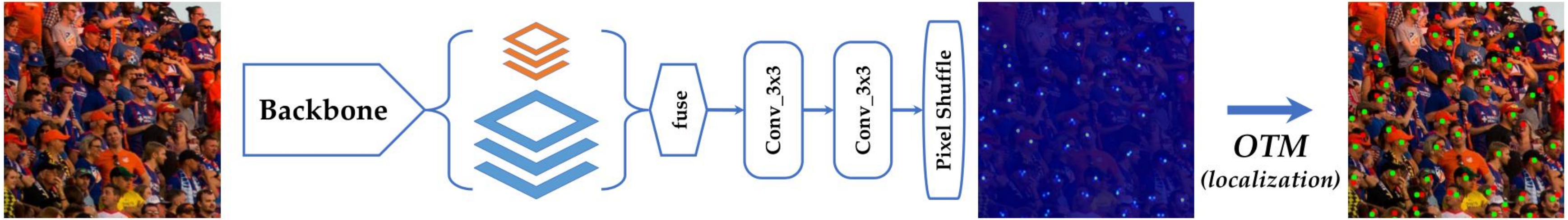
# Proximal Mapping Loss: Understanding Loss Functions in Crowd Counting & Localization

Wei Lin[1], Jia Wan[2], and Antoni B. Chan[1]

[1]Department of Computer Science, City University of Hong Kong
[2]School of Computer Science and Technology, Harbin Institute of Technology, Shenzhen
elonlin24@gmail.com, jiawan1998@gmail.com, abchan@cityu.edu.hk

# Supervised Crowd Counting

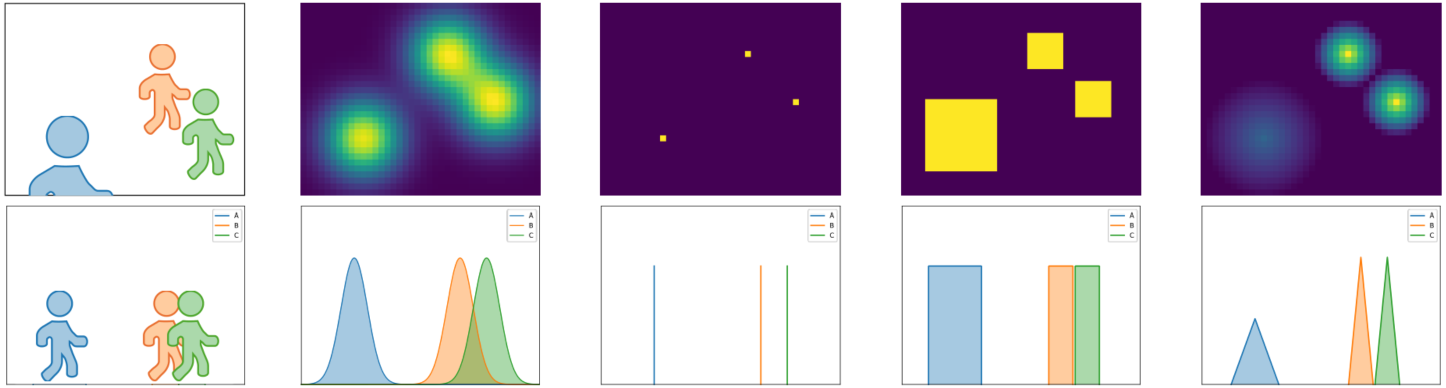



Supervised crowd counting is normally formulated as a regression task:

- **Input:** an image contain crowds;
- **Output:** a density map demonstrate the distribution and count of crowd in the input;
- **Ground Truth:** a point map in which a pixel with a value of one denotes a person's location.

- **OTM** is applied to regress the density map into a point map.
- **Applications:** video surveillance and public safety services, traffic congestion control, marine environmental monitoring

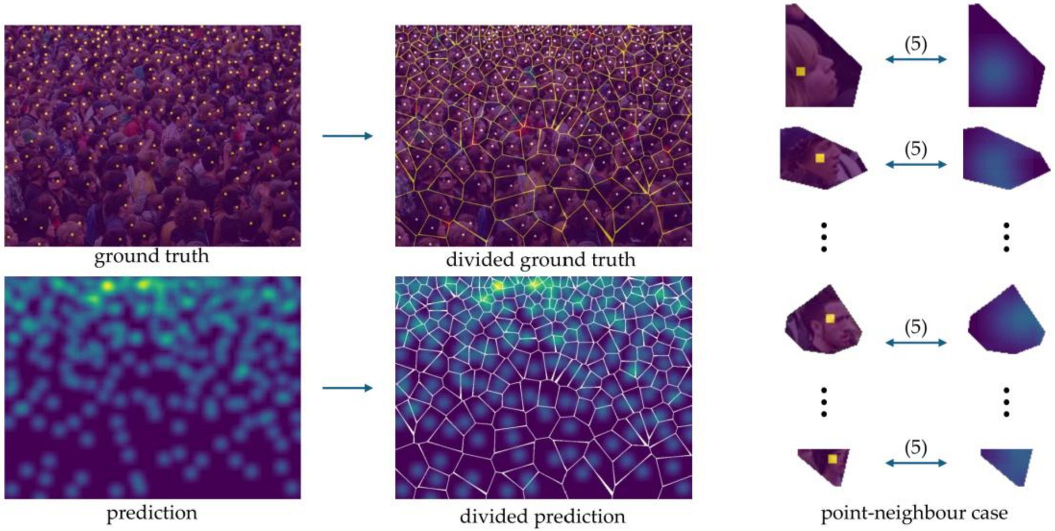
# Supervised Crowd Counting



(a) (b) (c) (d) (e)

(a) A synthetic input with three humans;
(b) Density regression with the intersection hypothesis (using a Gaussian prior), where one pixel may correspond to multiple objects;
(c) Point prediction;
(d) Head region segmentation;
(e) Density regression without the intersection hypothesis (proposed method), where one pixel corresponds to one object.

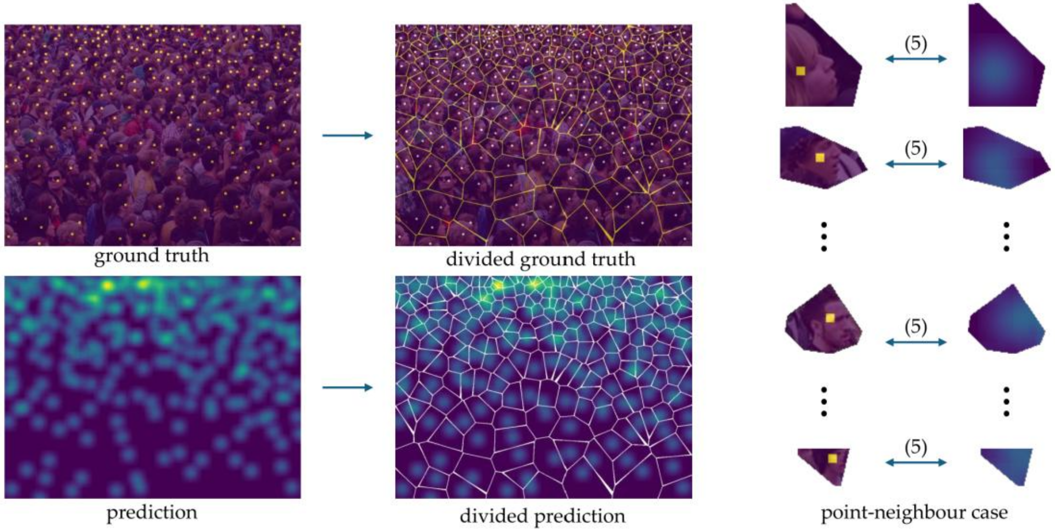
# Proximal Mapping Loss




**(I) Divide Stage**: assign each pixel to its nearest GT point **(II) Conquer Stage**: loss computation

- nearest neighbour is adopted to split the density map into multiple irregular patches without overlapping.
- In PML, the loss computation is divided into multiple simpler sub-problems, as each point-neighbour case can be handled independently.

# Proximal Mapping Loss




**(I) Divide Stage**: assign each pixel to its nearest GT point **(II) Conquer Stage**: loss computation

$$\mathcal{L}(\mathcal{A},\mathcal{B}) = \sum_{j=1}^{m} \tilde{\mathcal{L}}(\tilde{A}_j, \boldsymbol{b}_j), \quad \tilde{A}_j = \{(a_i, \boldsymbol{x}_i)\}_{i\in\mathcal{X}_j} \quad \boldsymbol{b}_j = (1, \boldsymbol{y}_j), \tag{1}$$

$$\mathcal{X}_j = \{\, i \mid \|\boldsymbol{x}_i - \boldsymbol{y}_j\|_2 \le \|\boldsymbol{x}_i - \boldsymbol{y}_k\|_2,\ \forall \boldsymbol{y}_k \in \mathcal{B} \,\}, \tag{2}$$

# Proximal Mapping Loss

By defining $\tilde{\boldsymbol{a}} = [\tilde{a}_i]_i^{\tilde{n}}$ constructed from $\tilde{\mathcal{A}}$, the objective inherited from GL is to minimize the transport $f(\tilde{\boldsymbol{a}}) = \boldsymbol{c}^\top \tilde{\boldsymbol{a}}$, where $\boldsymbol{c} = [c_i]_{i=1}^{\tilde{n}}$ measures the cost when moving a unit mass from $\boldsymbol{x}_i$ *to* $\boldsymbol{y}$.

***Proximal mapping:*** $\tilde{\boldsymbol{a}}_{t+1} \approx \operatorname{argmin}_{\boldsymbol{p}} \underbrace{f(\tilde{\boldsymbol{a}}_t) + \nabla f(\tilde{\boldsymbol{a}}_t)^\top (\boldsymbol{p} - \tilde{\boldsymbol{a}}_t)}_{\text{linear approximation of } f(\tilde{\boldsymbol{a}}_{t+1})} + \underbrace{\frac{\tau}{2}\|\boldsymbol{p} - \tilde{\boldsymbol{a}}_t\|^2}_{\text{regularizer}}$

$$\left.\begin{array}{r} \nabla f(\tilde{\boldsymbol{a}}_t) = \boldsymbol{c} \\ \xi \subseteq \{\boldsymbol{p} \mid \boldsymbol{p}^\top \mathbf{1} = 1, \boldsymbol{p} \in \mathbb{R}^{\tilde{n}}\} \end{array}\right\} \Rightarrow \mathcal{L}(\tilde{\mathcal{A}}, \boldsymbol{b}) = \min_{\boldsymbol{p} \in \xi}\ \boldsymbol{c}^\top \boldsymbol{p} + \frac{\tau}{2}\|\boldsymbol{p} - \tilde{\boldsymbol{a}}\|^2$$

$$\text{Bregman divergence } \mathcal{D}_\varphi(\boldsymbol{p}, \tilde{\boldsymbol{a}}) \Rightarrow \mathcal{L}(\tilde{\mathcal{A}}, \boldsymbol{b}) = \min_{\boldsymbol{p} \in \xi}\ \boldsymbol{p}^\top \boldsymbol{a} + \tau \mathcal{D}_\varphi(\boldsymbol{p}, \tilde{\boldsymbol{a}})$$

# Proximal Mapping Loss

$$\mathcal{L}(\tilde{A}, \boldsymbol{b}) = \min_{\boldsymbol{p}\in\xi} \ \boldsymbol{p}^\top \boldsymbol{a} + \tau \mathcal{D}_\varphi(\boldsymbol{p}, \tilde{\boldsymbol{a}})$$

| loss function | $\tau$ | $\mathcal{D}_\varphi(\boldsymbol{p}, \boldsymbol{a})$ | $\xi$ | $\boldsymbol{p}^*$ |
|---|---|---|---|---|
| L2 loss (Zhang et al., 2016) | $0$ | $\frac{1}{2}\lVert\boldsymbol{p}-\boldsymbol{a}\rVert^2$ | $\mathcal{N}(\mu\mid\Sigma)$ | $\mathcal{N}(0\mid\sigma\mathbf{1}_{2\times 2})$ |
| Bayesian loss (Ma et al., 2019) | $\frac{1}{\lvert\mathbf{1}^\top\boldsymbol{a}-1\rvert}$ | $\frac{1}{2}\lVert\boldsymbol{p}-\boldsymbol{a}\rVert^2$ | - | $\boldsymbol{a}-\frac{1}{\lvert\mathbf{1}^\top\boldsymbol{a}-1\rvert}\boldsymbol{c}+\eta$ |
| P2PNet (Song et al., 2021) | - | $\lVert\boldsymbol{p}-\boldsymbol{a}\rVert_1$ | $\delta(\cdot)$ | $\delta(\arg\min_j \boldsymbol{c}_j-\tau\boldsymbol{a}_j)$ |
| DM-Count (Wang et al., 2020a) | $\infty$ | $\mathtt{KL}(\boldsymbol{p}\mid\boldsymbol{a})$ | - | $\boldsymbol{a}/\lVert\boldsymbol{a}\rVert_1$ |

# PML & L2 loss

$$\mathcal{L}(\tilde{\mathcal{A}}, \boldsymbol{b}) = \min_{\boldsymbol{p}\in\xi}\ \boldsymbol{p}^\top \boldsymbol{a} + \tau \mathcal{D}_\varphi(\boldsymbol{p}, \tilde{\boldsymbol{a}}) \ \text{ with } \mathcal{D}_\varphi = \tfrac{1}{2}\|\boldsymbol{p} - \tilde{\boldsymbol{a}}\|^2$$

$$\mathcal{L}_2 = \min_{\boldsymbol{p}} \boldsymbol{c}^\top \boldsymbol{p} + \tfrac{\tau}{2}\|\boldsymbol{p} - \tilde{\boldsymbol{a}}\|_2^2, \quad s.t. \quad \boldsymbol{p}^\top \mathbf{1} = \sum_{i=0}^{\tilde{n}} p_i = 1.$$

$$\boldsymbol{p}^* = \tilde{\boldsymbol{a}} - \tfrac{1}{\tau_1}\boldsymbol{c} + \eta, \qquad \eta = \tfrac{1}{\tilde{n}}\left[1 - \left(\tilde{\boldsymbol{a}} - \tfrac{1}{\tau}\boldsymbol{c}\right)^\top \mathbf{1}\right]$$

Here $\eta$ takes the role as "filler" such that $\boldsymbol{p}^{*\top}\mathbf{1} = 1$.

$$\frac{\partial \mathcal{L}_2}{\partial \tilde{a}_i} = c_i - \tau\eta \quad \Rightarrow \quad \mathcal{L}_2 = \frac{\tau}{2}\|\boldsymbol{a} - \mathtt{detach}(\boldsymbol{p}^*)\|_2^2$$

# PML & L2 loss

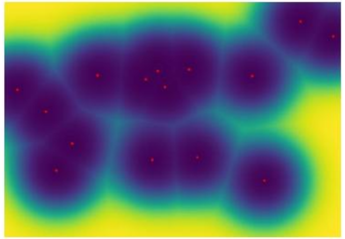
$$\mathcal{L}_2 = \frac{\tau}{2}\|\boldsymbol{a} - \texttt{detach}(\boldsymbol{p}^*)\|_2^2 \quad \text{where} \quad \boldsymbol{p}^* = \boldsymbol{a} - \tfrac{1}{\tau}\boldsymbol{c} + \eta$$

Dynamic L2 loss
$$\begin{cases} \lim_{\tau \to 0} \boldsymbol{p}^* &= \delta(\boldsymbol{y}) \\ \lim_{\tau \to \infty} \boldsymbol{p}^* &= \tilde{\boldsymbol{a}} + (1 - \boldsymbol{1}^\top \tilde{\boldsymbol{a}}) \end{cases}$$

Traditional L2 loss
$$\begin{cases} \tau = 0 & \\ \xi \subseteq \mathcal{N}(\boldsymbol{\mu}|\Sigma) & \Rightarrow \quad \boldsymbol{p}^{*\prime} \leftarrow \mathcal{N}(\boldsymbol{\mu} \mid (\boldsymbol{x} - \boldsymbol{\mu})^\top (\boldsymbol{x} - \boldsymbol{\mu})), \qquad \boldsymbol{\mu} = \boldsymbol{x}^\top \boldsymbol{p}^* \\ \Sigma \succcurlyeq \sigma^2 \mathbf{I}_{2\times 2} & \Rightarrow \quad \boldsymbol{p}^{*\prime} \leftarrow \mathcal{N}(\boldsymbol{y}_j | \sigma^2 \mathbf{I}_{2\times 2}) \end{cases}$$

# PML & Bayesian Loss

(a) image & density map



(b) Bayesian loss

$$\mathcal{L}_b = \sum_{i=1}^{n} q(\boldsymbol{y}_0|\boldsymbol{x}_i)a_i + \sum_{j=1}^{m} \left|\sum_{i=1}^{n} q(\boldsymbol{y}_j|\boldsymbol{x}_i)a_i - 1\right|$$

$$\Rightarrow \quad \mathcal{L}_b = \underbrace{\boldsymbol{q}^\top \boldsymbol{a}}_{\text{background loss}} + \underbrace{|\boldsymbol{a}^\top \mathbf{1} - 1|}_{\text{count loss}},$$

- **count loss** forces the sum of $\boldsymbol{a}$ to be close to 1;
- **background loss** forces the distribution of a to be close to $\delta(\boldsymbol{y})$;

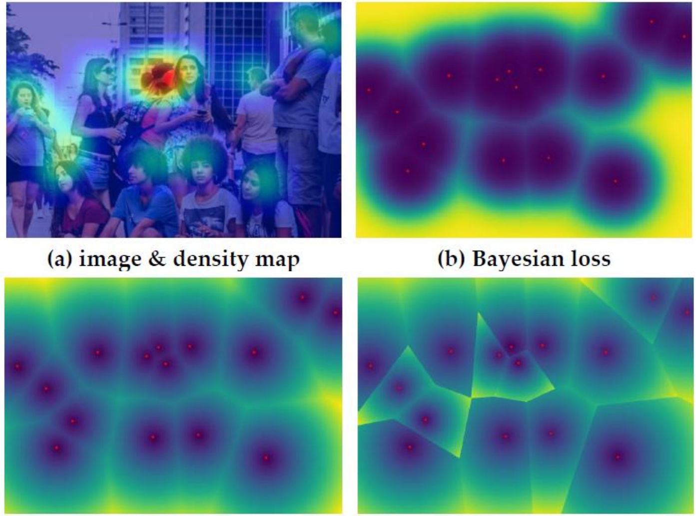
$$\boldsymbol{q}_{[i]} = q(\boldsymbol{y}_0|\boldsymbol{x}_i) = \frac{q(\boldsymbol{x}_i|\boldsymbol{y}_0)}{q(\boldsymbol{x}_i|\boldsymbol{y}_0)+1}, \qquad q(\boldsymbol{x}_i|\boldsymbol{y}_0) \stackrel{\text{def.}}{=\!=} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(d-c_i)^2}{2\sigma^2}\right]$$

$$\boldsymbol{q}_{[i]} \propto q(\boldsymbol{x}_i|\boldsymbol{y}_0) \quad \Rightarrow \quad \boldsymbol{q}_{[i]} > \boldsymbol{q}_{[j]} \text{ if } \boldsymbol{c}_i > \boldsymbol{c}_j$$

# PML & Bayesian Loss

$$\mathcal{L}(\tilde{A}, \boldsymbol{b}) = \min_{\boldsymbol{p}\in\xi} \ \boldsymbol{p}^\top \boldsymbol{a} + \tau \mathcal{D}_\varphi(\boldsymbol{p}, \tilde{\boldsymbol{a}}) \quad \text{with} \quad \mathcal{D}_\varphi = \tfrac{1}{2}\|\boldsymbol{p} - \tilde{\boldsymbol{a}}\|^2$$

$$\Rightarrow \quad \boldsymbol{p}^* = \tilde{\boldsymbol{a}} - \tfrac{1}{\tau_1}\boldsymbol{c} + \eta, \qquad \eta = \tfrac{1}{\tilde{n}}\left[1 - \left(\tilde{\boldsymbol{a}} - \tfrac{1}{\tau}\boldsymbol{c}\right)^\top \mathbf{1}\right]$$



(a) image & density map (b) Bayesian loss

(c) generalized loss (d) proximal mapping loss

$$\frac{\partial \mathcal{L}_2}{\partial \tilde{a}_i} = c_i - \tau\eta = c_i - \frac{1}{\tilde{n}}\boldsymbol{c}^\top \mathbf{1} + \frac{\tau}{\tilde{n}}(\tilde{\boldsymbol{a}}^\top \mathbf{1} - 1)$$

$$\Rightarrow \quad \mathcal{L}_2 = \underbrace{(\boldsymbol{c} - \bar{c})^\top \boldsymbol{a}}_{\text{background loss}} + \frac{\tau}{2\tilde{n}} \underbrace{\left(\tilde{\boldsymbol{a}}^\top \mathbf{1} - 1\right)^2}_{\text{count loss}},$$

# PML & Bayesian Loss

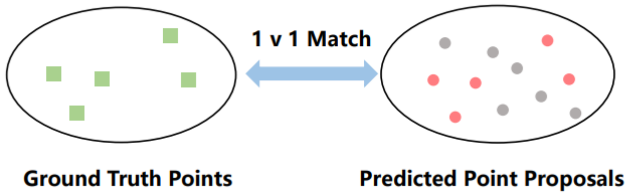
$$\mathcal{L}_2 = \underbrace{(\boldsymbol{c} - \bar{c})^\top \boldsymbol{a}}_{\text{background loss}} + \frac{\tau}{2\tilde{n}} \underbrace{\left(\tilde{\boldsymbol{a}}^\top \mathbf{1} - 1\right)^2}_{\text{count loss}},$$

$$\tau = \mathtt{detach}\left(2/|\mathbf{1}^\top \tilde{\boldsymbol{a}} - 1|\right) \begin{cases} \mathcal{L}_b' = (\boldsymbol{c} - \bar{c})\boldsymbol{a} + \frac{1}{\tilde{n}} \left|\mathbf{1}^\top \tilde{\boldsymbol{a}} - 1\right| \\ \Updownarrow \\ \boldsymbol{p}^* = \boldsymbol{a} - \left(\frac{1}{2}|\mathbf{1}^\top \tilde{\boldsymbol{a}} - 1|\right) \boldsymbol{c} + \eta \end{cases}$$

L1 norm is robust to noise annotation:

- If the predicted count is close to 1, $\boldsymbol{p}^*$ will be close to the distribution of $\boldsymbol{a}$;
- If the count is far from GT, $\boldsymbol{p}^*$ will be close to the distribution of $\delta(\boldsymbol{y})$

# PML & P2PNet




The matching is implemented via Hungarian algorithm with the cost matrix:

$$\mathcal{D}(\mathcal{P}, \hat{\mathcal{P}}) = \left(\tau \,||p_i - \hat{p}_j||_2 - \hat{c}_j\right)_{i \in N, j \in M}$$

$$\mathcal{L}_{p2p} = \min_{\boldsymbol{p} \in \xi_{p2p}} \boldsymbol{c}^\top \boldsymbol{p} + \tau \|\boldsymbol{p} - \tilde{\boldsymbol{a}}\|_1, \qquad \xi_{p2p} = \{\delta(\boldsymbol{x}_i) | \boldsymbol{x}_i \in \tilde{\mathcal{A}}\}$$

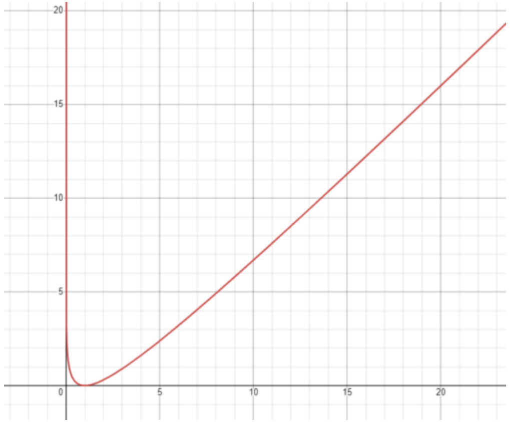
$$\begin{aligned} \mathcal{L}_{p2p} &= \min_{\boldsymbol{p} \in \xi_{p2p}} \boldsymbol{c}^\top \boldsymbol{p} + \tau \left[\boldsymbol{p}^\top (1 - \boldsymbol{a}) + (1 - \boldsymbol{p})^\top \tilde{\boldsymbol{a}}\right] \\ &= \min_{\boldsymbol{p} \in \xi_{p2p}} \underbrace{(\boldsymbol{c} - 2\tau\tilde{\boldsymbol{a}})^\top \boldsymbol{p}}_{\text{matching strategy}} + \underbrace{\tau(\boldsymbol{p}^\top 1 + \tilde{\boldsymbol{a}}^\top 1)}_{\text{constant } \tau(1+\|\tilde{\boldsymbol{a}}\|_1)}, \end{aligned}$$

$$\boldsymbol{p}^* = \delta(\operatorname{argmin}_j c_j - 2\tau\tilde{a}_j)$$

# PML & DMC

$$\mathcal{L}_{dmc} = \underbrace{\boldsymbol{c}^\top \frac{\boldsymbol{a}}{\|\boldsymbol{a}\|_1}}_{\text{OT loss}} + \tau \underbrace{|\boldsymbol{1}^\top \boldsymbol{a} - 1|}_{\text{count loss}}$$

$$\varphi(\boldsymbol{x}) = \boldsymbol{x}^\top \log \boldsymbol{x} - \boldsymbol{x}^\top \boldsymbol{1} \quad \Rightarrow \quad \mathcal{D}_\varphi(\boldsymbol{p}, \boldsymbol{a}) = \boldsymbol{p}^\top \log \tfrac{\boldsymbol{p}}{\boldsymbol{a}} - \boldsymbol{p}^\top \boldsymbol{1} + \boldsymbol{a}^\top \boldsymbol{1}.$$




$$\Rightarrow \mathcal{L}_{dmc} = \min_{\boldsymbol{p}} \boldsymbol{c}^\top \boldsymbol{p} + \tau(\boldsymbol{p}^\top \log \tfrac{\boldsymbol{p}}{\boldsymbol{a}} - \boldsymbol{p}^\top \boldsymbol{1} + \boldsymbol{a}^\top \boldsymbol{1}) \quad s.t. \quad \boldsymbol{p}^\top \boldsymbol{1} = 1,$$

$$\Rightarrow p_i^* = \frac{a_i \exp(-c_i/\tau)}{\eta}, \qquad \eta = \sum\nolimits_j a_j \exp(-c_i/\tau).$$

$\eta$ also serves as "filler", ensuring the sum of elements in $\boldsymbol{p}^*$ equals 1.

$$\tau \to \infty \quad \Rightarrow \quad \boldsymbol{p} = \frac{\boldsymbol{a}}{\|\boldsymbol{a}\|_1} \quad \Rightarrow \quad \mathcal{L}_{dmc}^{(\varphi)} = \underbrace{\boldsymbol{c}^\top \frac{\boldsymbol{a}}{\|\boldsymbol{a}\|_1}}_{\text{OT loss}} + \tau \underbrace{\left(\boldsymbol{1}^\top \boldsymbol{a} - \log\left(\boldsymbol{1}^\top \boldsymbol{a}\right) - 1\right)}_{\text{count loss}}$$

# Experiments

| METHOD | (backbone) | ShTech A MAE | ShTech A MSE | ShTech B MAE | ShTech B MSE | UCF-QNRF MAE | UCF-QNRF MSE | JHU ++ MAE | JHU ++ MSE | NWPU MAE | NWPU MSE |
|---|---|---|---|---|---|---|---|---|---|---|---|
| MCNN (Zhang et al., 2016) | | 232.5 | 714.6 | 110.2 | 173.2 | 277.0 | 426.0 | 188.9 | 483.4 | 232.5 | 714.6 |
| CSRNet (Li et al., 2018) | (VGG-16) | 68.2 | 115.0 | 10.6 | 16.0 | 110.6 | 190.1 | 85.9 | 309.2 | 121.3 | 387.8 |
| SFCN (Wang et al., 2019) | (ResNet-101) | 64.8 | 107.5 | 7.6 | 13.0 | 102.0 | 171.4 | 77.5 | 297.6 | 105.7 | 424.1 |
| BL (Ma et al., 2019) | (VGG-19) | 62.8 | 101.8 | 7.7 | 12.7 | 88.7 | 154.8 | 75.0 | 299.9 | 105.4 | 454.2 |
| KDMG (Wan et al., 2020) | (VGG-19) | 63.8 | 99.2 | 7.8 | 12.7 | 99.5 | 173.0 | 69.7 | 268.3 | 100.5 | 415.5 |
| DMC (Wang et al., 2020a) | (VGG-19) | 59.7 | 95.7 | 7.4 | 11.8 | 85.6 | 148.3 | 68.4 | 283.3 | 88.4 | 357.6 |
| NoiseCC (Wan & Chan, 2020) | (VGG-19) | 61.9 | 99.6 | 7.4 | 11.3 | 85.8 | 150.6 | 67.7 | 258.5 | 96.9 | 534.2 |
| P2PNet (Song et al., 2021) | (VGG-16bn) | 52.7 | 85.6 | 6.3 | 9.9 | 85.3 | 154.5 | - | - | 77.4 | 362.0 |
| UOTCC (Ma et al., 2021) | (VGG-19) | 58.1 | 95.9 | 6.5 | 10.2 | 83.3 | 142.3 | 60.5 | 252.7 | 87.8 | 387.5 |
| GL (Wan et al., 2021) | (VGG-19) | 61.3 | 95.4 | 7.3 | 11.7 | 84.3 | 147.5 | 59.9 | 259.5 | 79.3 | 346.1 |
| ChfL (Shu et al., 2022) | (VGG-19bn) | 57.5 | 94.3 | 6.9 | 11.0 | 80.3 | 137.6 | 57.0 | 235.7 | 76.8 | 343.0 |
| PET (Liu et al., 2023) | (VGG-16bn) | **49.3** | **78.8** | 6.2 | 9.7 | 79.5 | 144.3 | 58.5 | 238.0 | 74.4 | 328.5 |
| STEERER (Han et al., 2023) | (HRNet) | 54.5 | 86.9 | 5.8 | 8.5 | 74.3 | 128.3 | 54.3 | 238.1 | **63.7** | 309.3 |
| PML (ours) | (VGG-16bn) | 50.6 | 80.7 | 6.1 | 9.7 | 79.5 | 142.7 | 58.9 | 249.6 | 75.7 | 353.1 |
| PML (ours) | (VGG-19) | 55.5 | 89.0 | 6.0 | 9.3 | 76.6 | 132.2 | 57.4 | **227.4** | 73.6 | 338.6 |
| PML (ours) | (HRNet) | 52.3 | 84.7 | **5.4** | **8.2** | **73.2** | **127.5** | **52.6** | 230.8 | 63.8 | **306.9** |

Table 2: Comparison of our PML with recent crowd counting methods.

# Experiments

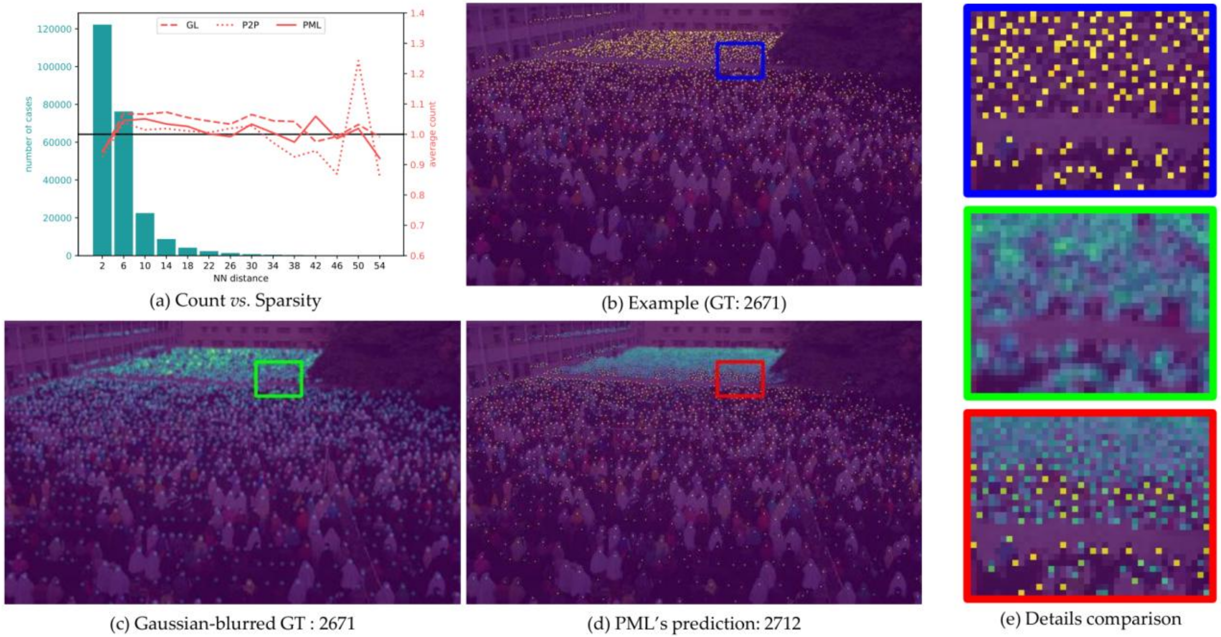



Figure 7: The performance difference of PML when handling sparse and dense crowds.

# Experiments

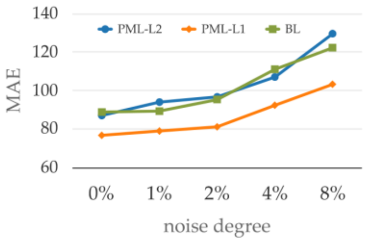
| | F1-meas. | Prec. | Rec. |
|---|---|---|---|
| RAZNet | 0.599 | 0.666 | 0.543 |
| GL+LM | 0.660 | 0.800 | 0.562 |
| GL+OTM | 0.683 | 0.710 | 0.658 |
| P2PNet | 0.729 | 0.676 | 0.685 |
| PET | 0.742 | 0.752 | 0.732 |
| STEERER+LM | 0.770 | **0.814** | 0.730 |
| PML(VGG-19)+OTM | 0.735 | 0.776 | 0.698 |
| PML(HRNet)+OTM | **0.802** | 0.809 | **0.795** |
| PML(HRNet)+LM | 0.790 | 0.803 | 0.777 |

Table 3: Localization on NWPU-Crowd.

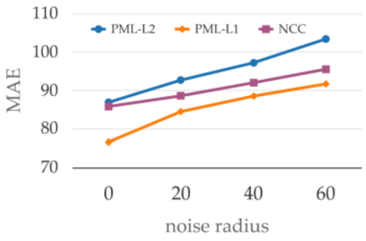
| | MCNN | | CSRNet | | VGG19 | | HRNet | |
|---|---|---|---|---|---|---|---|---|
| | MAE | MSE | MAE | MSE | MAE | MSE | MAE | MSE |
| L2 Loss | 186.4 | 283.6 | 110.6 | 190.1 | 98.7 | 176.1 | 92.03 | 157.49 |
| BL | 190.6 | 272.3 | 107.5 | 184.3 | 88.8 | 154.8 | 85.52 | 149.55 |
| NoiseCC | 177.4 | 259.0 | 96.5 | 163.3 | 85.8 | 150.6 | - | - |
| DMC | 176.1 | 263.3 | 103.6 | 180.6 | 85.6 | 148.3 | 82.07 | 144.84 |
| GL | 142.8 | 227.9 | 92.0 | 165.7 | 84.3 | 147.5 | 78.37 | 140.23 |
| GCFL | - | - | 83.0 | 139.8 | 80.3 | 137.6 | - | - |
| PML (ours) | **138.9** | **215.7** | **82.1** | **139.0** | **77.6** | **132.8** | **73.17** | **127.45** |

Table 4: Comparison of loss functions and backbones on UCF-QNRF dataset.

# Experiments

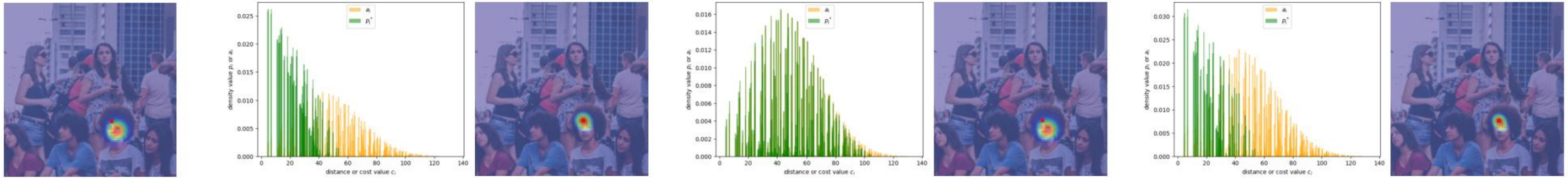



(a) Adding noise following BL




(b) Adding noise following NCC


(c) MAE of PML with L2- and L1-norm during training

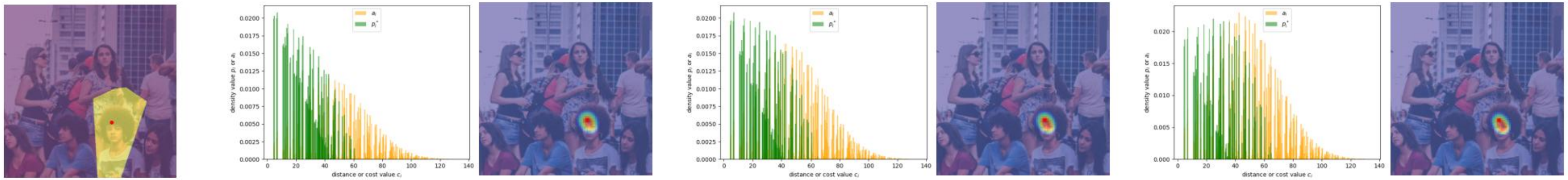


(a) Image & prediction

(c) Under estimation (L1-norm)

(d) Well estimation (L1-norm)

(e) Over estimation (L1-norm)



(b) GT & neighbor pixels

(f) Under estimation (L2-norm)

(g) Well estimation (L2-norm)

(h) Over estimation (L2-norm)

# Experiments

# Thanks

Wei Lin[1], Jia Wan[2], and Antoni B. Chan[1]

[1]Department of Computer Science, City University of Hong Kong
[2]School of Computer Science and Technology, Harbin Institute of Technology, Shenzhen
elonlin24@gmail.com, jiawan1998@gmail.com, abchan@cityu.edu.hk