



计算机视觉青年开发者榜单活动

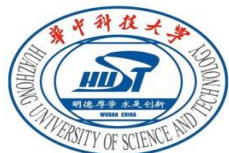
# Learn to Scale: Generating Multipolar Normalized Density Map for Crowd Counting

参赛队员：梁定康、徐晨丰、徐志良、陈习武、杨贤

指导老师：许永超、白翔

参赛单位：华中科技大学(MCLAB VLR\_Group)

参加赛道：平安城市——人流密度统计



# 人流密度统计在公共安全、监控等扮演重要角色

本次实战项目需要参赛者基于给定的场景图片，开发出能够同时适用于**密集**和**稀疏**等多种复杂场景的人数统计算法，准确输出图片中的**总人数**



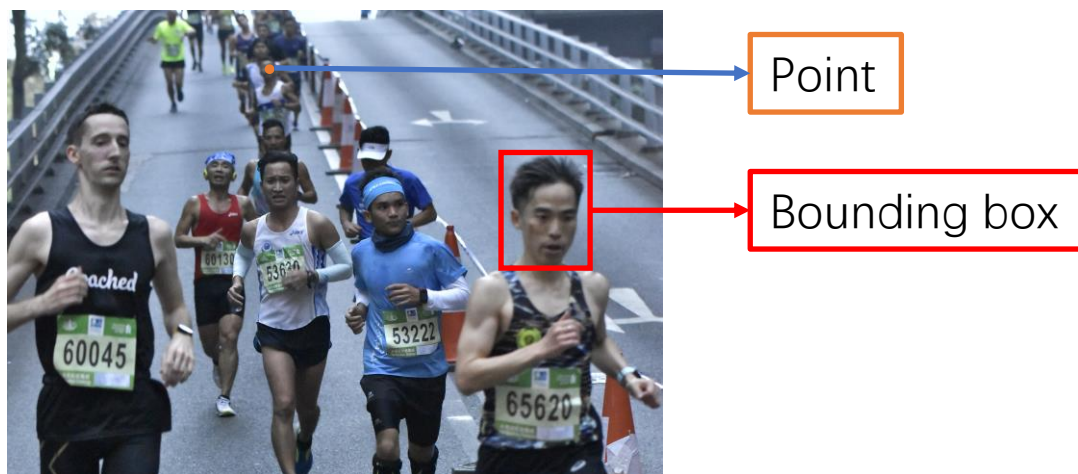
面临**多尺度**、**光照**、**遮挡**、**模糊**等多种复杂问题

# 人头检测 or 回归密度图

<http://www.vlrlab.net/>

比赛数据集提供了两种标注方式

- 提供bounding box标注 (大人头)
- 提供point标注 (小人头)



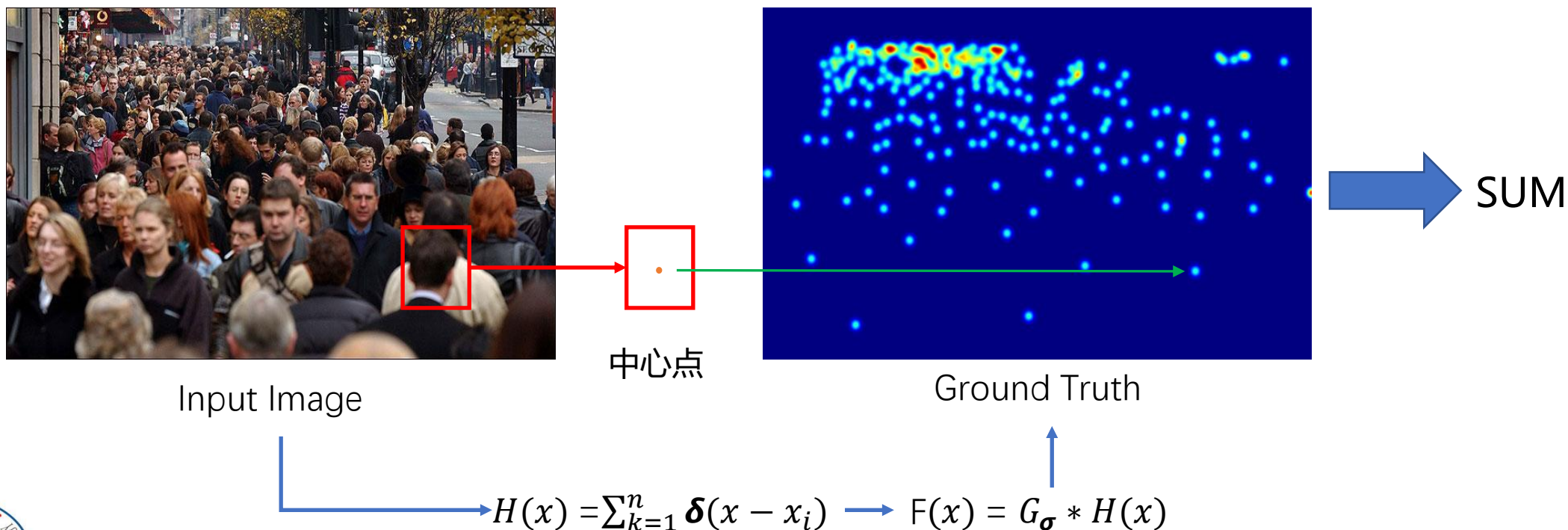
- 基于检测的方法, 由于有些人头没有给出 bounding box标注, 因此在密集区域, 很难训练检测网络, 同时, 人头与人头之间还会存在多尺度、遮挡、光照等复杂问题, 实际测试时难以检测出人头位置。
- 基于回归的方法, 通过回归密度图的方式, 能较好的解决密集、遮挡等问题, 同时也仅需提供 point标注。



# 生成Ground Truth

<http://www.vlrlab.net/>

- 统一用Point标注, 对于bounding box则求取中心
- 对每个坐标点, 做高斯变换得到密度图
- 对密度图积分得到总人数



# 普通人如何计数 / 工作人员如何标注数据

<http://www.vlrlab.net/>





# 让网络学会去放大 (learn to scale)

<http://www.vlrlab.net/>



过大



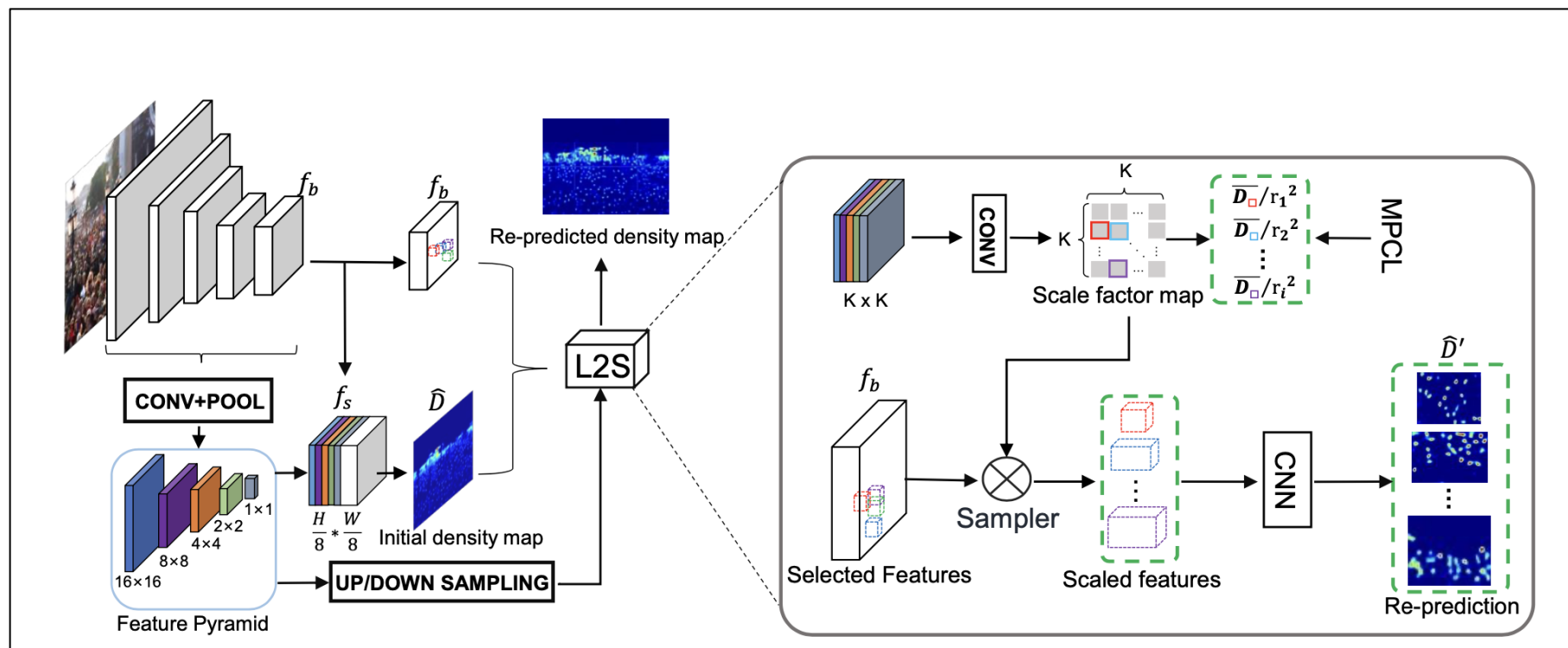
合适



过小

比赛数据集多为监控摄像头所拍摄，成像质量较差，放大会失真，放小不利于计数

# L2SM : Generating Multipolar Normalized Density Map for Crowd Counting



本次比赛采用的网络基于本团队发表于ICCV 2019的工作改进而来

<https://arxiv.org/abs/1907.12428>

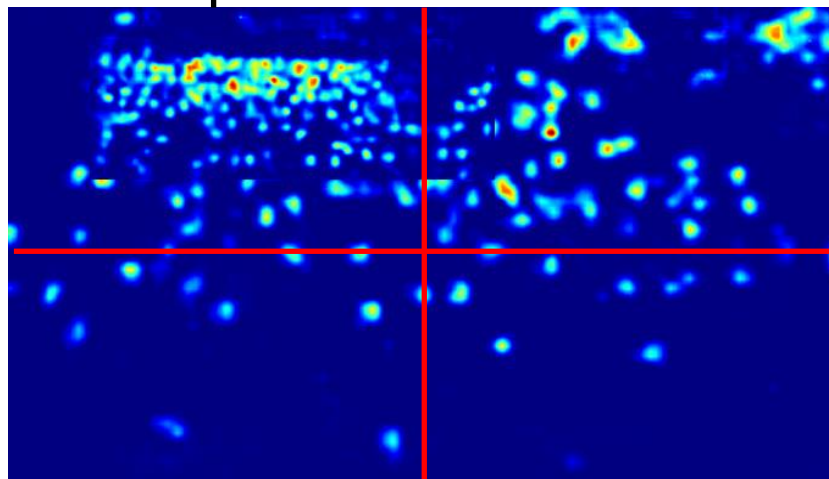
# 核心思想：优化密集区域

<http://www.vlrlab.net/>

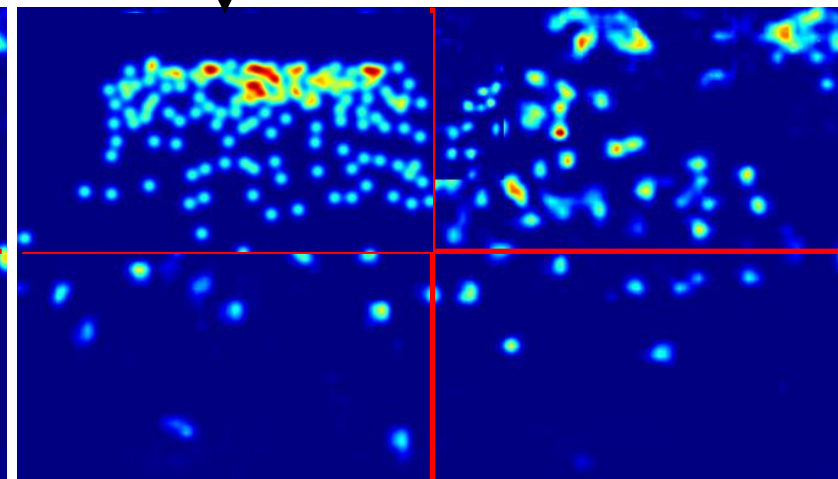
密集区域放大优化



原始图片



初始预测密度图



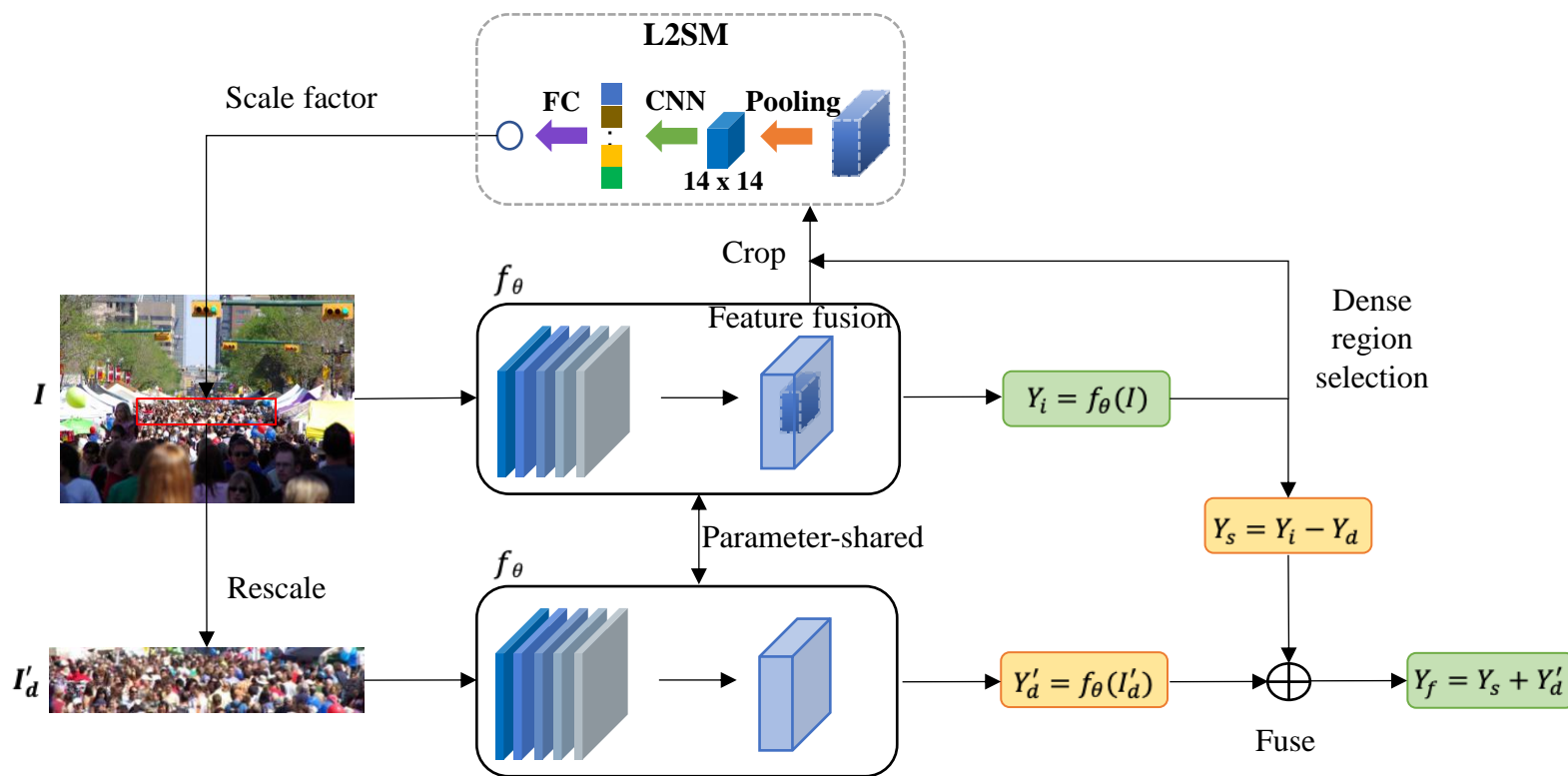
最终预测密度图

本质在于如何能同时优化稀疏、密集区域



# L2SM改进工作——本次比赛所用网络

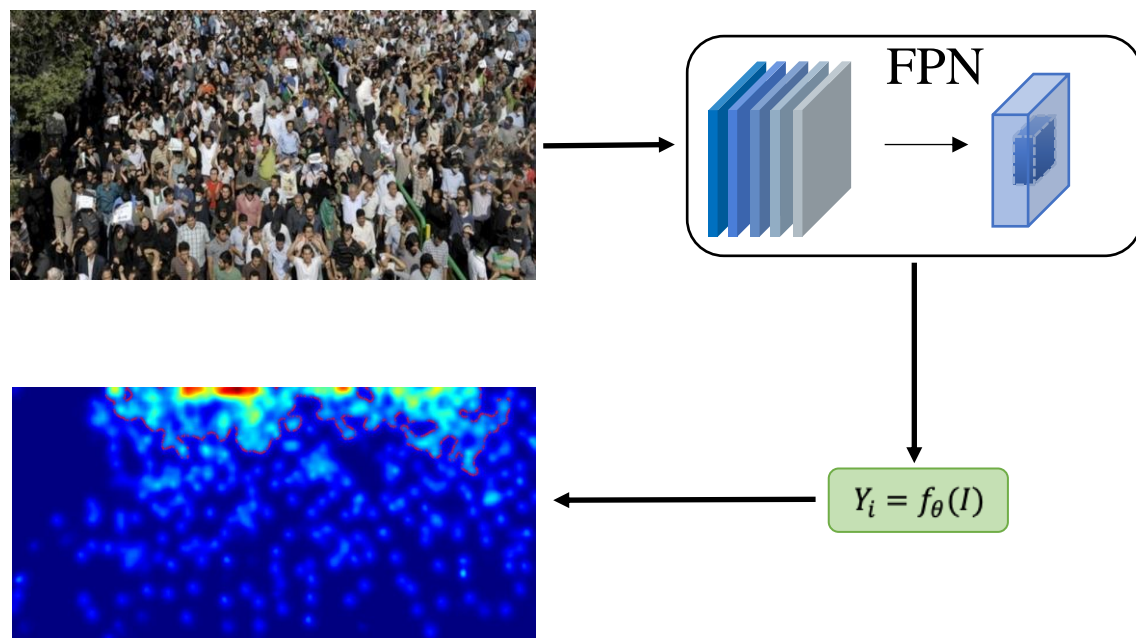
<http://www.vlrlab.net/>



- 直接寻找 **Attention region**
- 更好的baseline
- Refine原图，而非feature
- **共享权重**
- 速度更快 (**10FPS**)

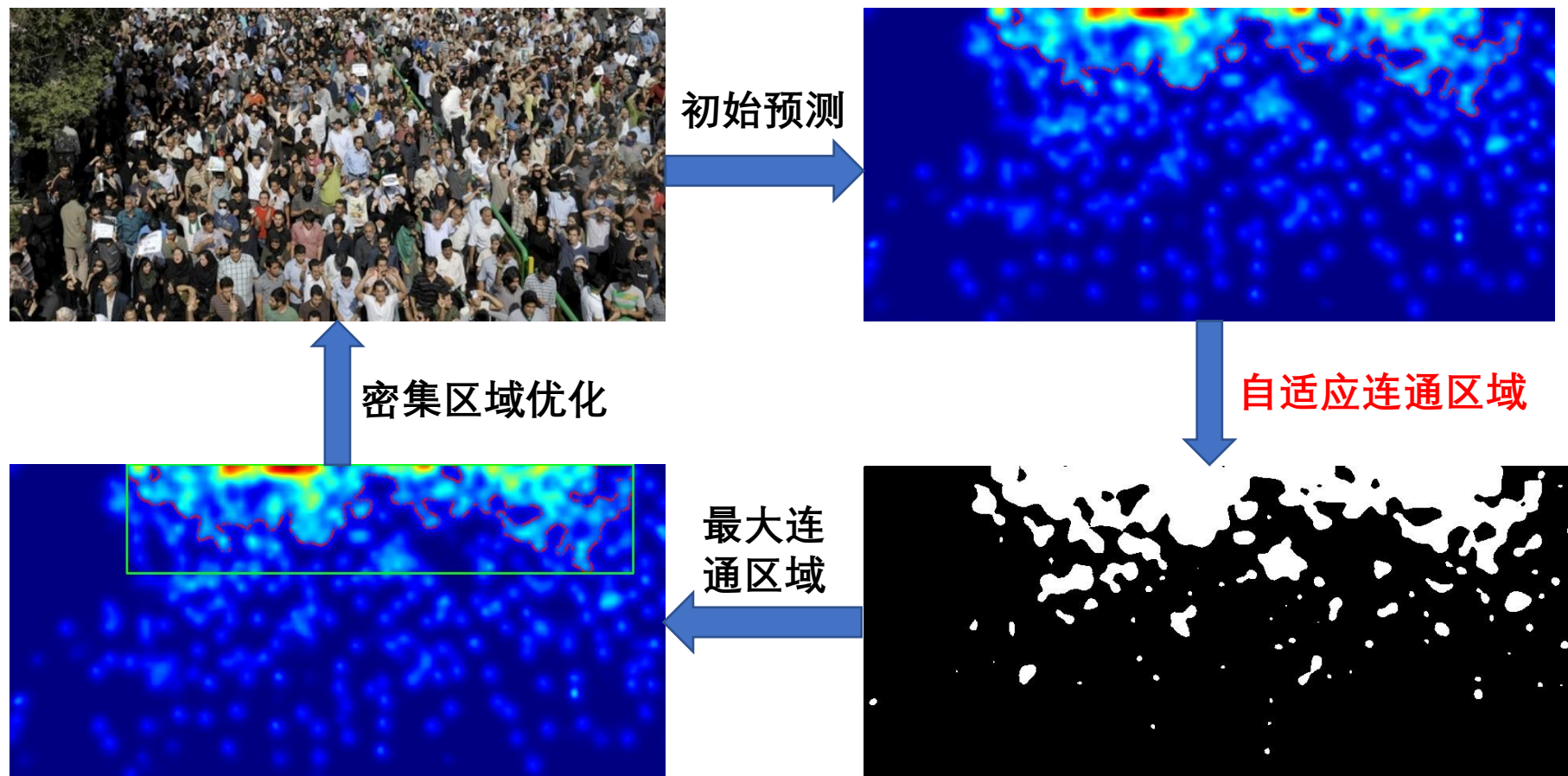
# Baseline

<http://www.vlrlab.net/>



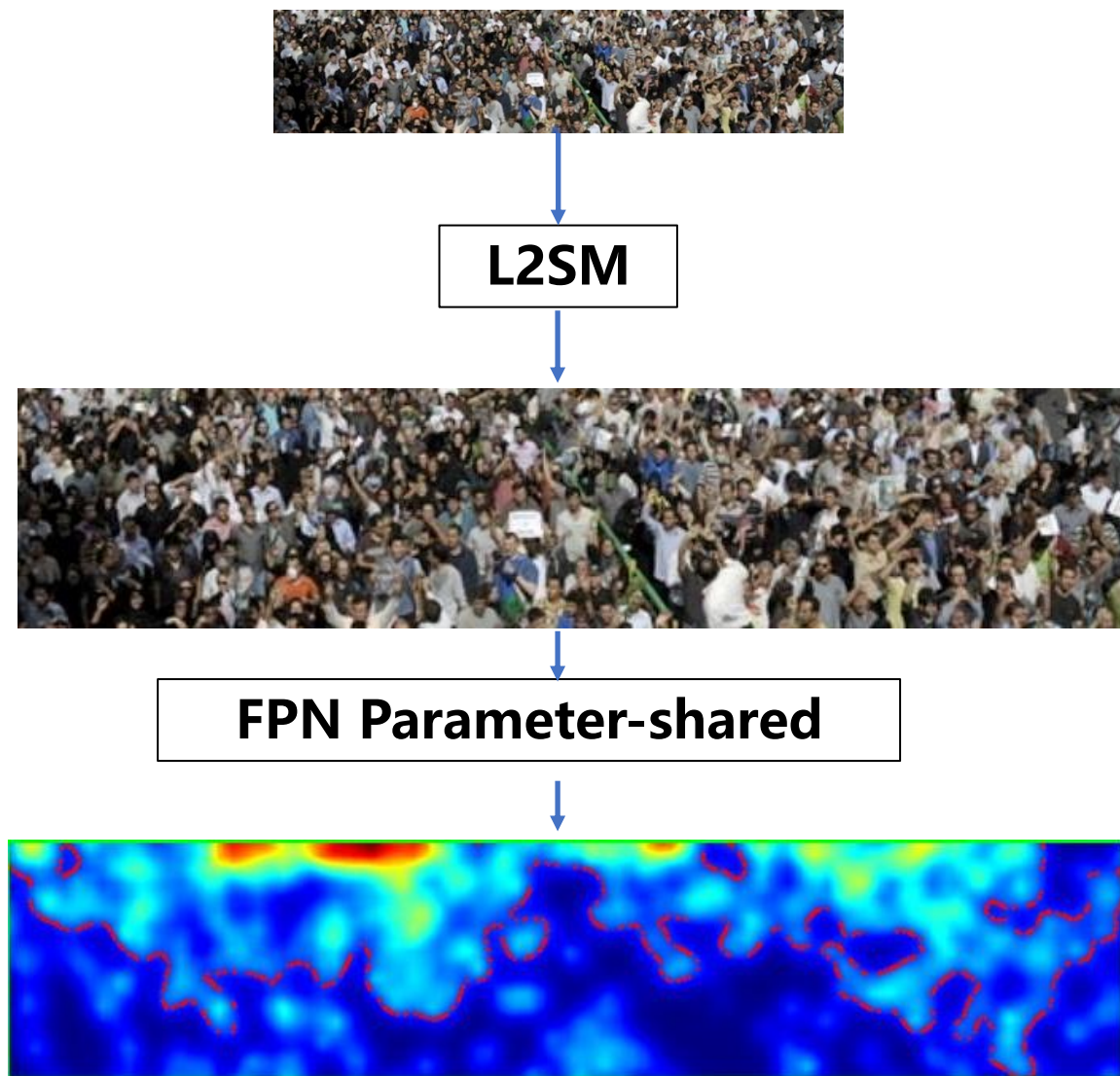
- 采用FPN生成初始预测密度图
- VGG16 作为Backbone
- 经过FPN得到的feature，即用于生成密度图，也用于后续训练放大网络（L2SM）

# 寻找密集区域——结合传统方法





# 对密集区域优化（放大再预测）



- 简而言之，希望所有的密集区域的密度能趋于一个**密度中心**。
- 此处的密度采用平均**KNN距离**衡量（k取1）
- 结合**Center Loss**，优化不同密集区域趋于一个密度中心
- 对密集区域放大，通过**共享**FPN生成新的密度图

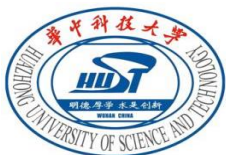
# 损失函数

$$L_m = ||D - \hat{D}||_2$$

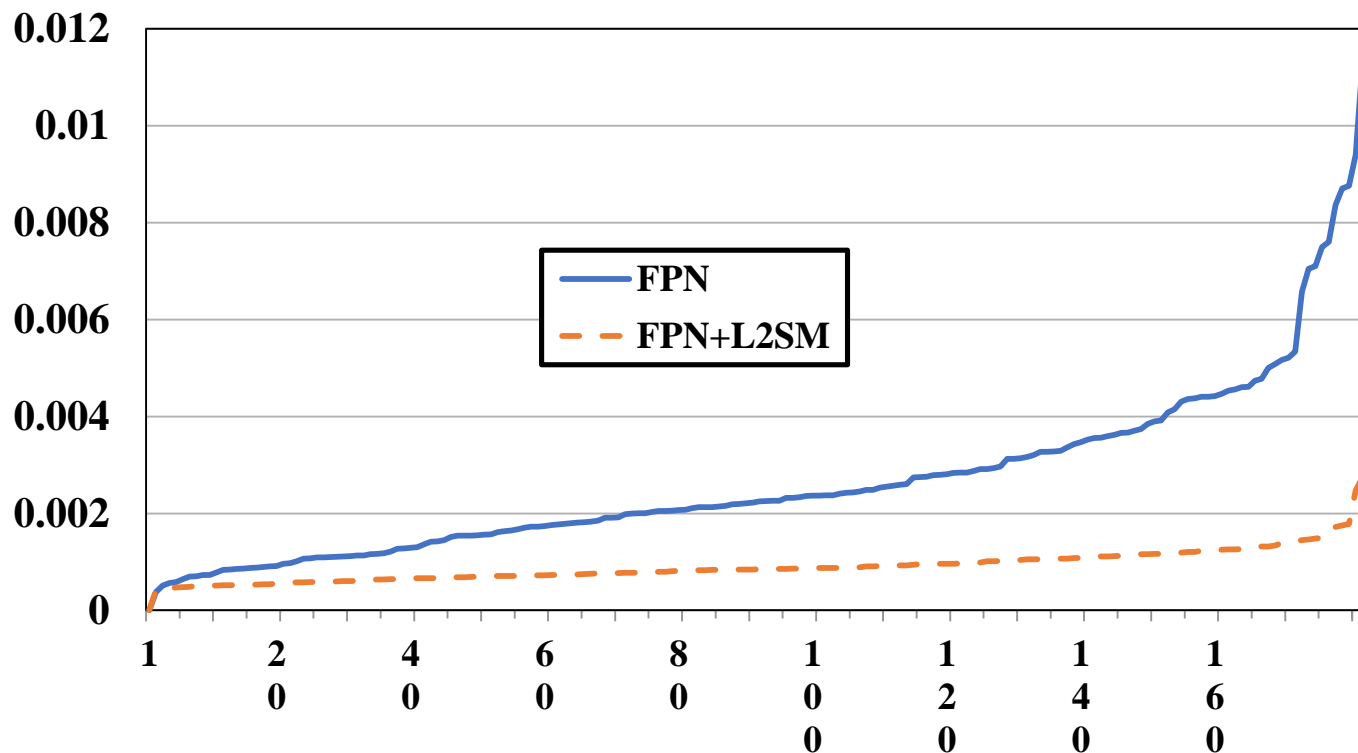
$$L_c = \frac{1}{2} \sum_{i=1}^M ||S_i \times r_i^2 - \overline{S}||_2^2, M \leq N$$

$$L_{total} = L_{m_{init}} + L_{m_{repre}} + \lambda_1 \times L_c$$

- $L_m$ 代表MSE Loss
- $L_c$ 代表center loss
- $\lambda$ 取0.1



# L2SM前后密度分布



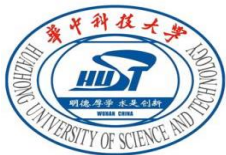
- Baseline中不同密集区域之间的密度差异**非常大**。
- 经过L2SM之后，不同密集区域的密度差异**变小**，符合现实中人喜欢把不同密集区域缩放到**相似**密度的特点。



# 结果

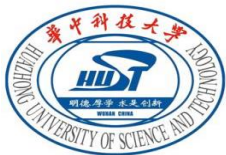
排名	用户名	score	
1	dklianghust1	0.09011	
2	pcalab4049	0.11172	
3	click00	0.11191	
4	commissarma2019	0.12088	
5	kongbangwa11	0.12266	
6	csmanlyt666	0.12966	
7	xuwei2838	0.19765	
8	yy798533133	0.24572	
9	fengjin2019	0.34845	
10	iseekseek58	0.39211	

- FPN + L2SM取得了**第一名**的成绩。
- 经过对比测试，采用所提出的L2S模块后结果会**提高**近2个点 (0.1098->0.09011)



# 实现细节

- 比赛数据集较大，普遍为1920 \* 1080的分辨率，易爆显存，将长边resize至1024，短边按照相同比例resieze。
- 数据增强：0.8-1.3倍随机缩放、50%概率水平翻转、50%概率增加椒盐噪声。实际测试效果：随机缩放>椒盐噪声>翻转
- 寻找联通区域阈值为density map均值
- 学习率1e-5，每50 epoch 衰减0.5倍



# 谢谢聆听!

<http://www.vlrlab.net/>

