



VGG-Embedded Adaptive Layer-Normalized Crowd Counting Net with Scale-Shuffling Modules

Dewen Guo, Jie Feng, Bingfeng Zhou

Peking University

CHN

Background

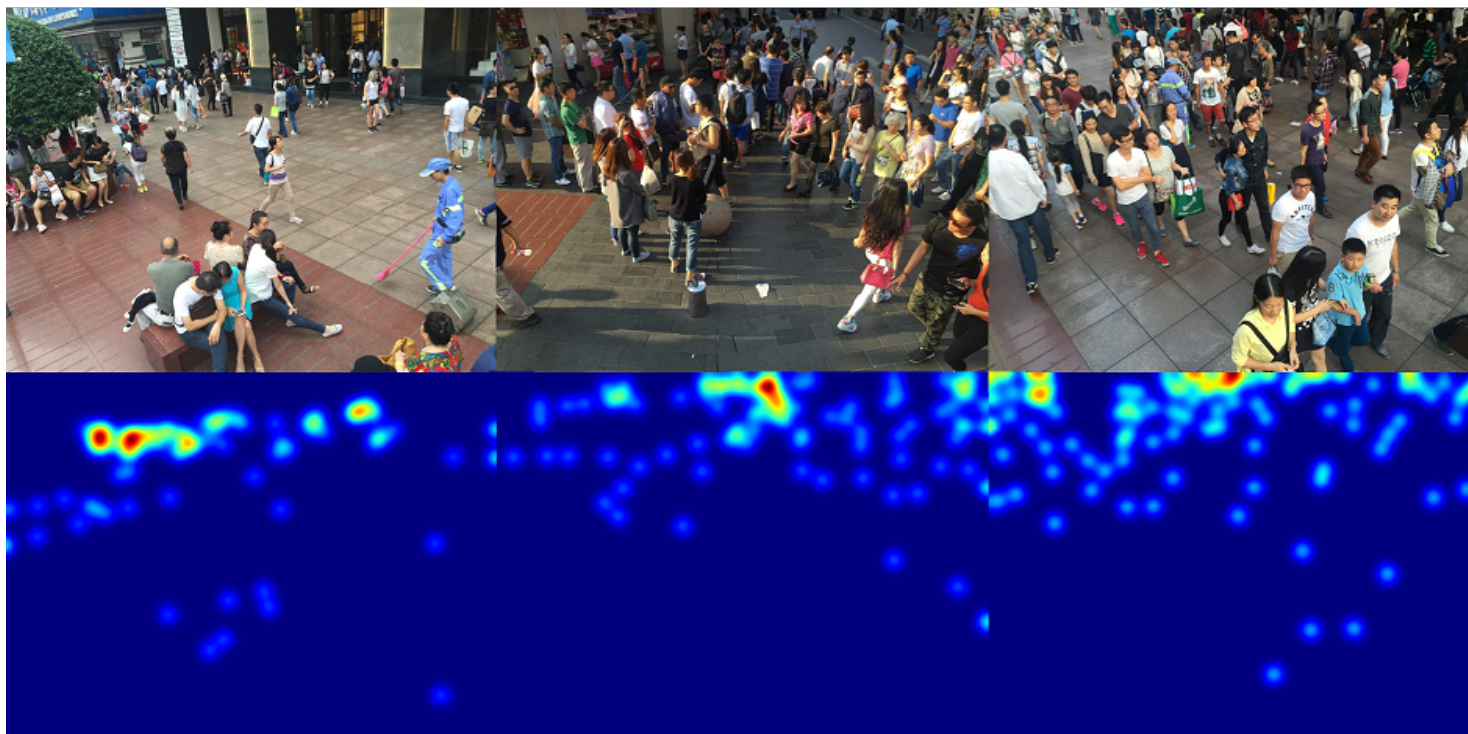


Figure from CSRNet

Background

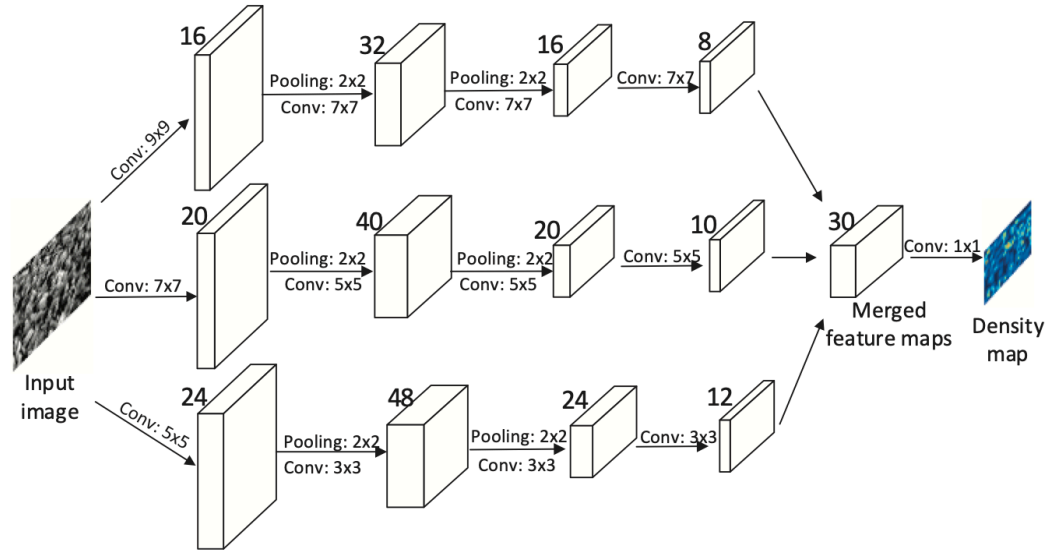


Figure from MCNN

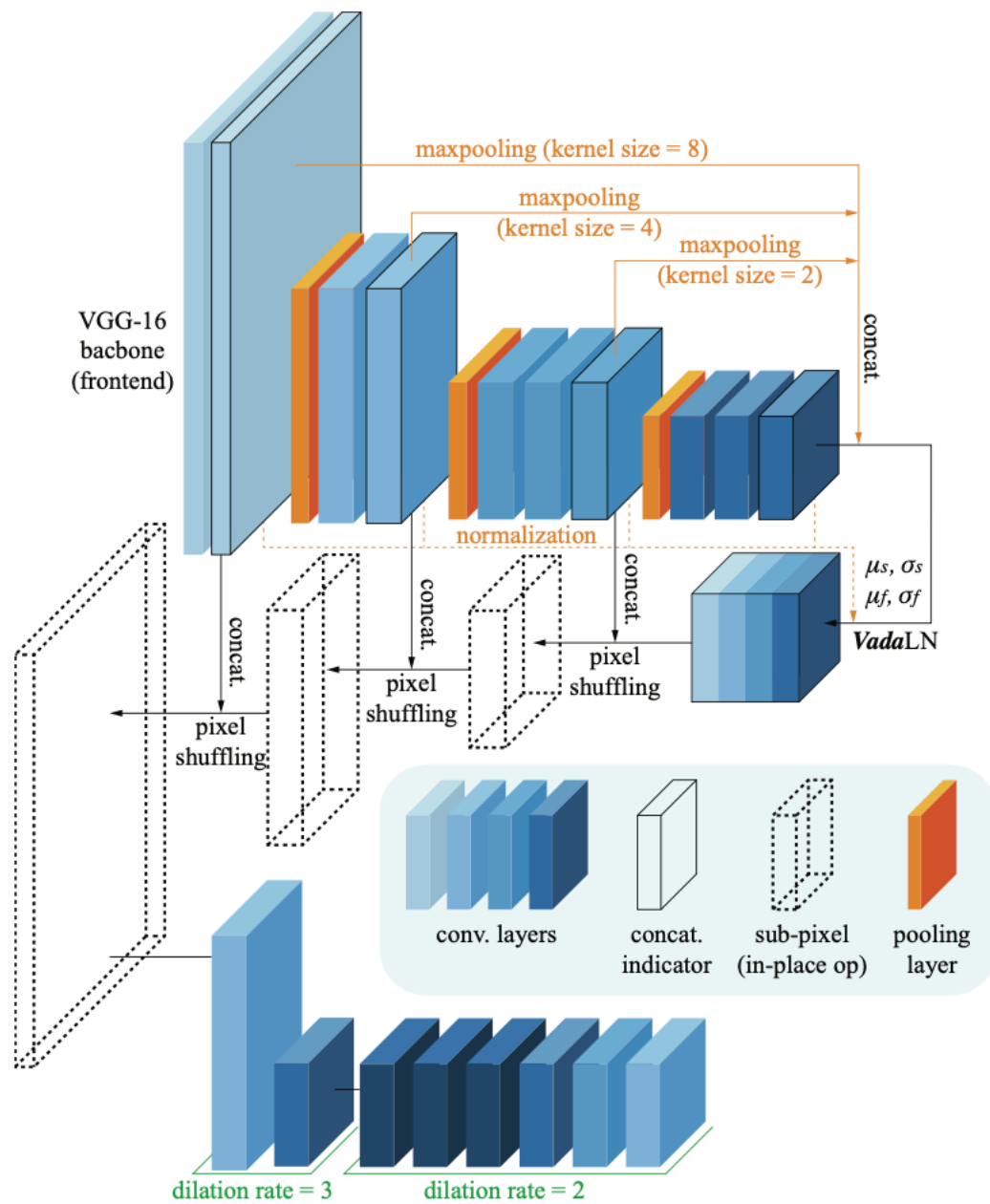
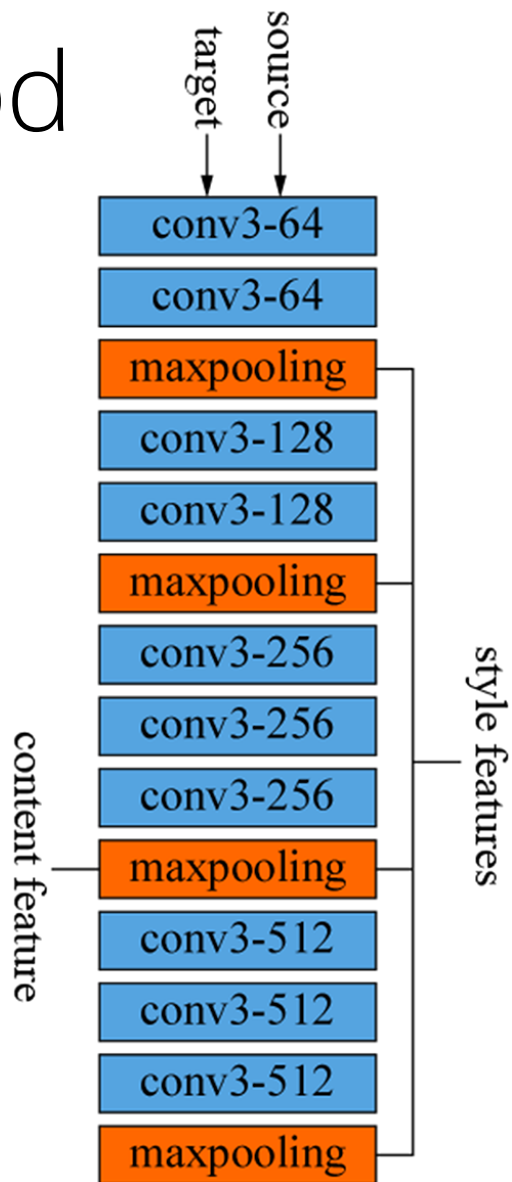
Configurations of CSRNet			
A	B	C	D
input(unfixed-resolution color image)			
front-end (fine-tuned from VGG-16)			
conv3-64-1			
conv3-64-1			
max-pooling			
conv3-128-1			
conv3-128-1			
max-pooling			
conv3-256-1			
conv3-256-1			
conv3-256-1			
max-pooling			
conv3-512-1			
conv3-512-1			
conv3-512-1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

Table from CSRNet

Background



Method



VadaLN

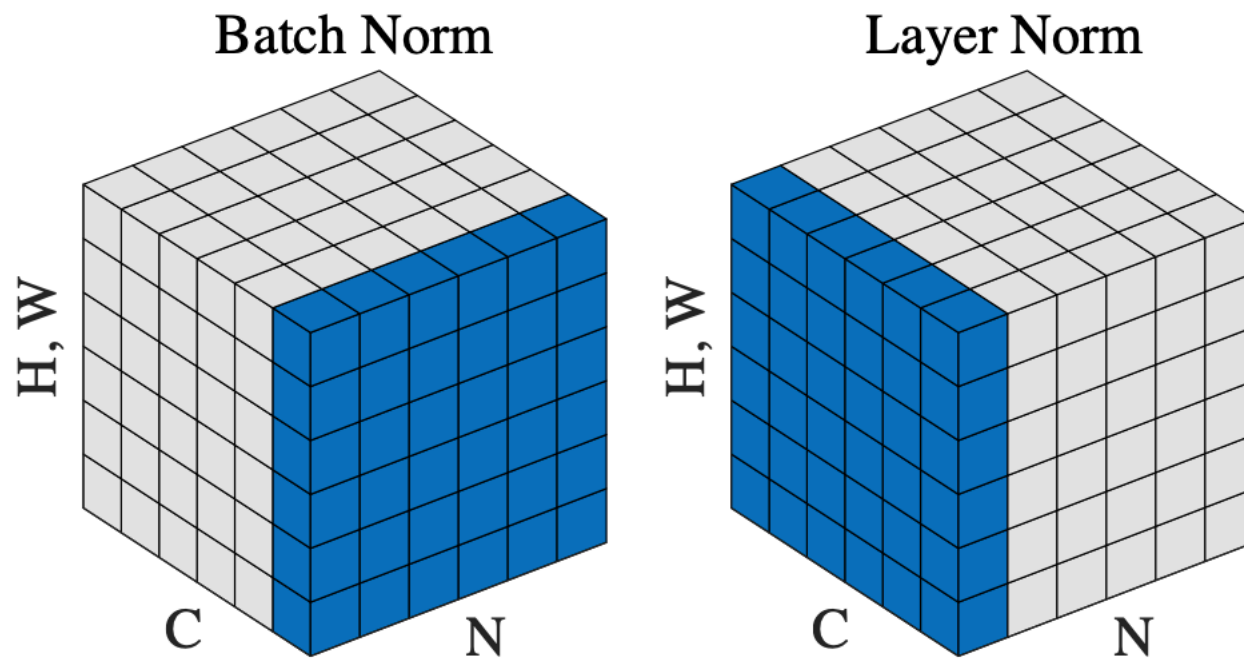


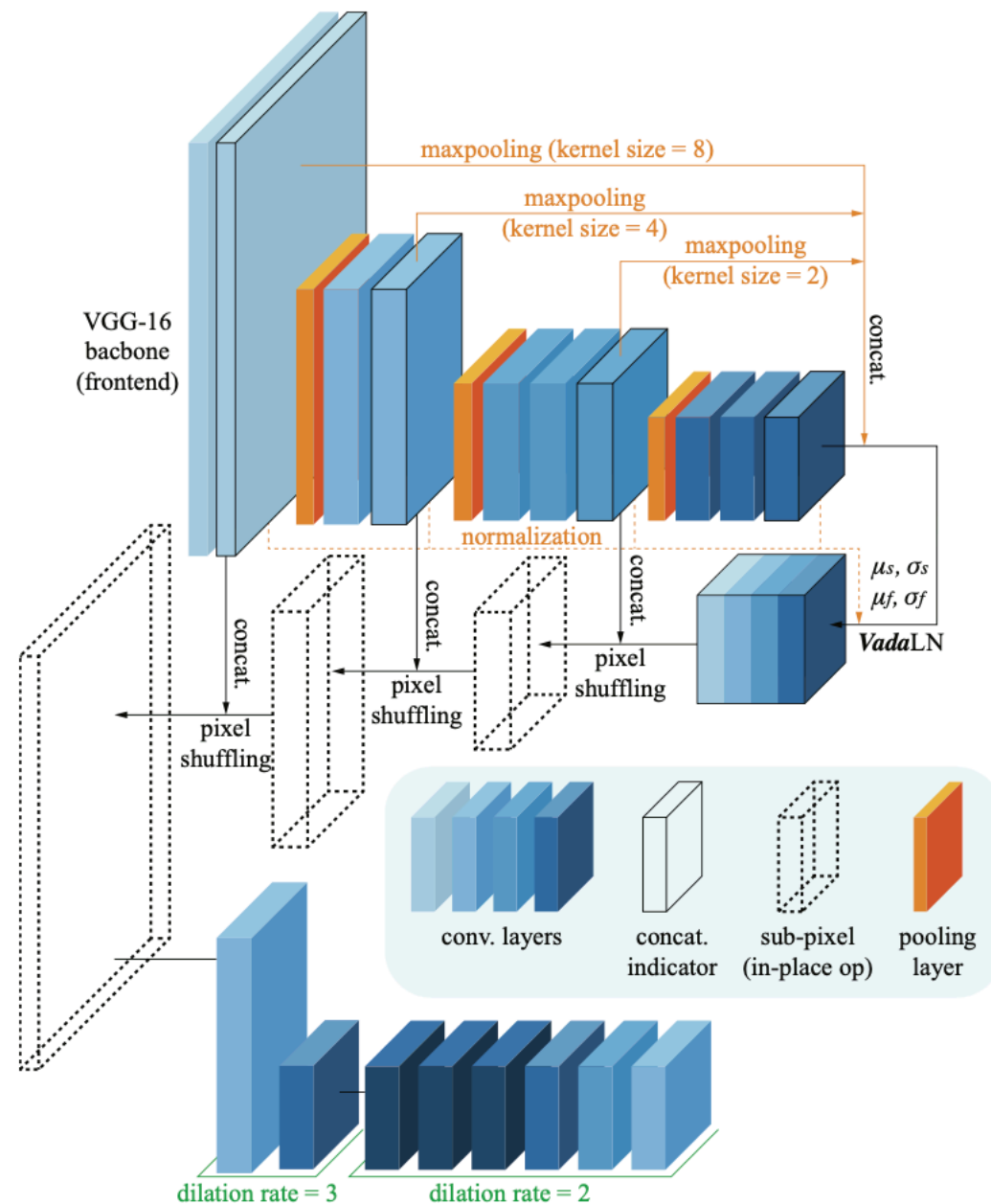
Figure from Group Normalization

VadaLN

$$\mu(x) = \frac{1}{H} \sum_{i=1}^H x_i,$$

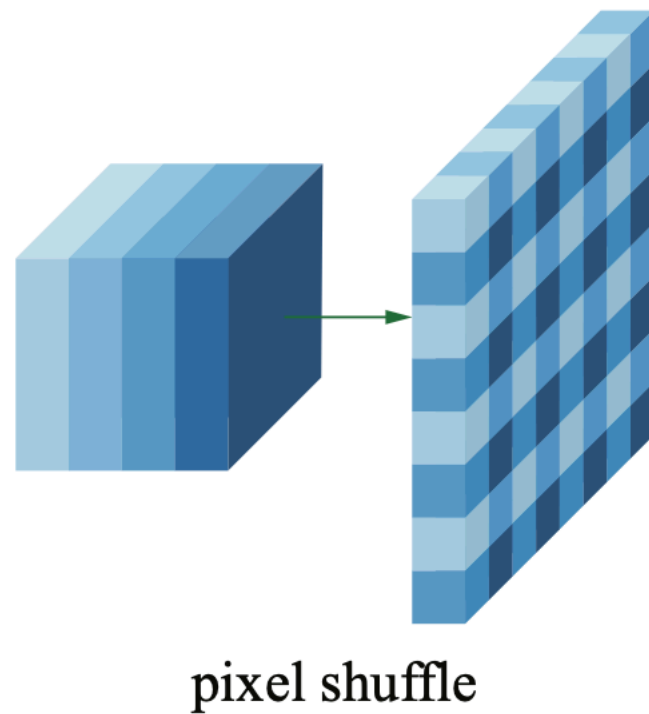
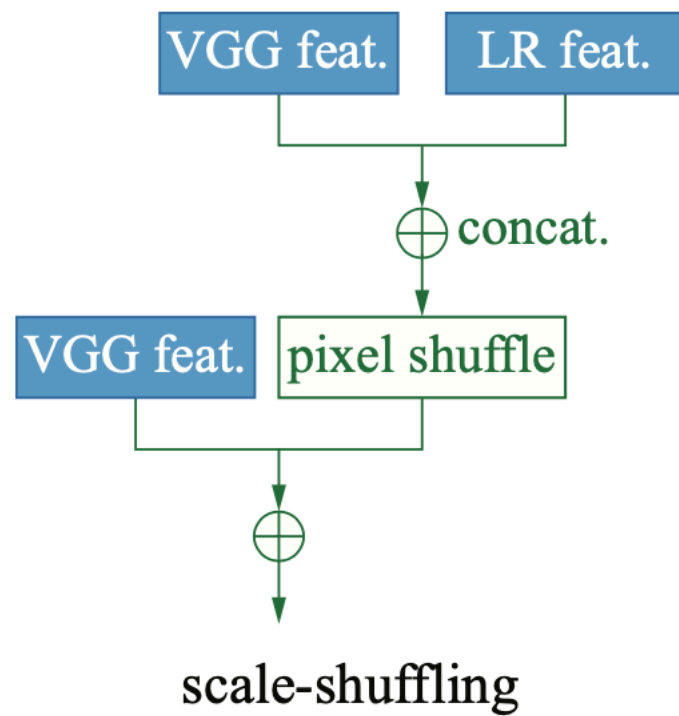
$$\sigma(x) = \sqrt{\frac{1}{H} \sum_{i=1}^H (x_i - \mu(x))^2},$$

$$\text{VadaLN}(x) = \sigma_f \frac{\text{VGG}_{16}(x) - \mu_s}{\sigma_s} + \mu_f.$$



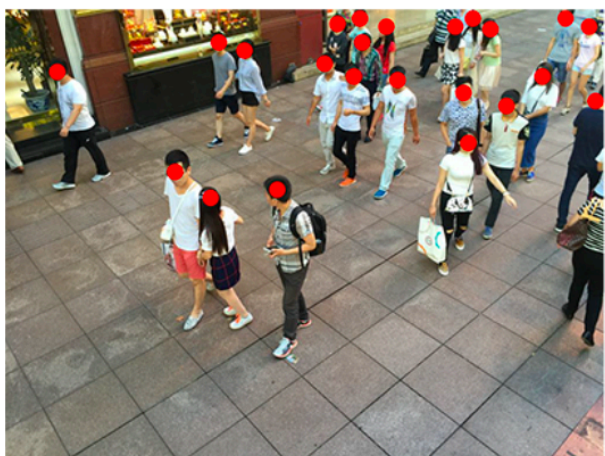
SSM

$$\text{conv}_{\text{sub-pixel}}(\text{feat.})_{x,y,c} = \text{feat.}_{x/r, y/r, c \cdot y + c \cdot x/r}$$

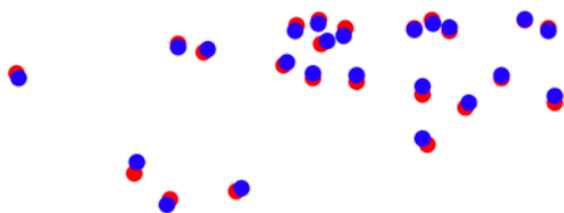


ACL

manual annotation 1



manual annotation 2

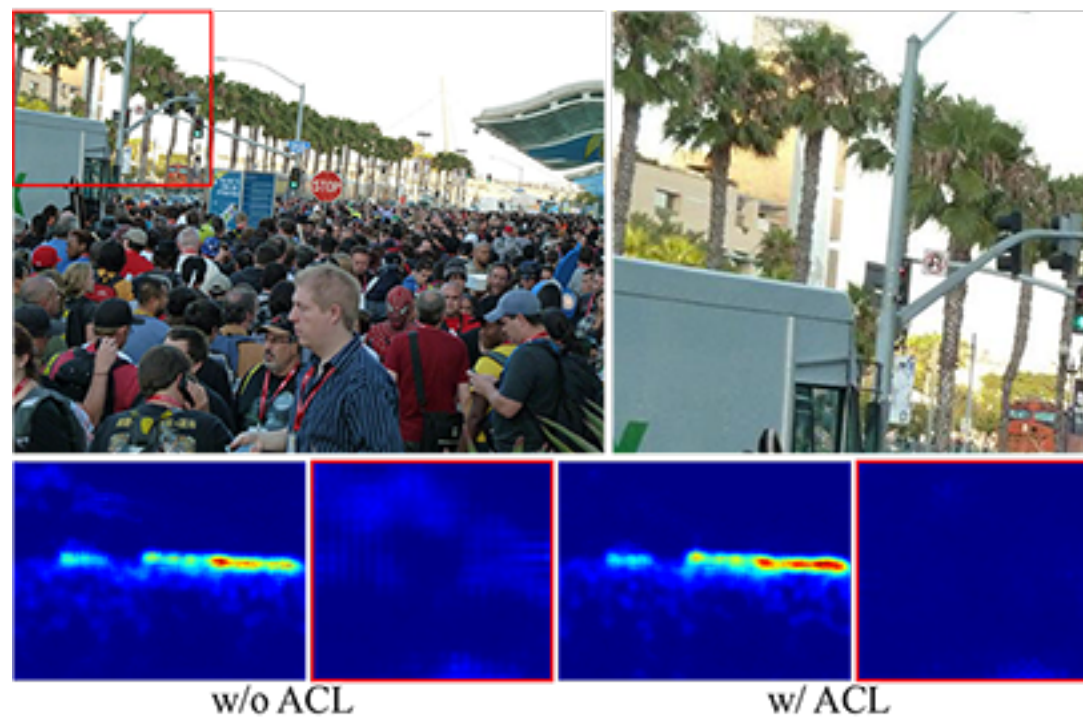
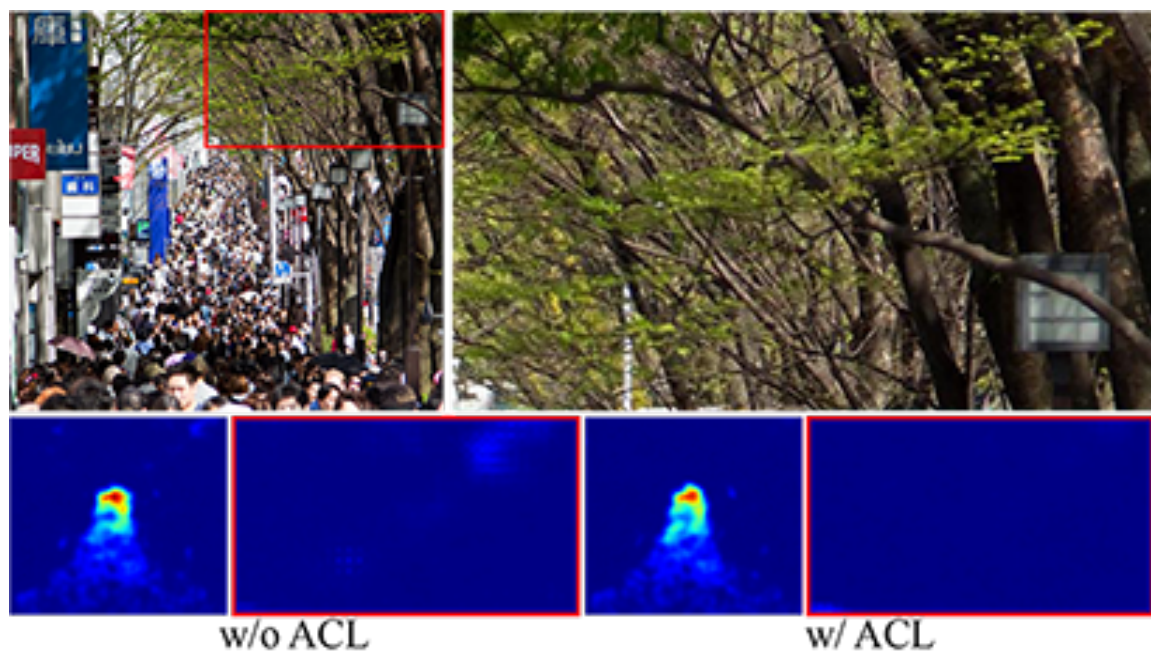


$$\text{CX}(x, y) = \text{CX}(X, Y) = \frac{1}{N} \sum_j \max_i \text{CX}_{ij},$$

$$\mathcal{L}_{\text{CX}}(x, y, l) = -\log (\text{CX} (\Phi^l(x), \Phi^l(y)))$$

$$\begin{aligned} \text{ACL}(x, y) = & -\lambda_1 \log (\text{CX} (\Phi^{\text{conv}_{3-3}}(x), \Phi^{\text{conv}_{3-3}}(y))) \\ & -\lambda_2 \log (\text{CX} (\Phi^{\text{conv}_{4-2}}(x), \Phi^{\text{conv}_{4-2}}(y))) \\ & + \lambda_3 \mathbb{E}_{x \sim p_{\text{pred.}}} [\log D_{\text{ACL}}^*(x)] \\ & + \lambda_3 \mathbb{E}_{x \sim p_{gt}} [\log (1 - D_{\text{ACL}}^*(x))] \end{aligned}$$

Results



Results

