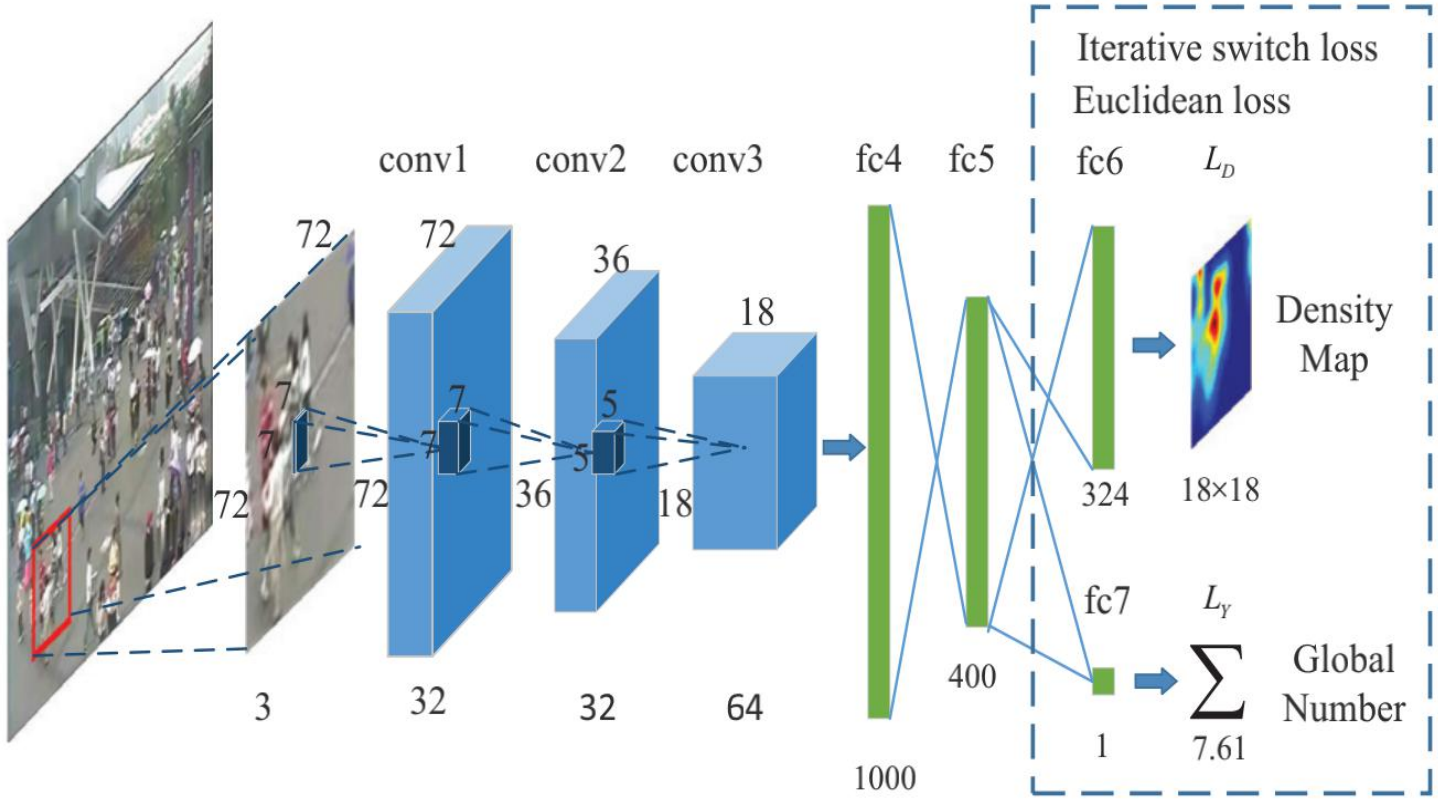


基于CNN



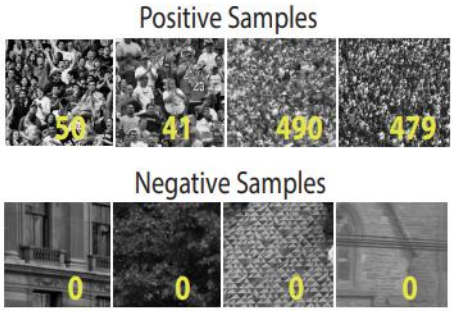
Fu, Min, et al. "Fast crowd density estimation with convolutional neural networks." Engineering Applications of Artificial Intelligence 43 (2015): 81-88.

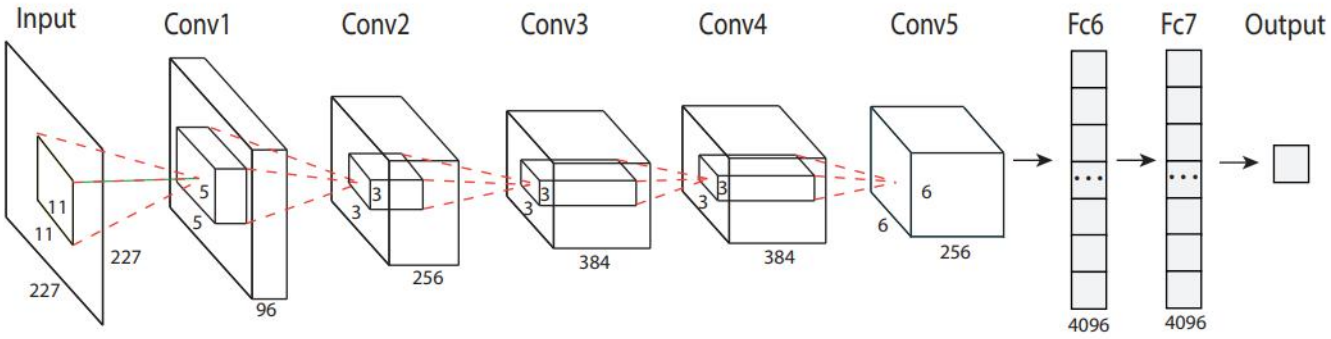
基于CNN



Zhang, Cong, et al. "Cross-scene crowd counting via deep convolutional neural networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.

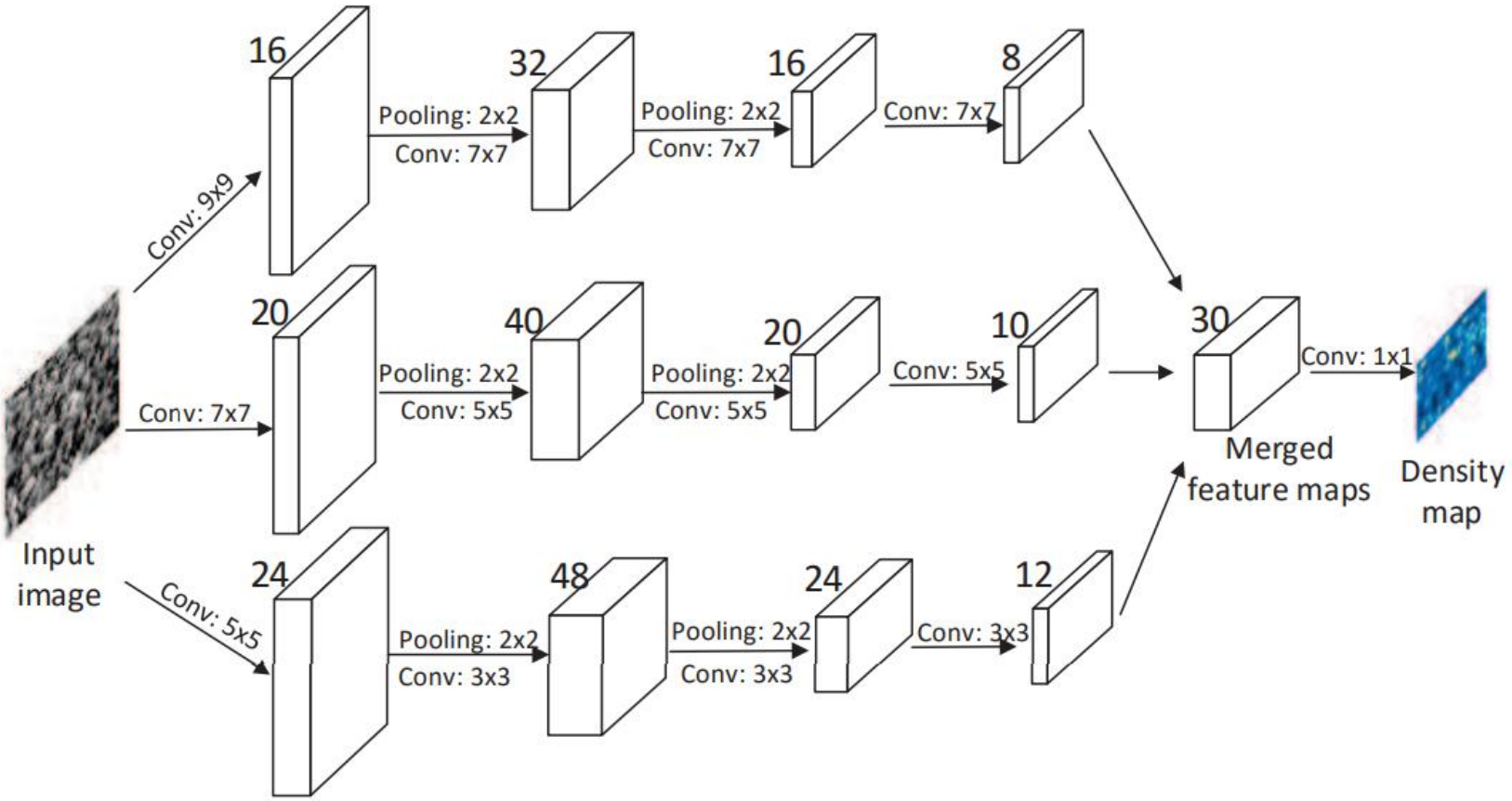
基于CNN





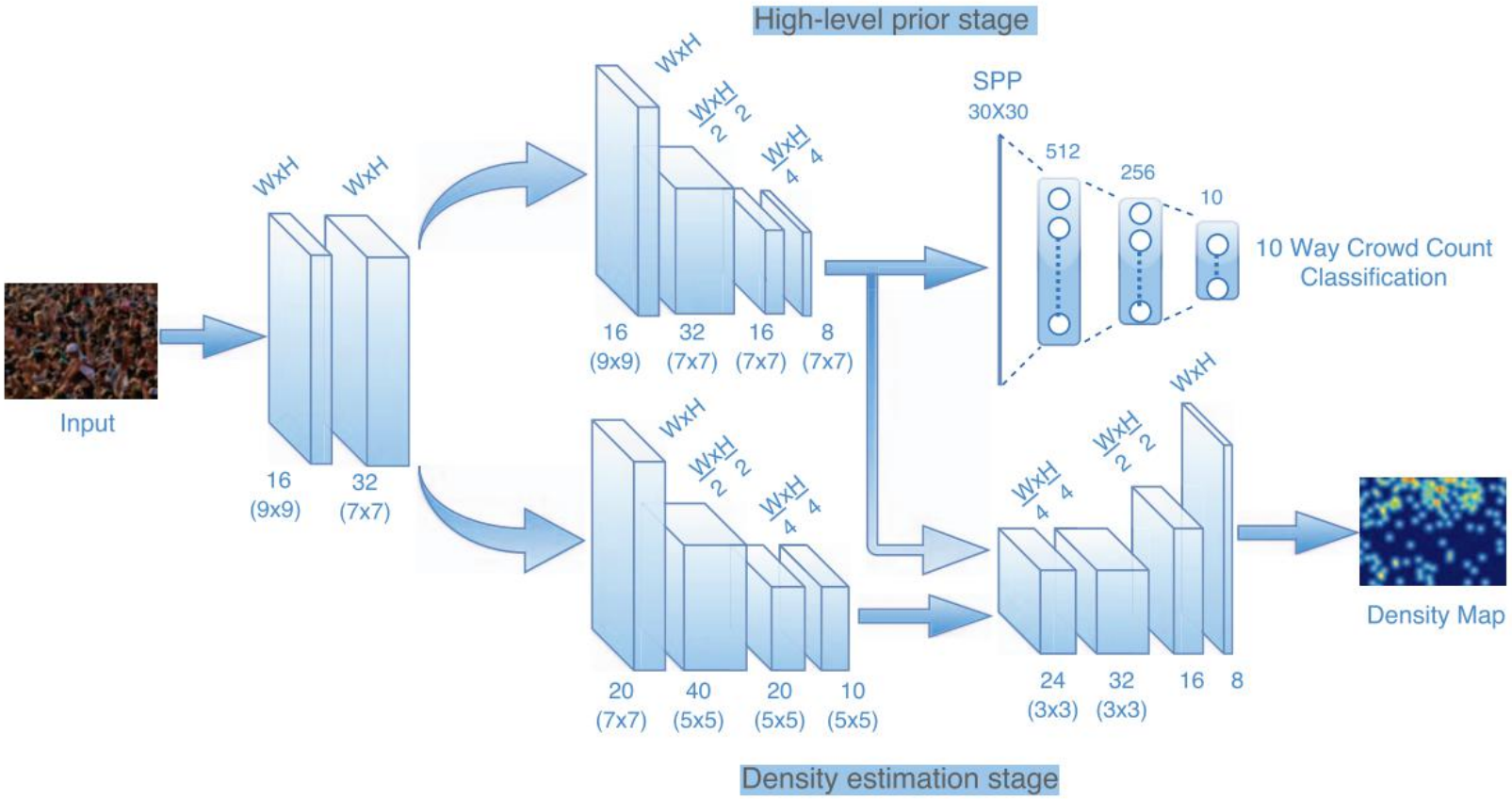
Wang, Chuan, et al. "Deep people counting in extremely dense crowds." Proceedings of the 23rd ACM international conference on Multimedia. 2015.

基于CNN



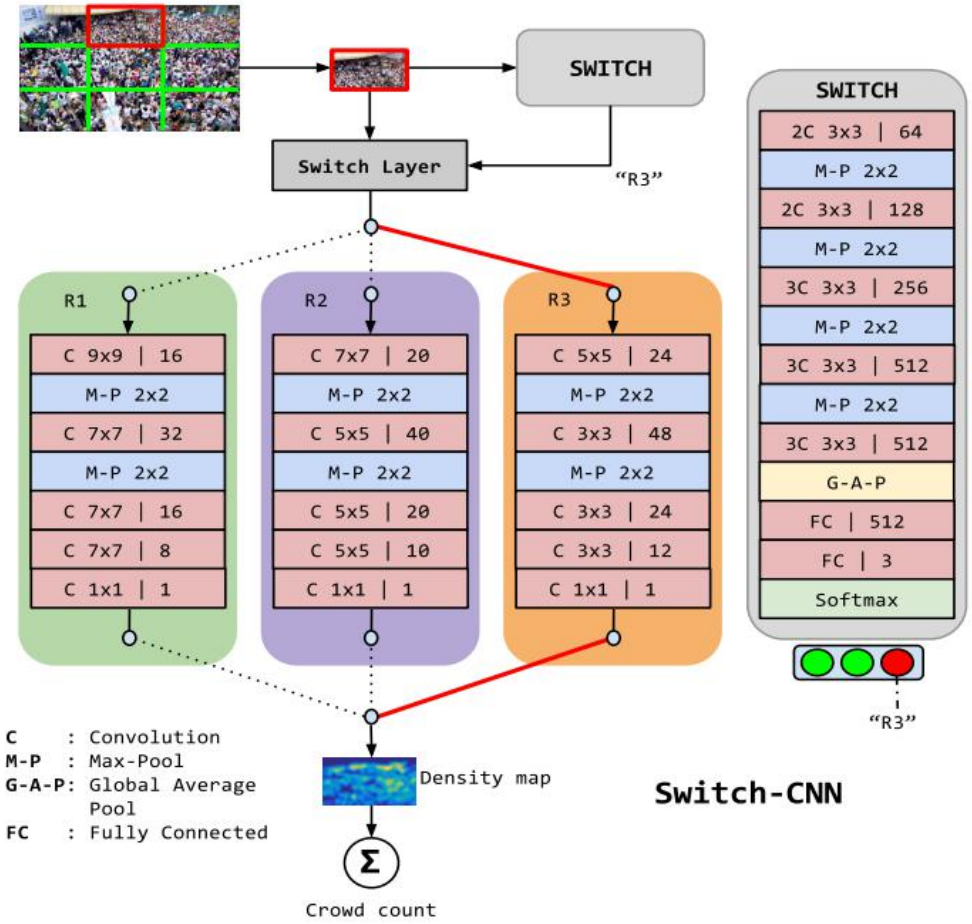
Zhang, Yingying, et al. "Single-image crowd counting via multi-column convolutional neural network." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

基于CNN



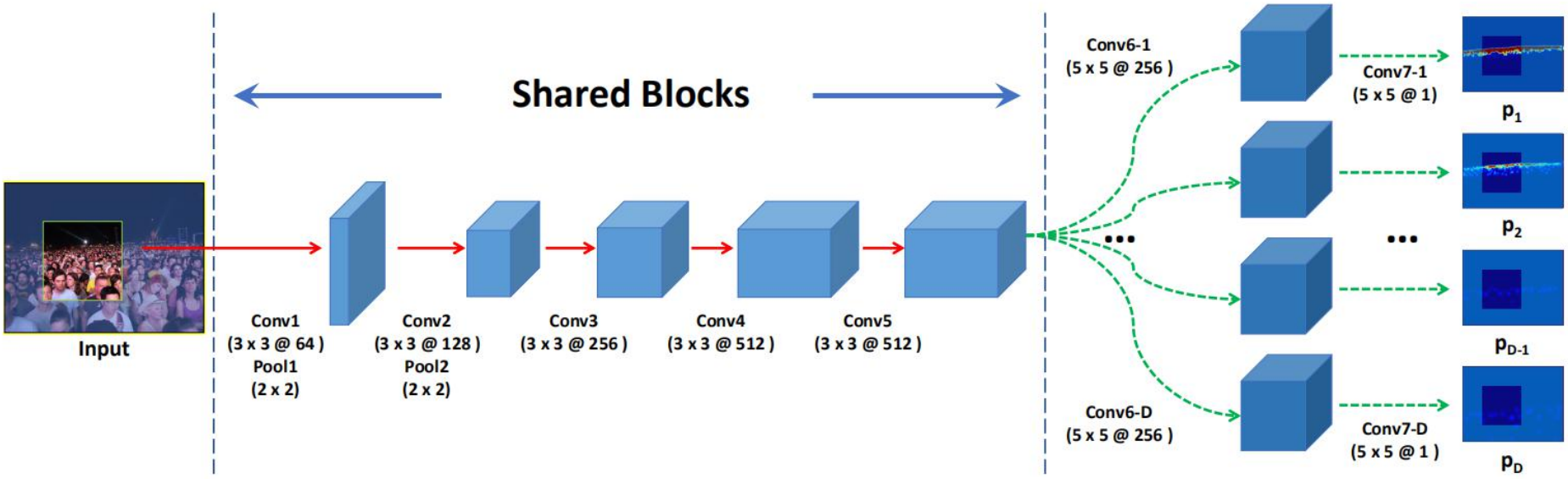
Sindagi, Vishwanath A., and Vishal M. Patel. "Cnn-based cascaded multi-task learning of high-level prior and density estimation for crowd counting." 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). IEEE, 2017.

基于CNN



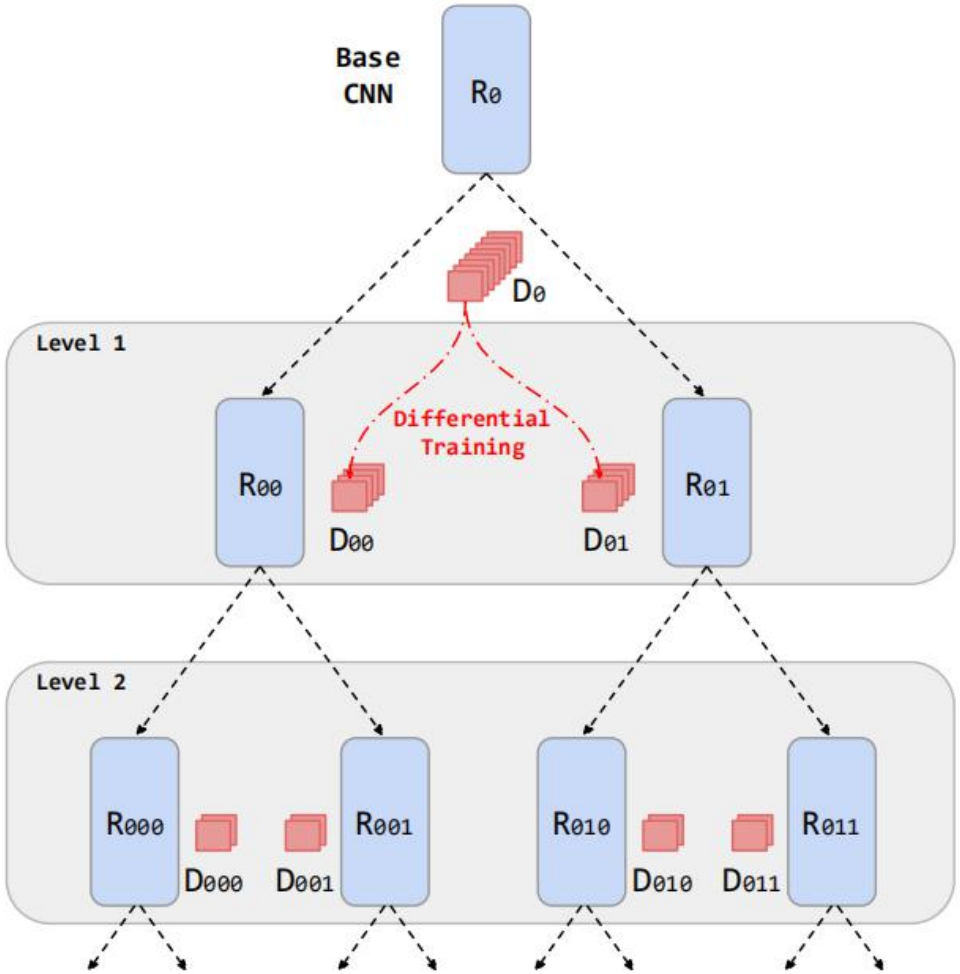
Sam, Deepak Babu, Shiv Surya, and R. Venkatesh Babu. "Switching convolutional neural network for crowd counting." 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2017.

基于CNN



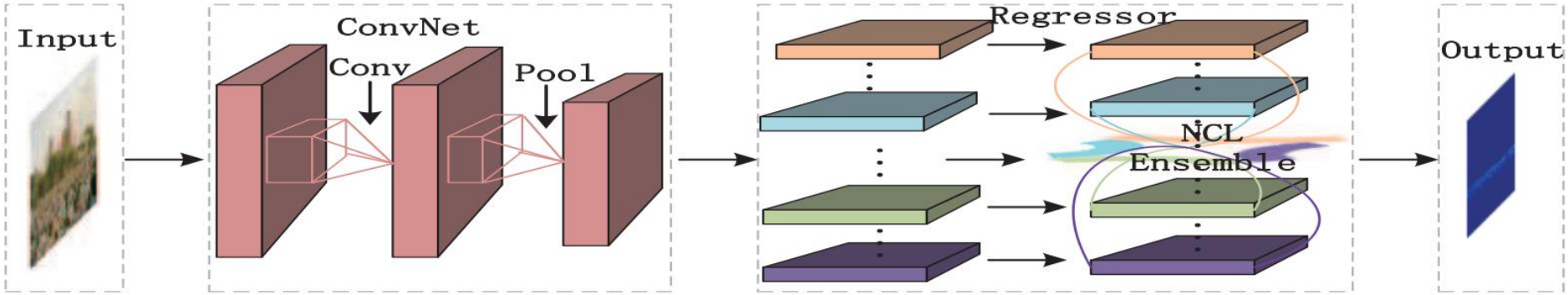
Li, Hanhui, et al. "Structured inhomogeneous density map learning for crowd counting." arXiv preprint arXiv:1801.06642 (2018).

基于CNN



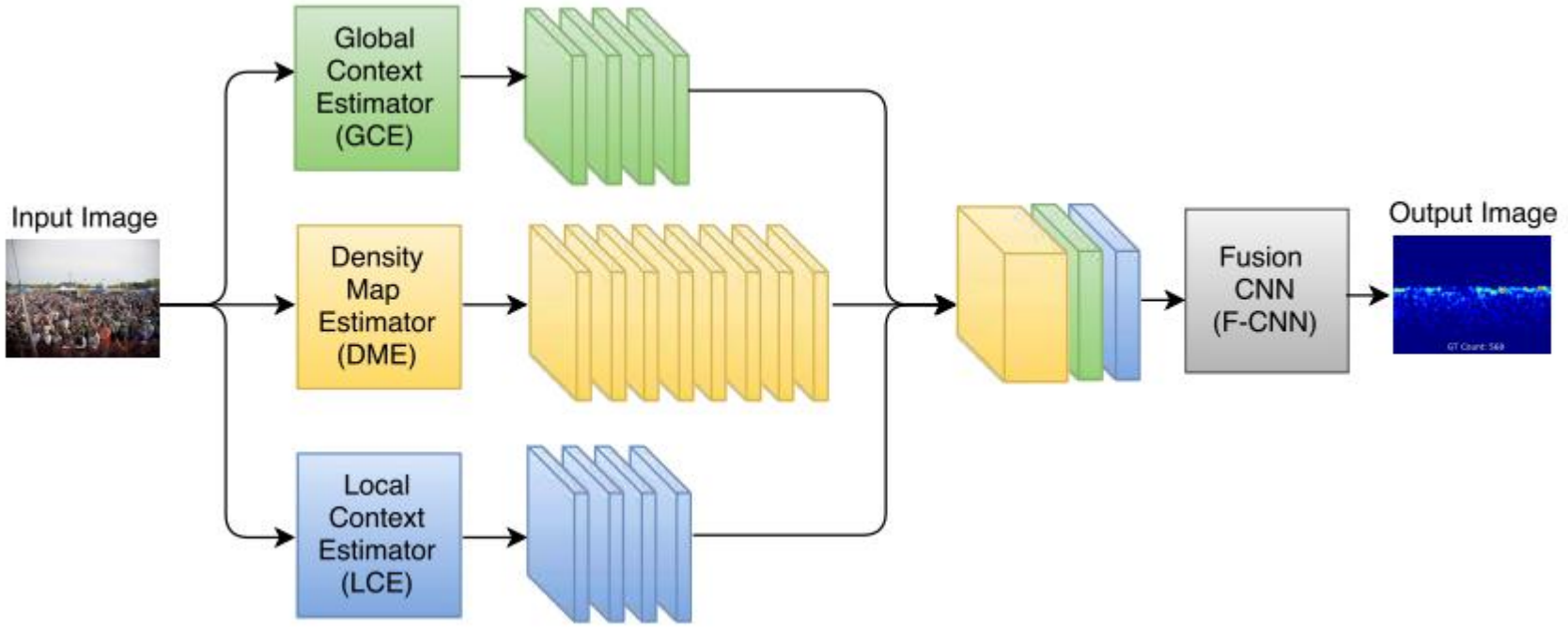
Babu Sam, Deepak, et al. "Divide and grow: Capturing huge diversity in crowd images with incrementally growing cnn." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.

基于CNN



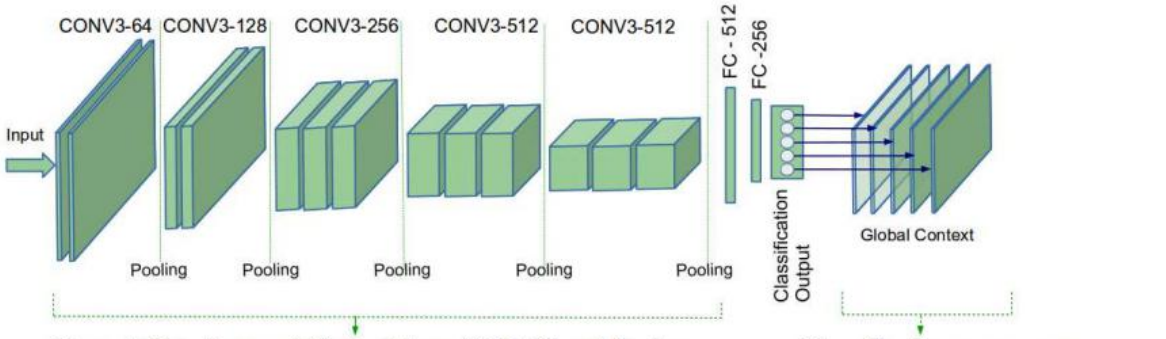
Shi, Zenglin, et al. "Crowd counting with deep negative correlation learning." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.

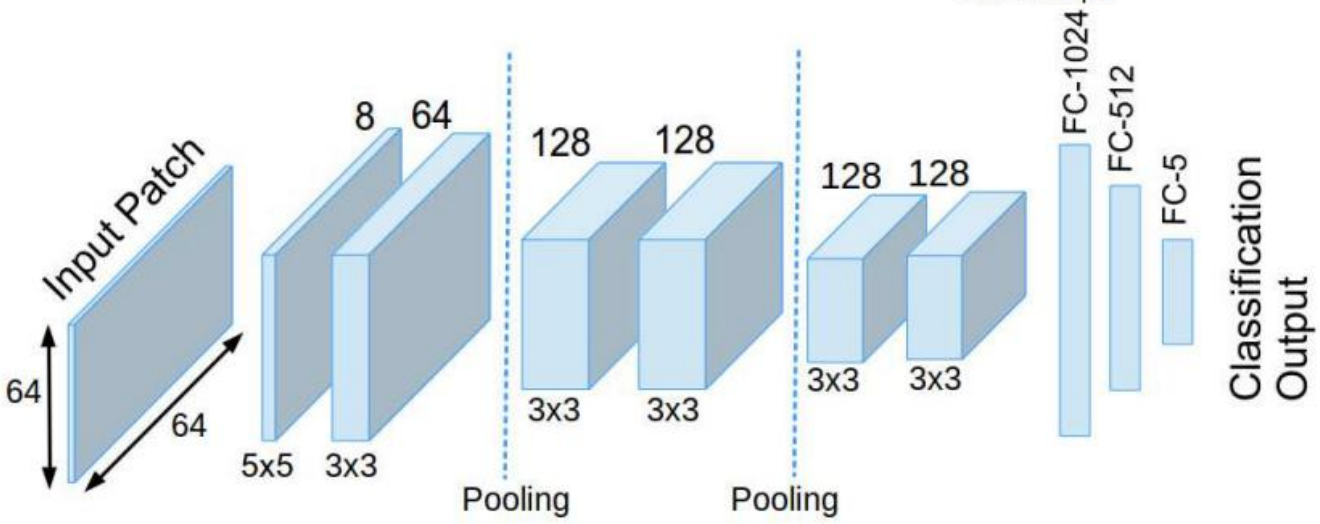
基于CNN



Sindagi, Vishwanath A., and Vishal M. Patel. "Generating high-quality crowd density maps using contextual pyramid cnns." Proceedings of the IEEE International Conference on Computer Vision. 2017.

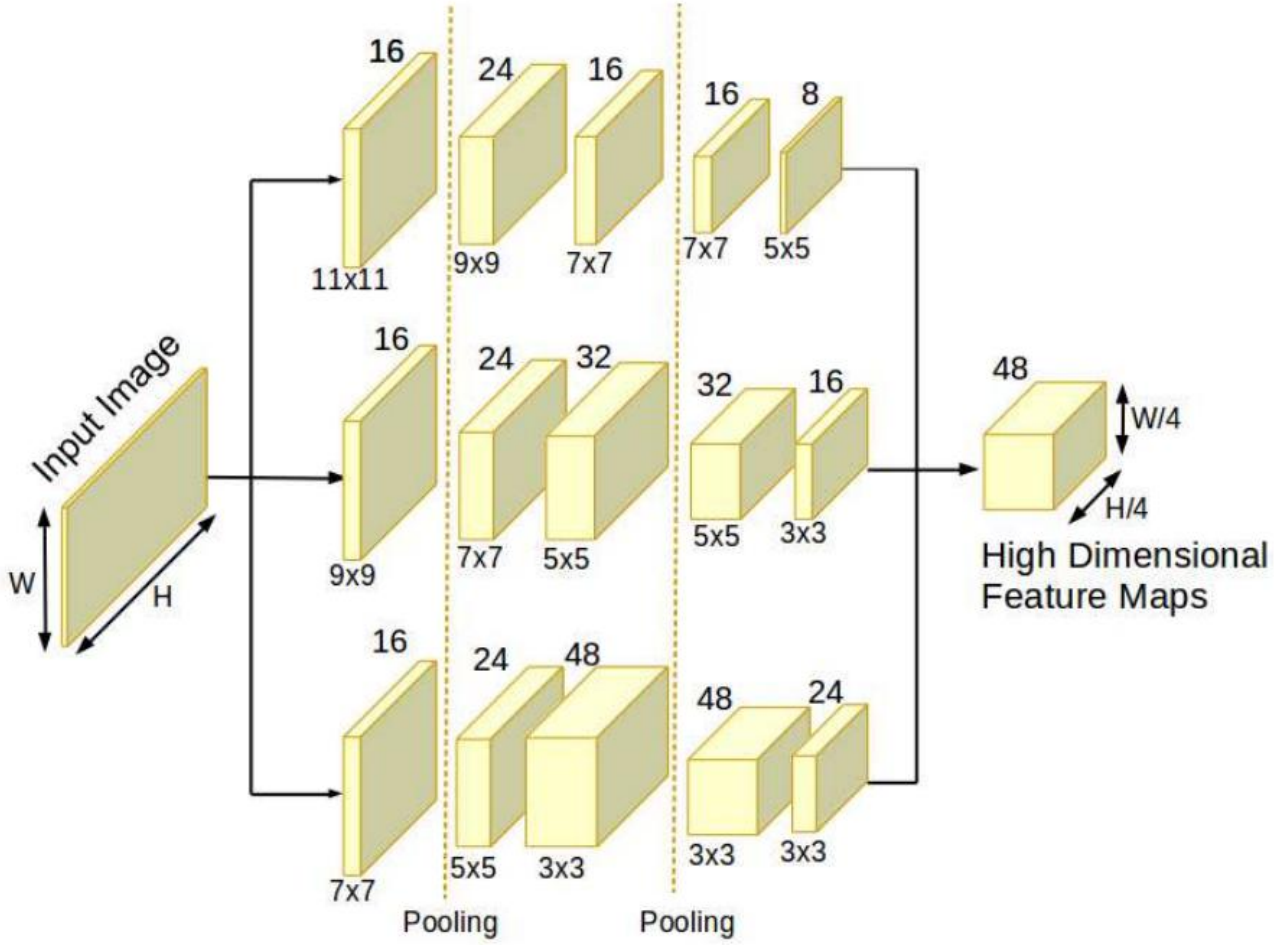
基于CNN





Sindagi, Vishwanath A., and Vishal M. Patel. "Generating high-quality crowd density maps using contextual pyramid cnns." Proceedings of the IEEE International Conference on Computer Vision. 2017.

基于CNN



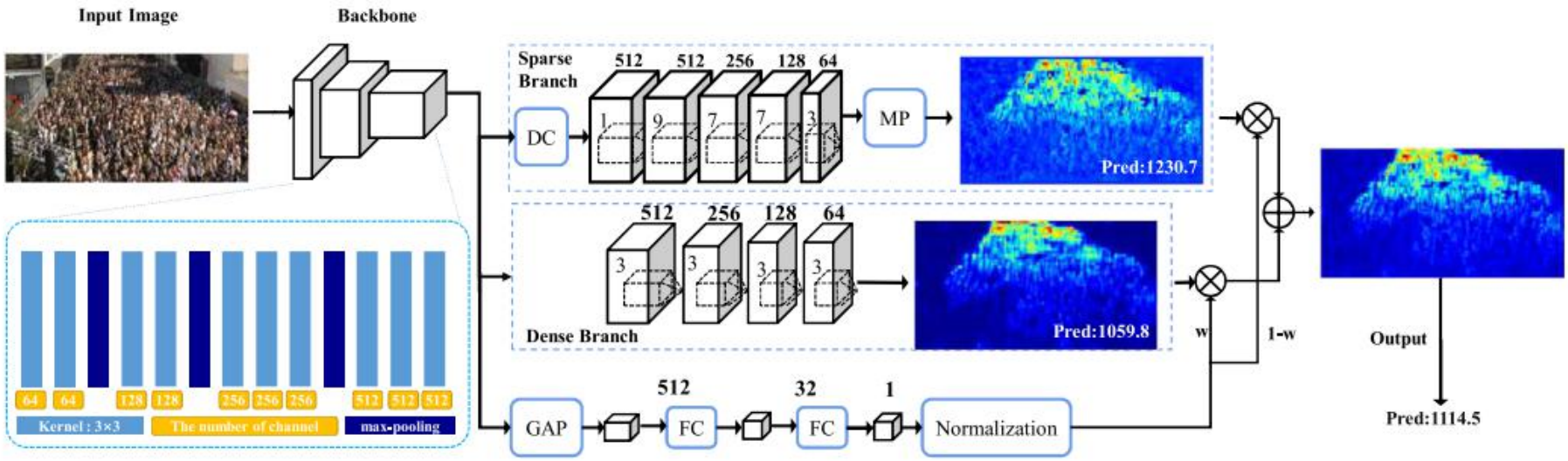
Sindagi, Vishwanath A., and Vishal M. Patel. "Generating high-quality crowd density maps using contextual pyramid cnns." Proceedings of the IEEE International Conference on Computer Vision. 2017.

基于CNN

Configurations of CSRNet			
A	B	C	D
input(unfixed-resolution color image)			
front-end (fine-tuned from VGG-16)			
conv3-64-1 conv3-64-1			
max-pooling			
conv3-128-1 conv3-128-1			
max-pooling			
conv3-256-1 conv3-256-1 conv3-256-1			
max-pooling			
conv3-512-1 conv3-512-1 conv3-512-1			
back-end (four different configurations)			
conv3-512-1 conv3-512-1 conv3-512-1 conv3-256-1 conv3-128-1 conv3-64-1	conv3-512-2 conv3-512-2 conv3-512-2 conv3-256-2 conv3-128-2 conv3-64-2	conv3-512-2 conv3-512-2 conv3-512-2 conv3-256-4 conv3-128-4 conv3-64-4	conv3-512-4 conv3-512-4 conv3-512-4 conv3-256-4 conv3-128-4 conv3-64-4
conv1-1-1			

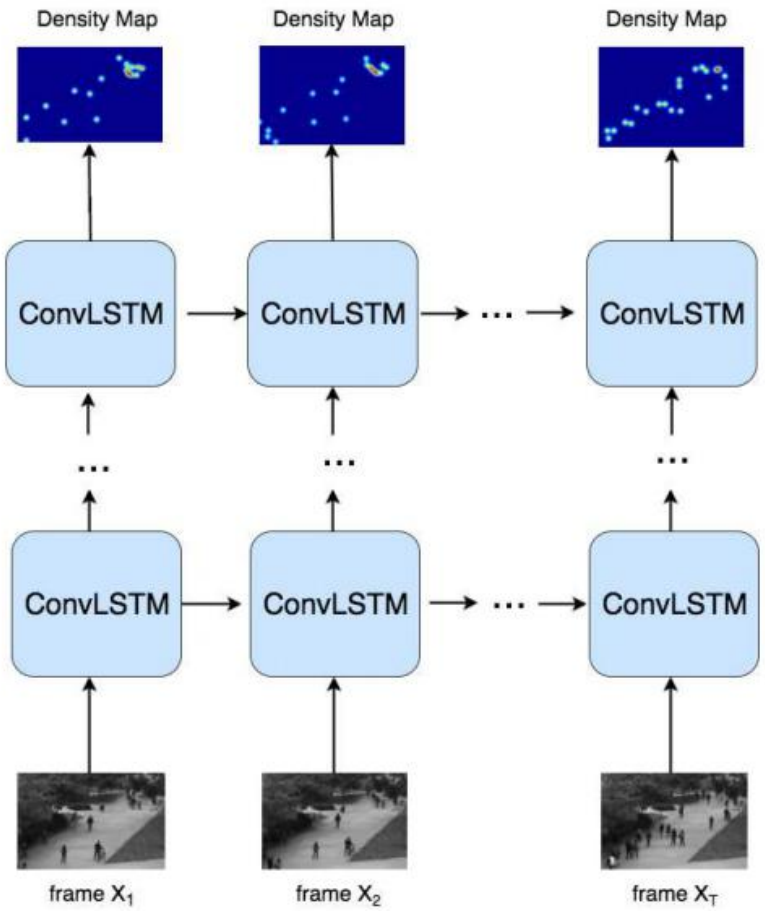
Li, Yuhong, Xiaofan Zhang, and Deming Chen. "Csrnet: Dilated convolutional neural networks for understanding the highly congested scenes." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.

基于CNN



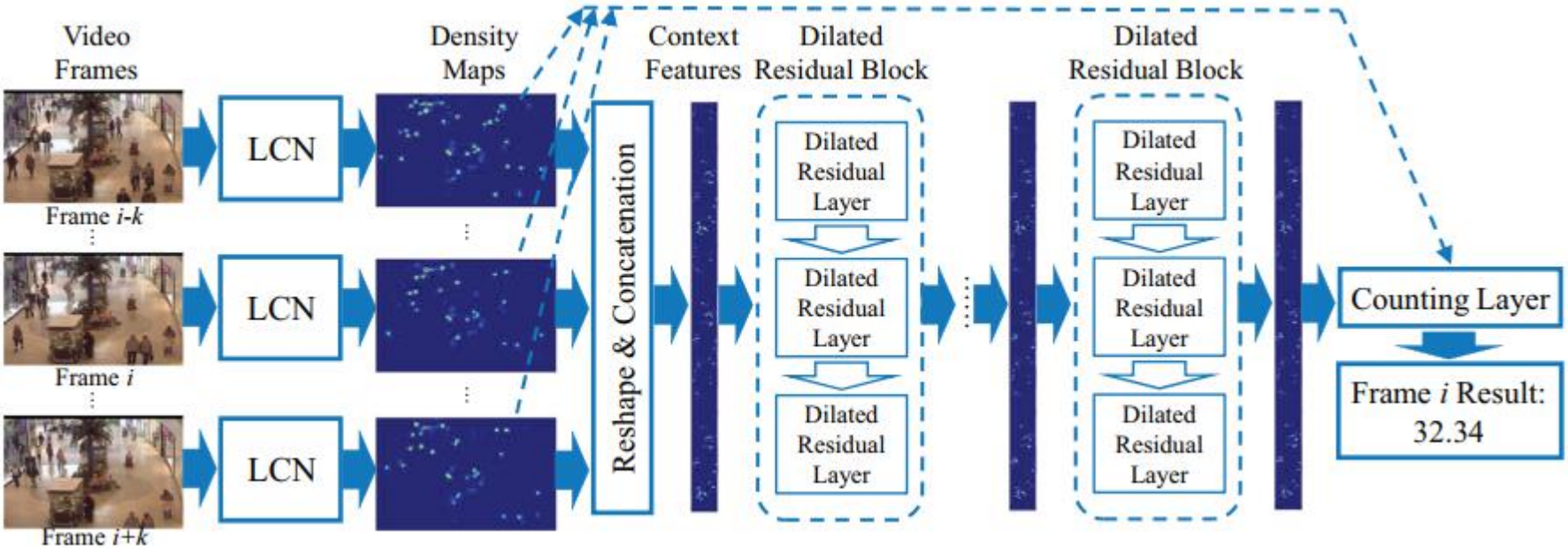
Wu, Xingjiao, et al. "Counting crowds with varying densities via adaptive scenario discovery framework." Neurocomputing (2020).

video



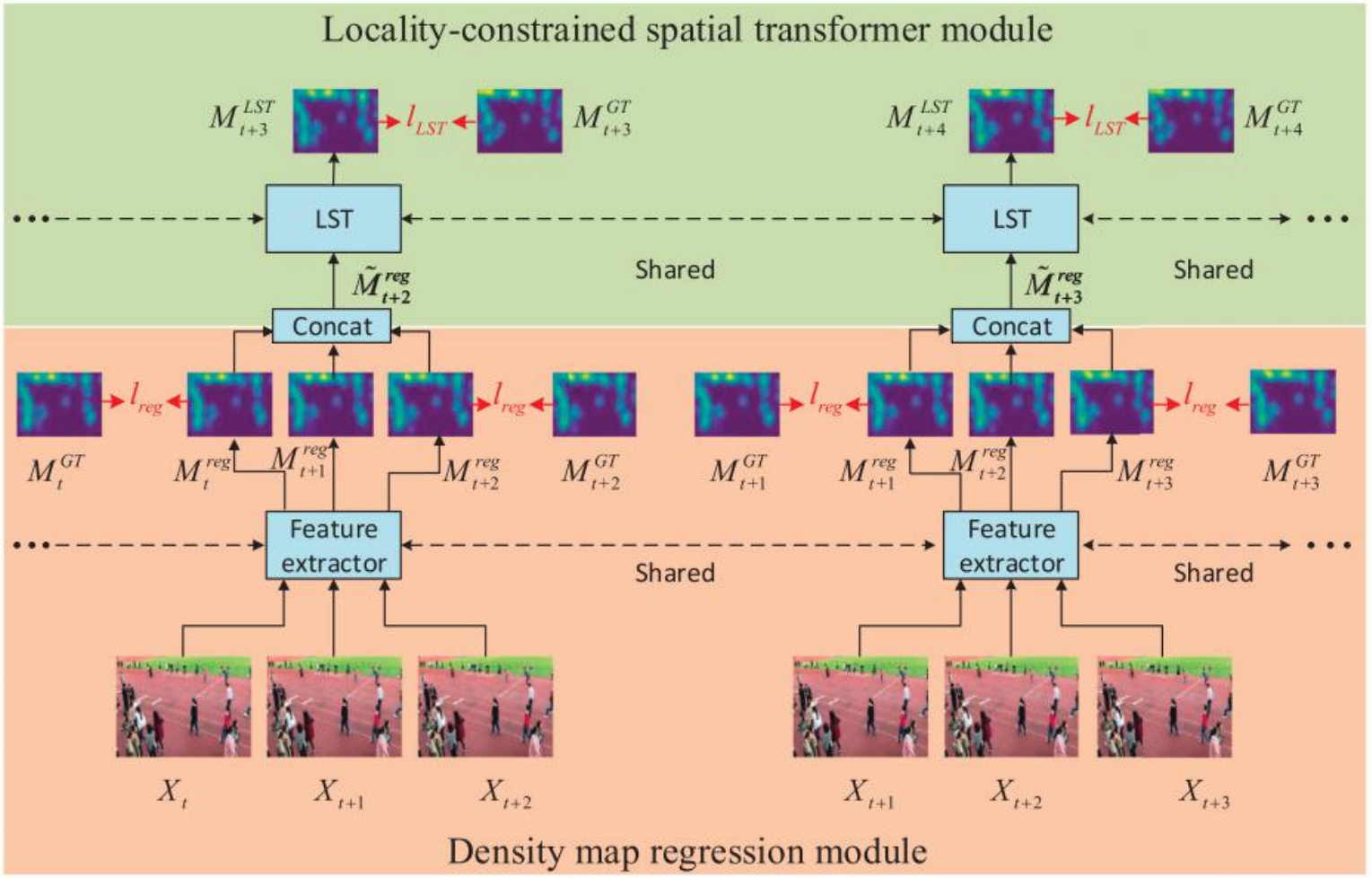
Xiong, Feng, Xingjian Shi, and Dit-Yan Yeung. "Spatiotemporal modeling for crowd counting in videos." Proceedings of the IEEE International Conference on Computer Vision. 2017.

video



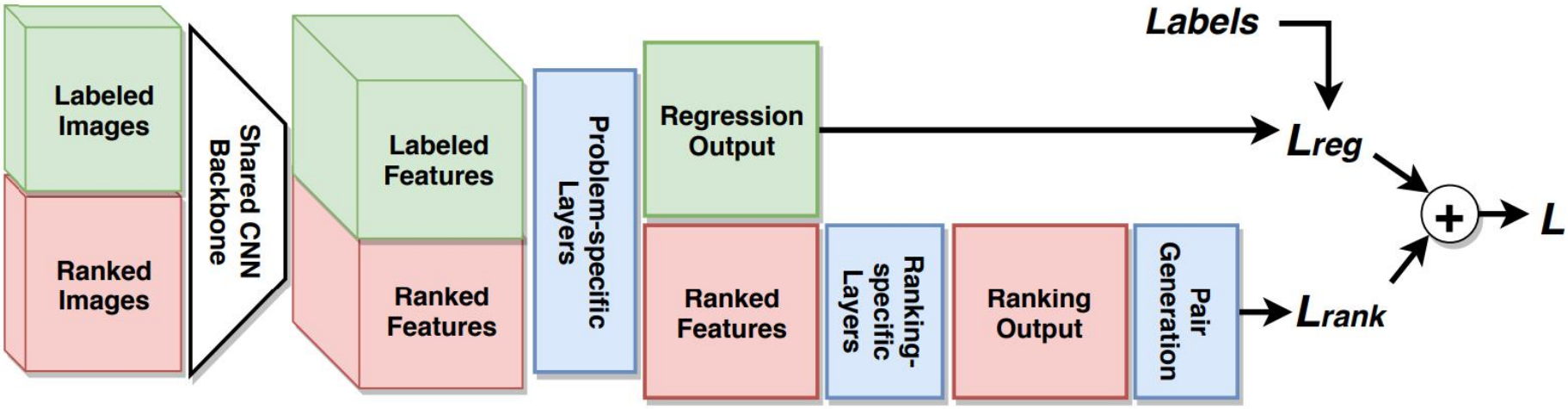
Wu, Xingjiao, et al. "Fast video crowd counting with a temporal aware network." Neurocomputing (2020).

video



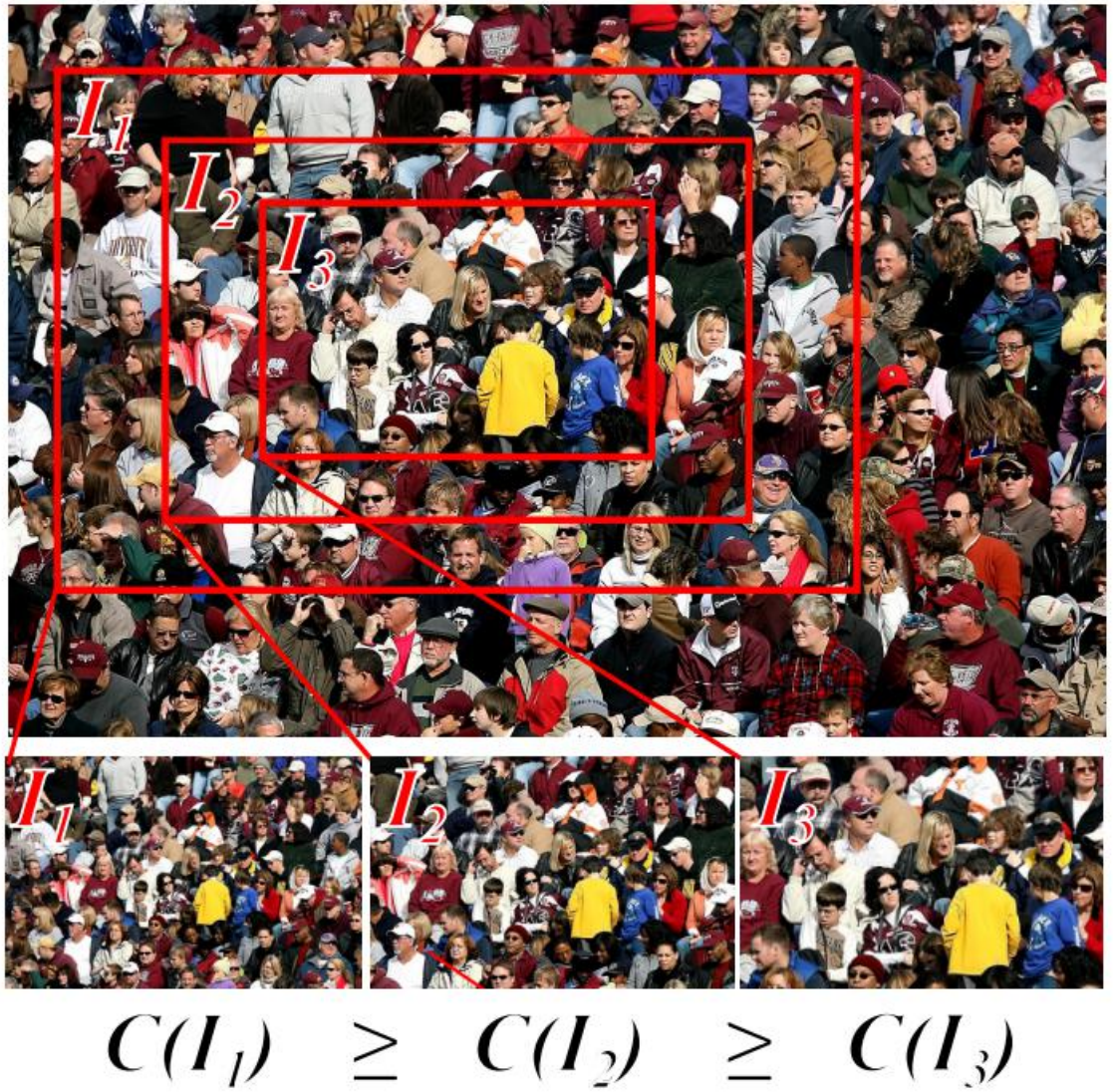
Fang, Yanyan, et al. "Multi-level feature fusion based Locality-Constrained Spatial Transformer network for video crowd counting." Neurocomputing (2020).

无监督



Liu, Xialei, Joost Van De Weijer, and Andrew D. Bagdanov. "Leveraging unlabeled data for crowd counting by learning to rank." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018.
Liu, Xialei, Joost Van De Weijer, and Andrew D. Bagdanov. "Exploiting unlabeled data in cnns by self-supervised learning to rank." IEEE transactions on pattern analysis and machine intelligence 41.8 (2019): 1862-1878.

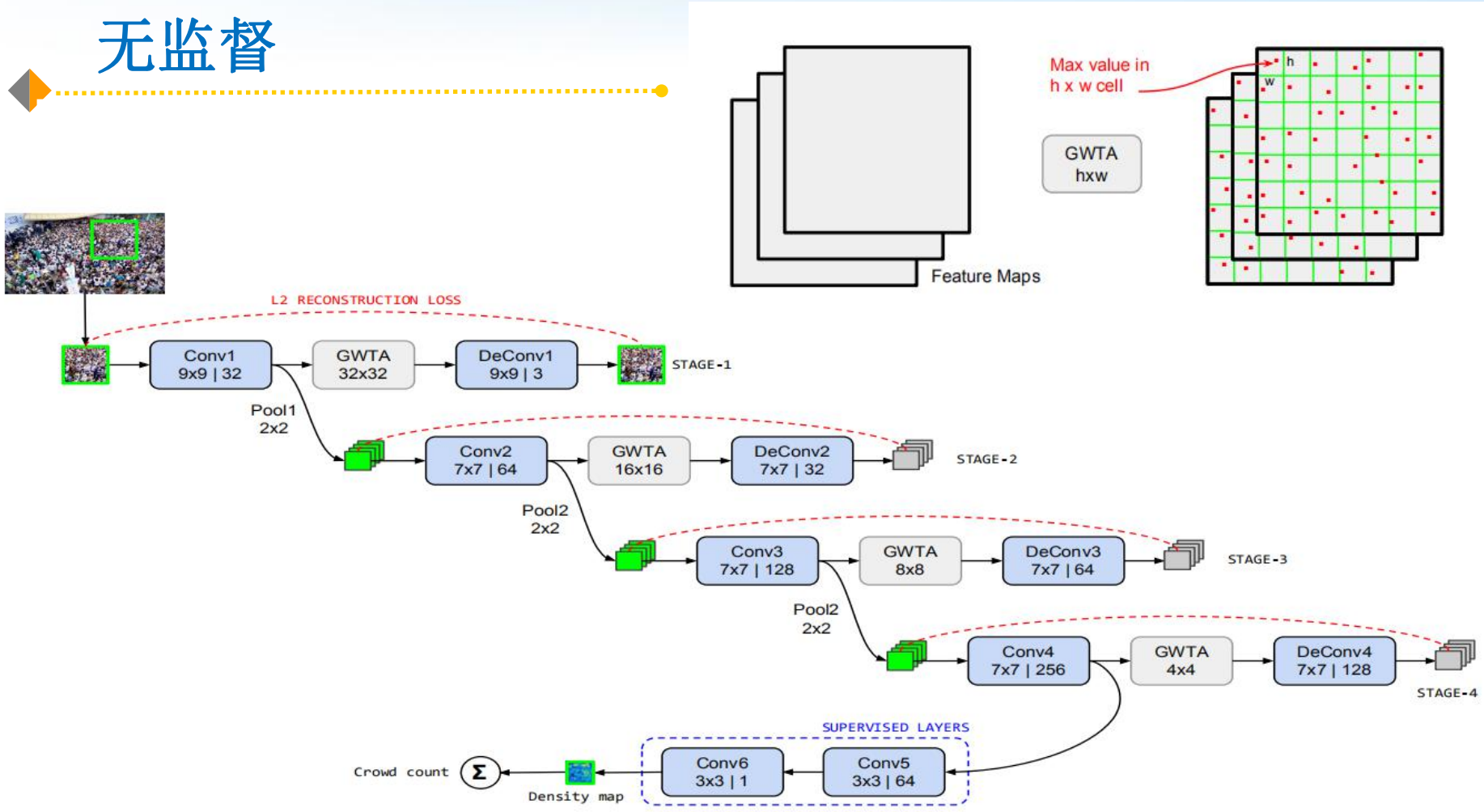
无监督



$$C(I_1) \geq C(I_2) \geq C(I_3)$$

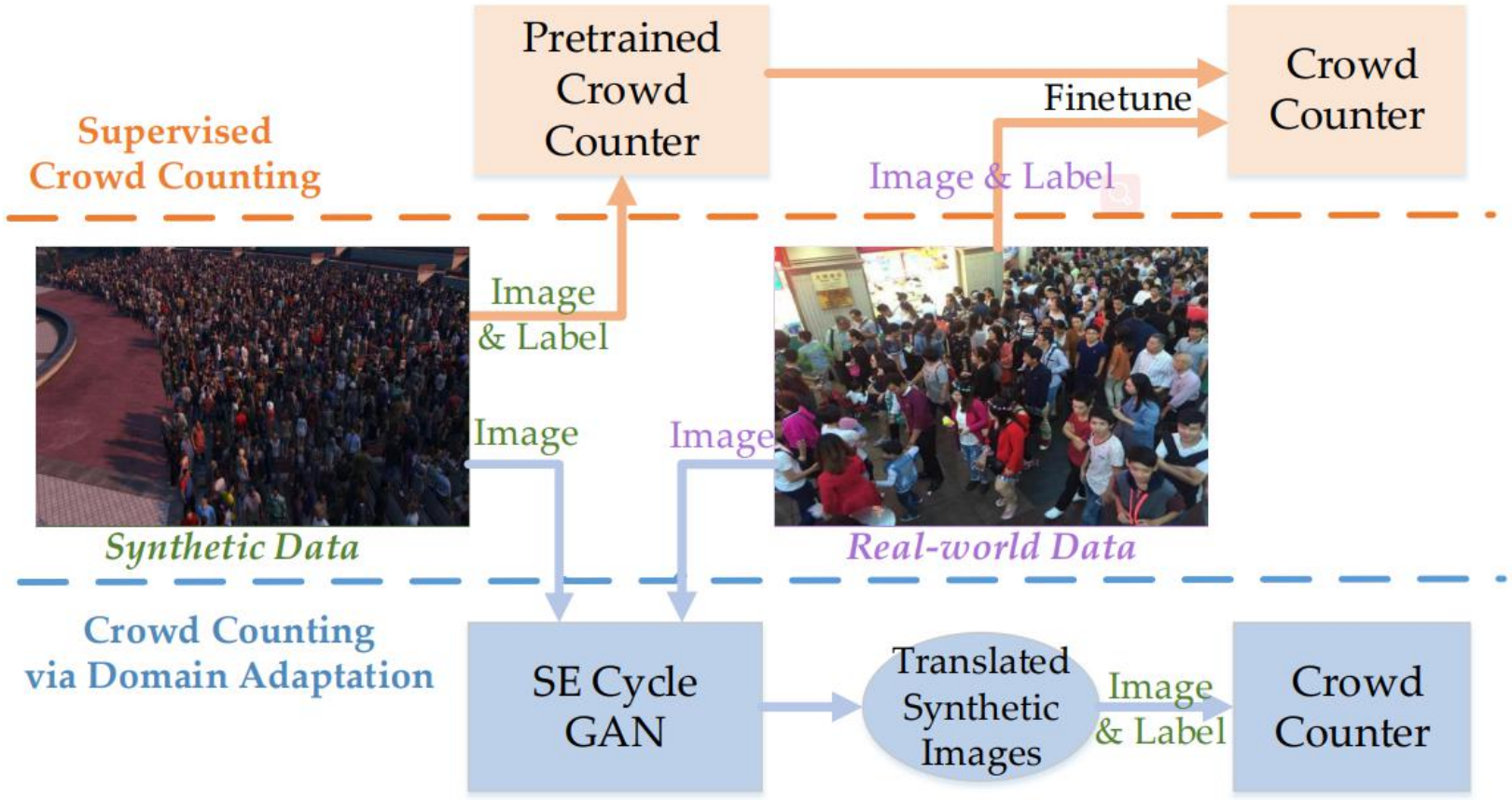
Liu, Xialei, Joost Van De Weijer, and Andrew D. Bagdanov. "Leveraging unlabeled data for crowd counting by learning to rank." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018.
Liu, Xialei, Joost Van De Weijer, and Andrew D. Bagdanov. "Exploiting unlabeled data in cnns by self-supervised learning to rank." IEEE transactions on pattern analysis and machine intelligence 41.8 (2019): 1862-1878.

无监督



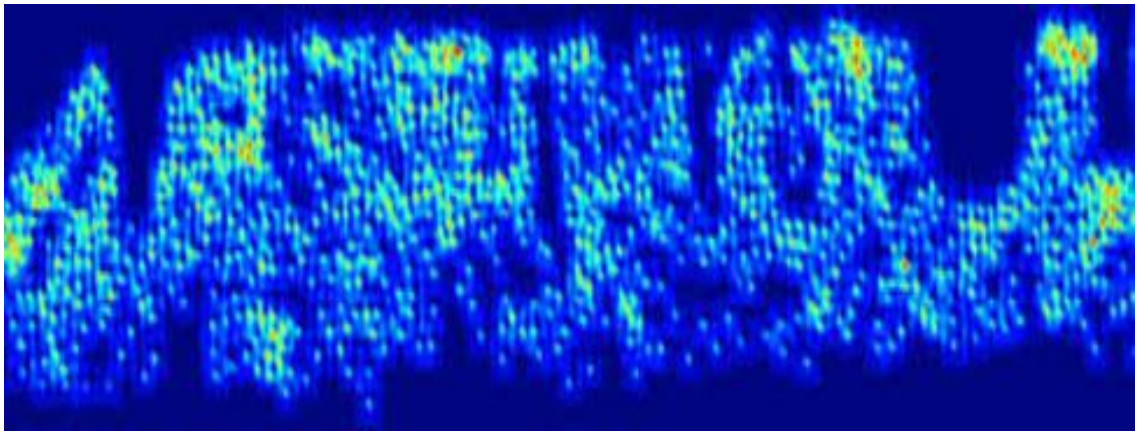
Sam, Deepak Babu, et al. "Almost unsupervised learning for dense crowd counting." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 33. 2019.

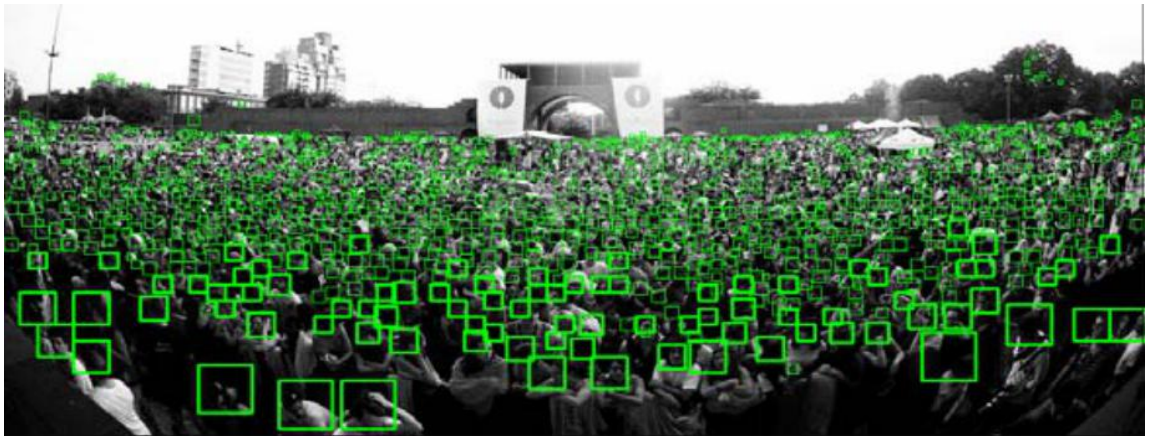
数据集构建



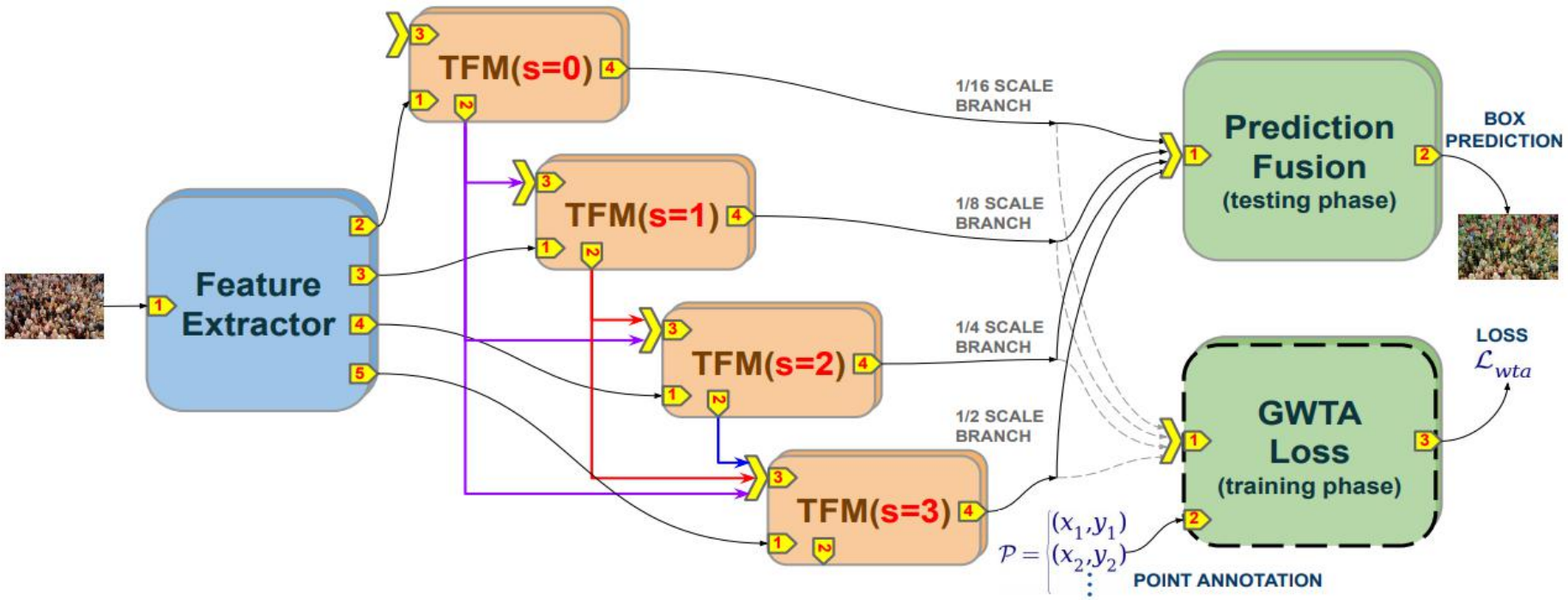
Wang, Qi, et al. "Learning from synthetic data for crowd counting in the wild." Proceedings of the IEEE conference on computer vision and pattern recognition. 2019.

定位



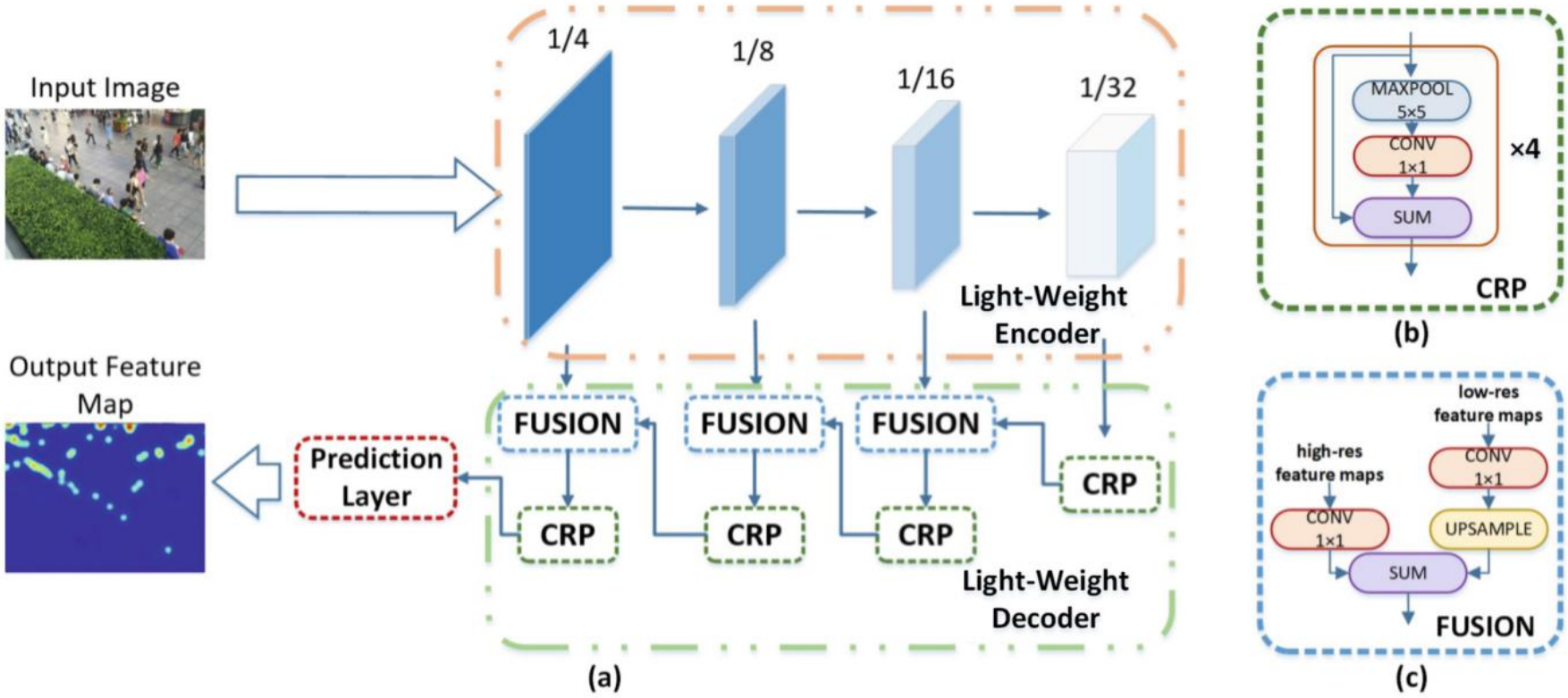


定位



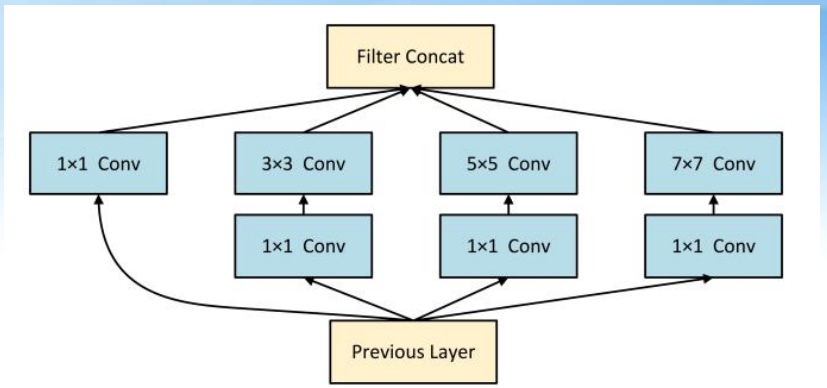
Sam, Deepak Babu, et al. "Locate, Size and Count: Accurately Resolving People in Dense Crowds via Detection." IEEE Transactions on Pattern Analysis and Machine Intelligence (2020).

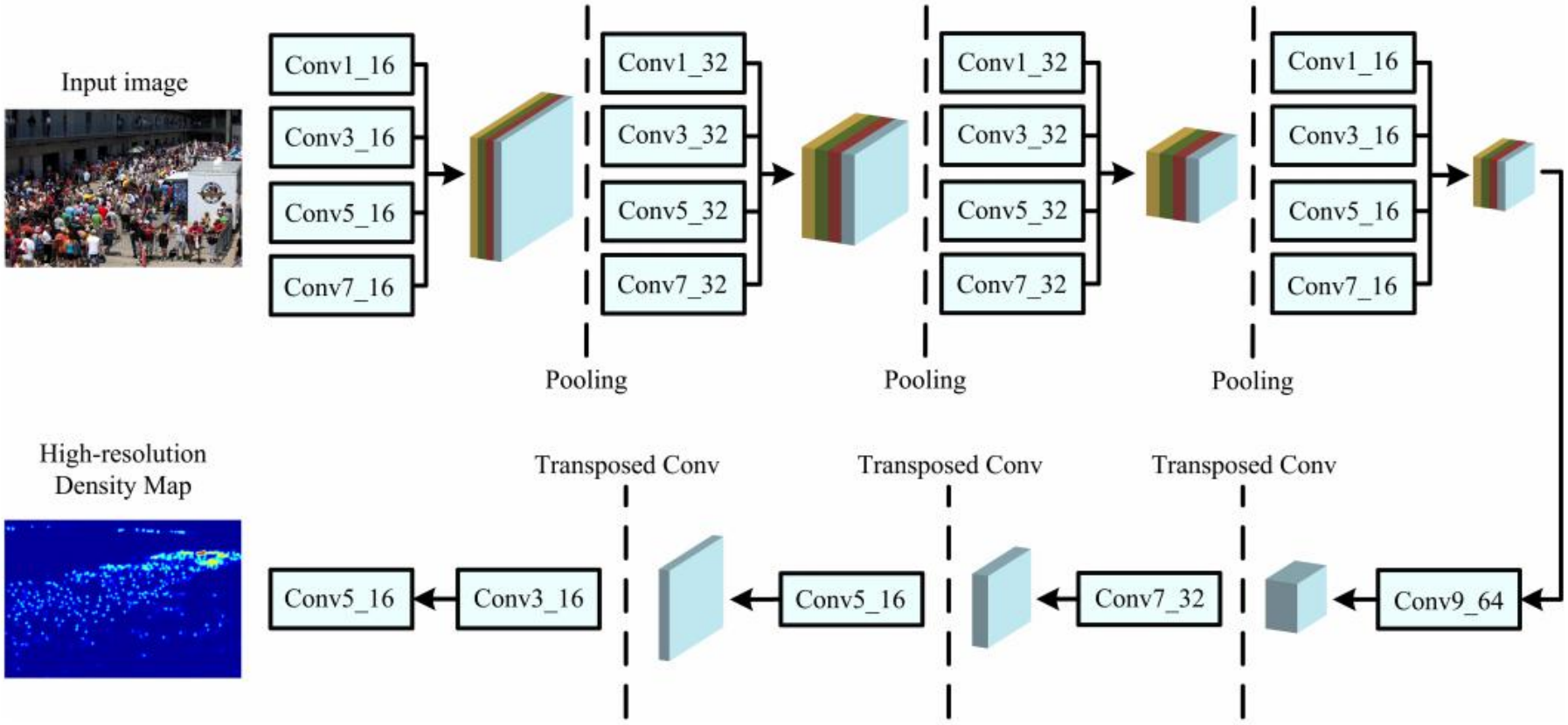
轻量级



Wang, Peng, et al. "MobileCount: An efficient encoder-decoder framework for real-time crowd counting." Neurocomputing 407 (2020): 292-299.

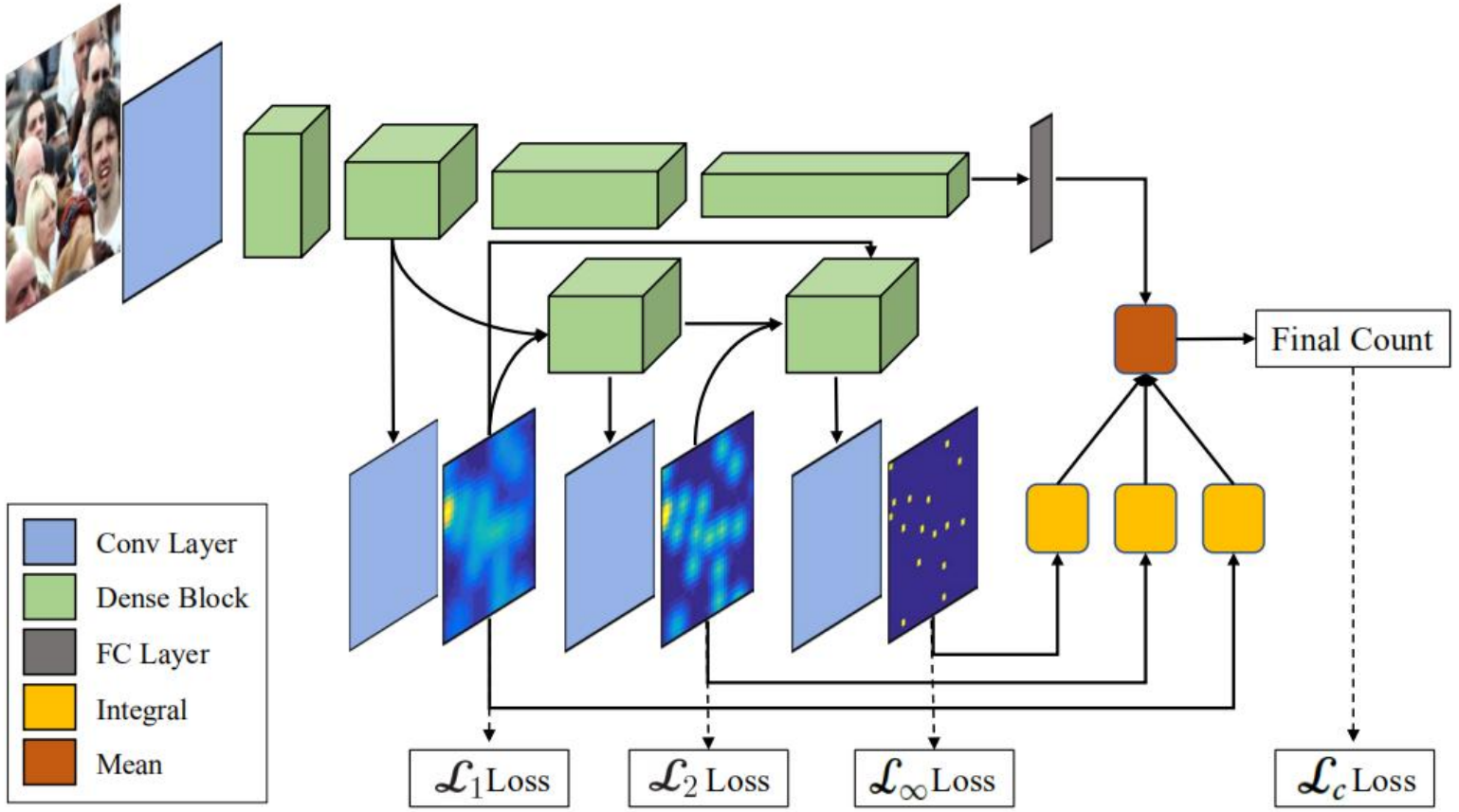
轻量级





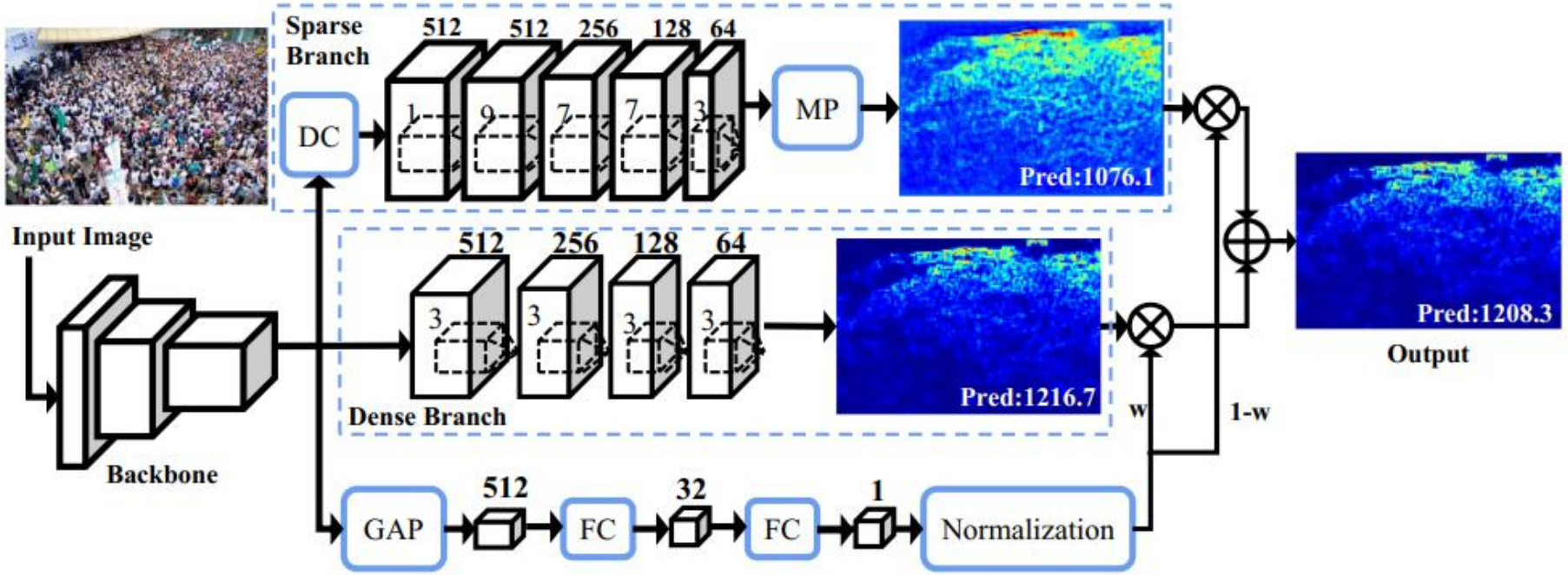
Cao, Xinkun, et al. "Scale aggregation network for accurate and efficient crowd counting." Proceedings of the European Conference on Computer Vision (ECCV). 2018.

其他

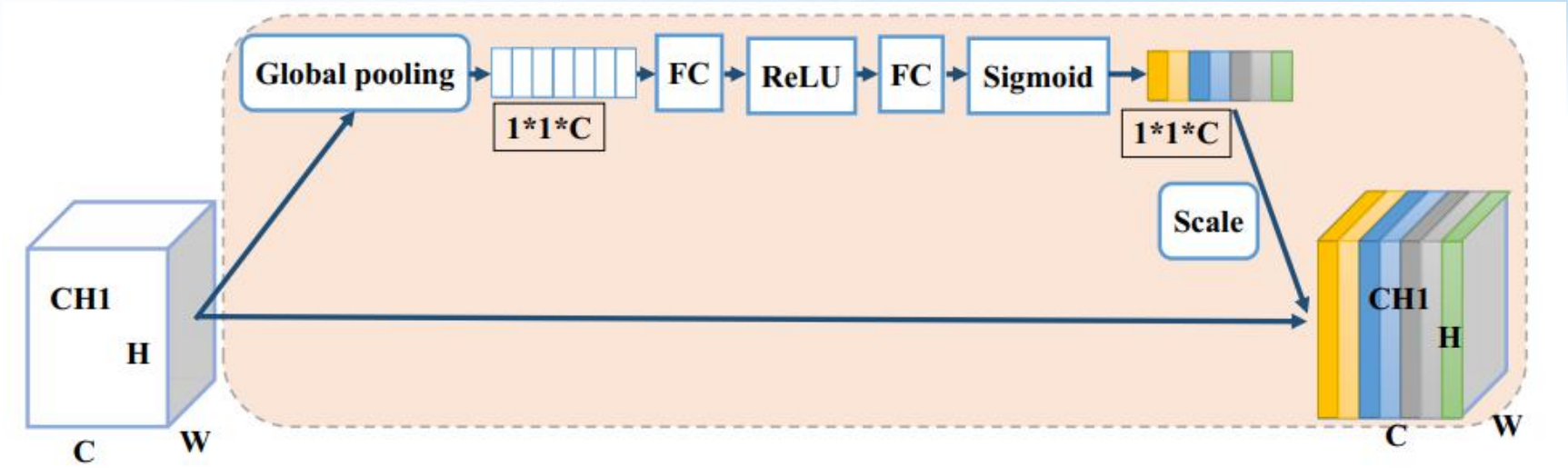


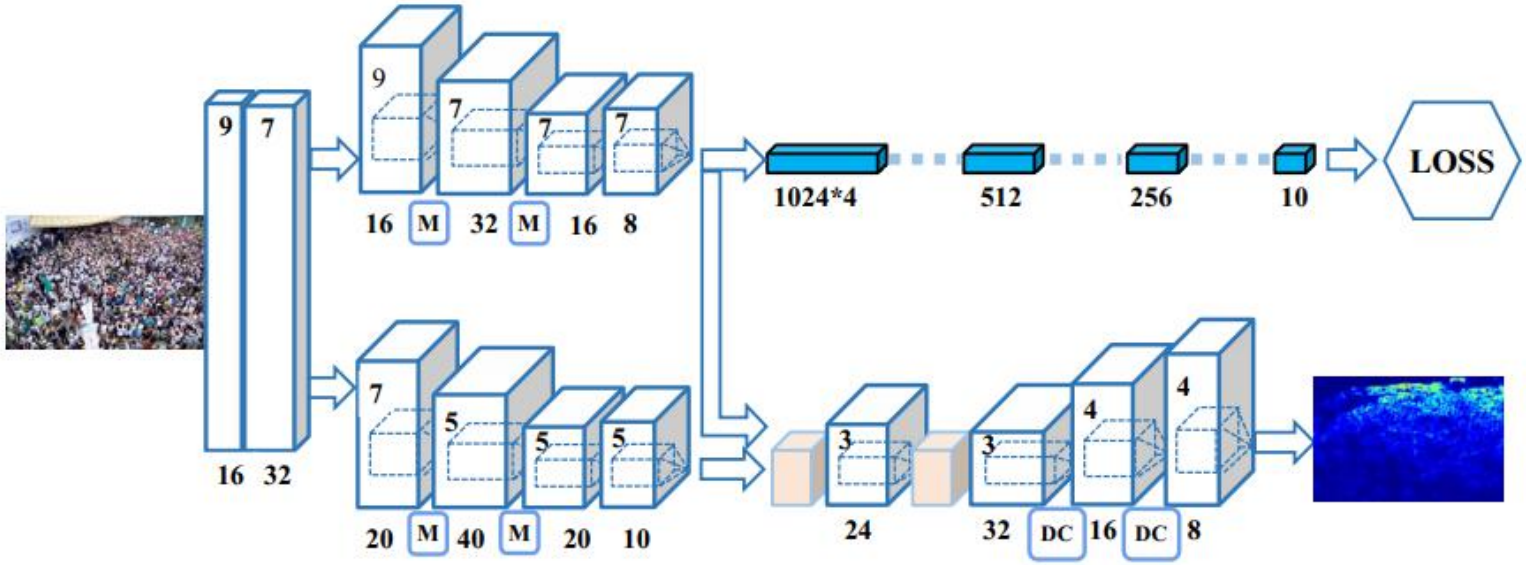
Idrees, Haroon, et al. "Composition loss for counting, density map estimation and localization in dense crowds." Proceedings of the European Conference on Computer Vision (ECCV). 2018.

其他



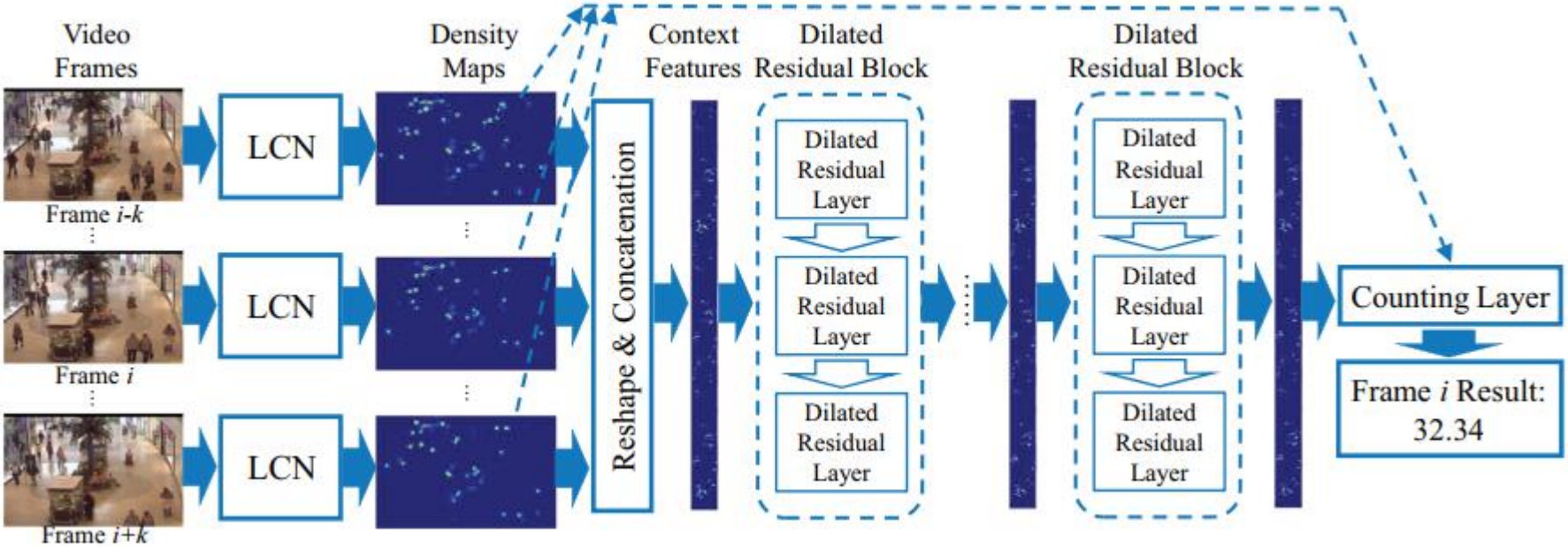
Wu, Xingjiao, et al. "Adaptive scenario discovery for crowd counting." ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2019.





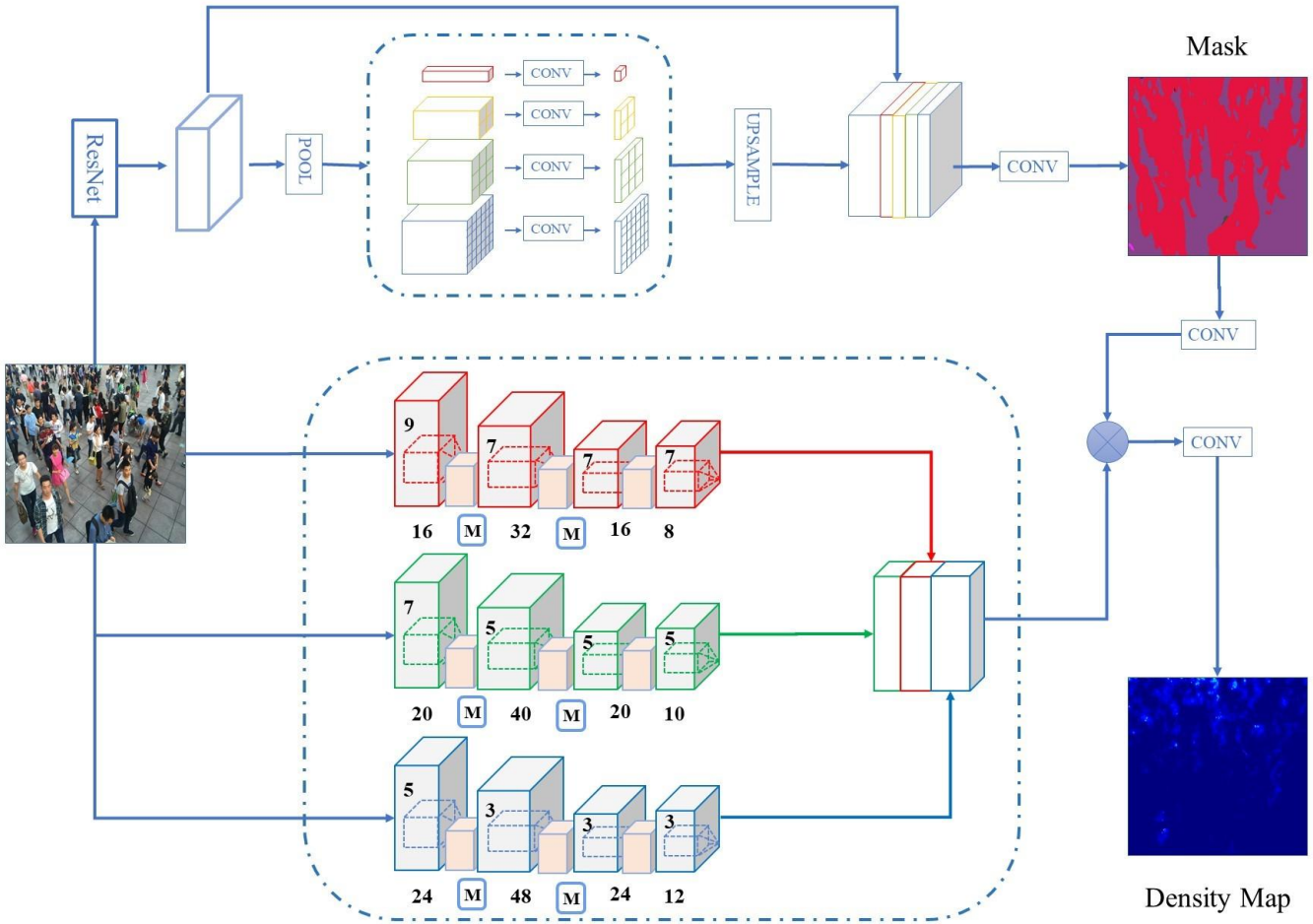
Wu, Xingjiao, et al. "Feature channel enhancement for crowd counting." IET Image Processing 14.11 (2020): 2376-2382.

其他



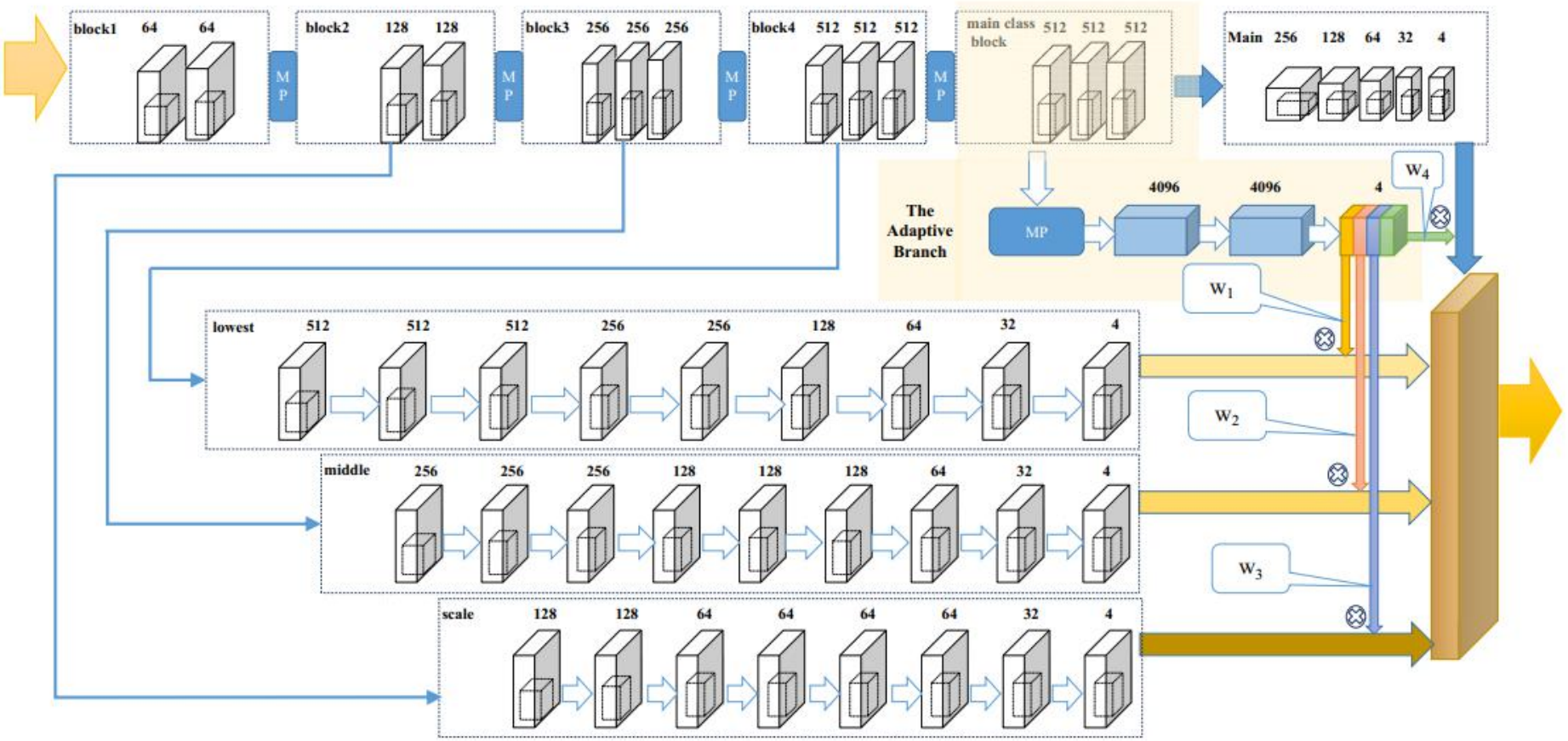
Wu, Xingjiao, et al. "Fast video crowd counting with a temporal aware network." Neurocomputing (2020).

其他



Jie He, Xingjiao Wu, Jing Yang, Wenxin Hu, CPSPNet: Crowd Counting via Semantic Segmentation Framework[C]//International Conference on Tools with Artificial Intelligence (ICTAI). IEEE, 2020

其他



Sun, Yixuan, et al. "Counting Crowds with Perspective Distortion Correction via Adaptive Learning." Sensors 20.13 (2020): 3781.

汇报完毕，谢谢！