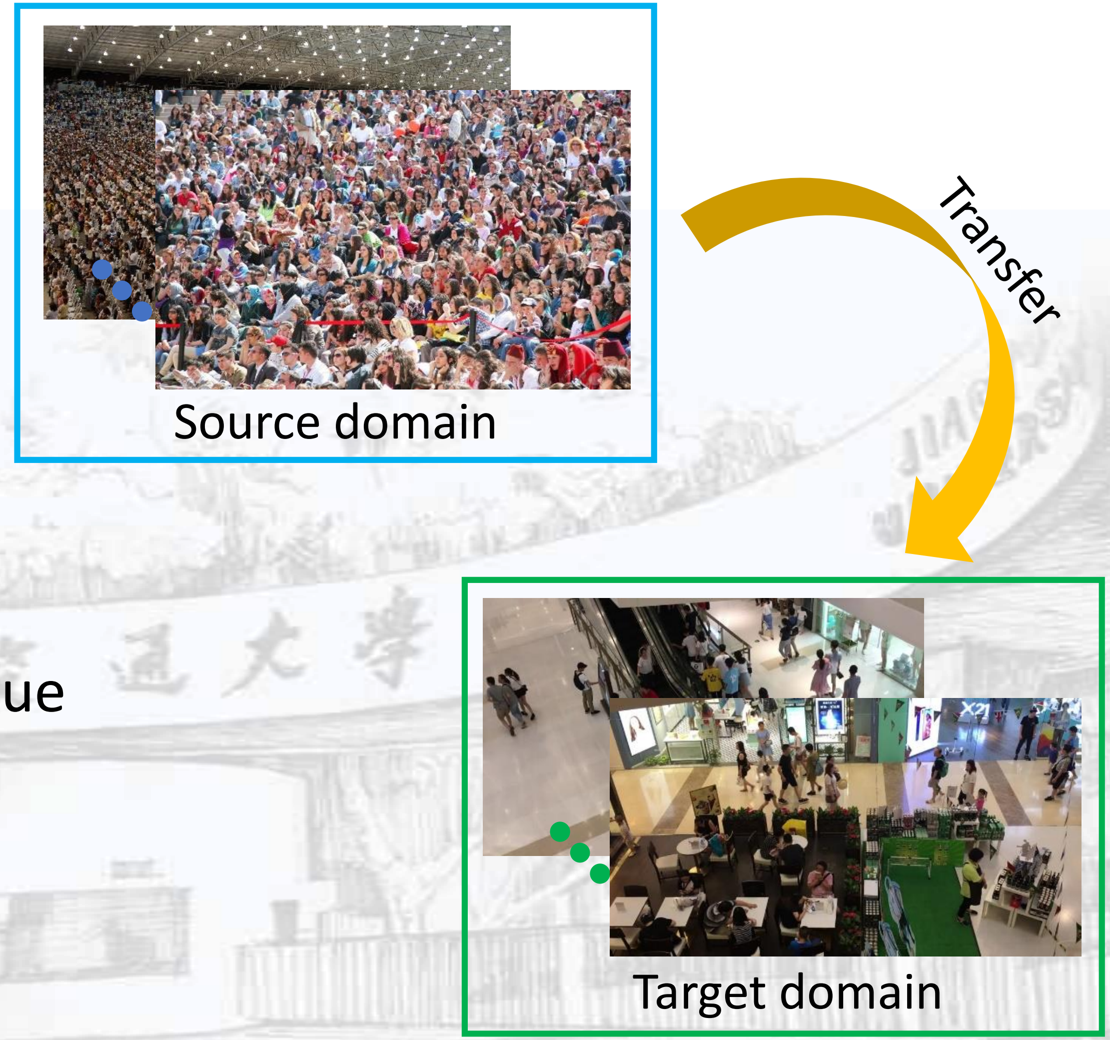
西安交通大學
XI'AN JIAOTONG UNIVERSITY

Error-Aware Density Isomorphism Reconstruction for Unsupervised Cross-Domain Crowd Counting

Yuhang He, Zhiheng Ma, Xing Wei, Xiaopeng Hong, Wei Ke, Yihong Gong
Xi'an Jiaotong University (XJTU), P.R.China

1. Unsupervised Cross-Domain Crowd Counting

- Goal
  - Transfer source domain pre-trained counting model to target domains using unlabeled data.



- Importance
  - Pre-trained model is inferior in practice due to different data distributions.
  - Annotation is expensive and laborious.

1. Unsupervised Cross-Domain Crowd Counting

➢ Challenge

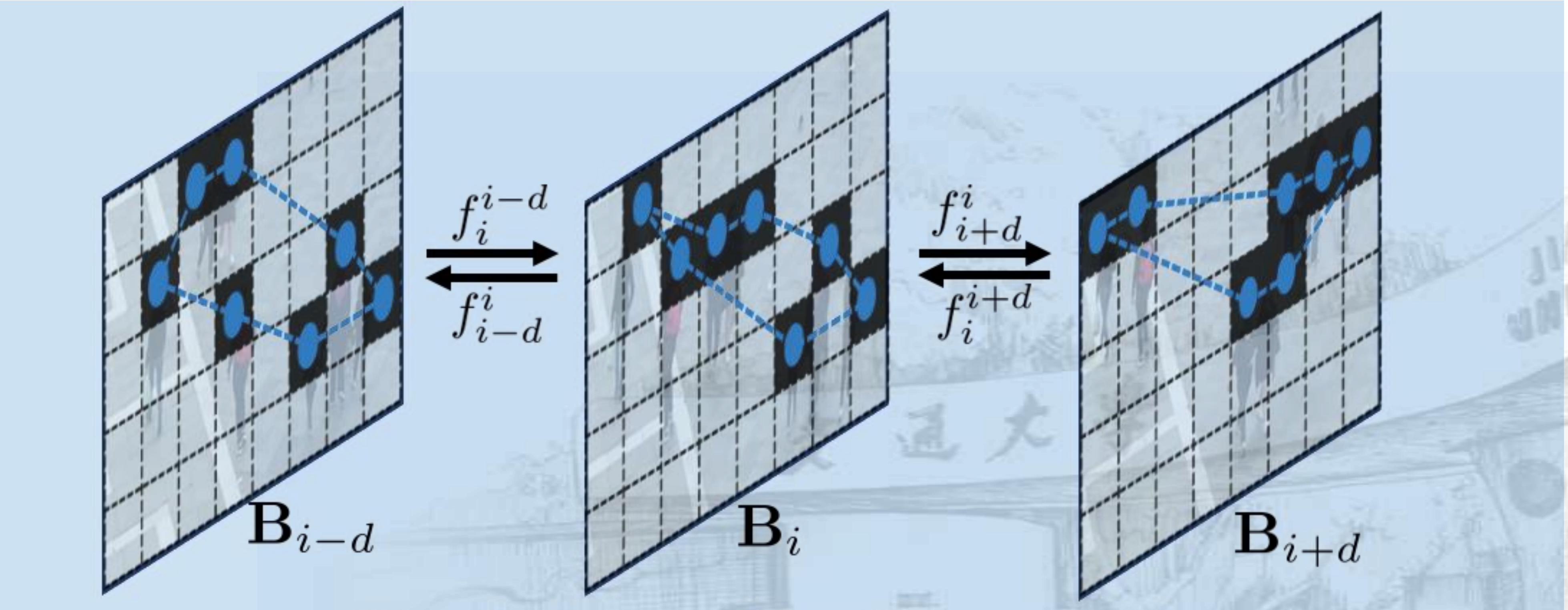
- How to explore concealed information in unlabeled data for knowledge transfer from source domain to unseen target domains?
- How to measure the reliability of supervision signal?

➢ Solutions

- Regard density isomorphism reconstruction as self-supervised signal.
- Model reconstruction erroneousness using estimation-reconstruction consistency.

2. Error-Aware Density Isomorphism Reconstruction for Unsupervised Cross-Domain Crowd Counting

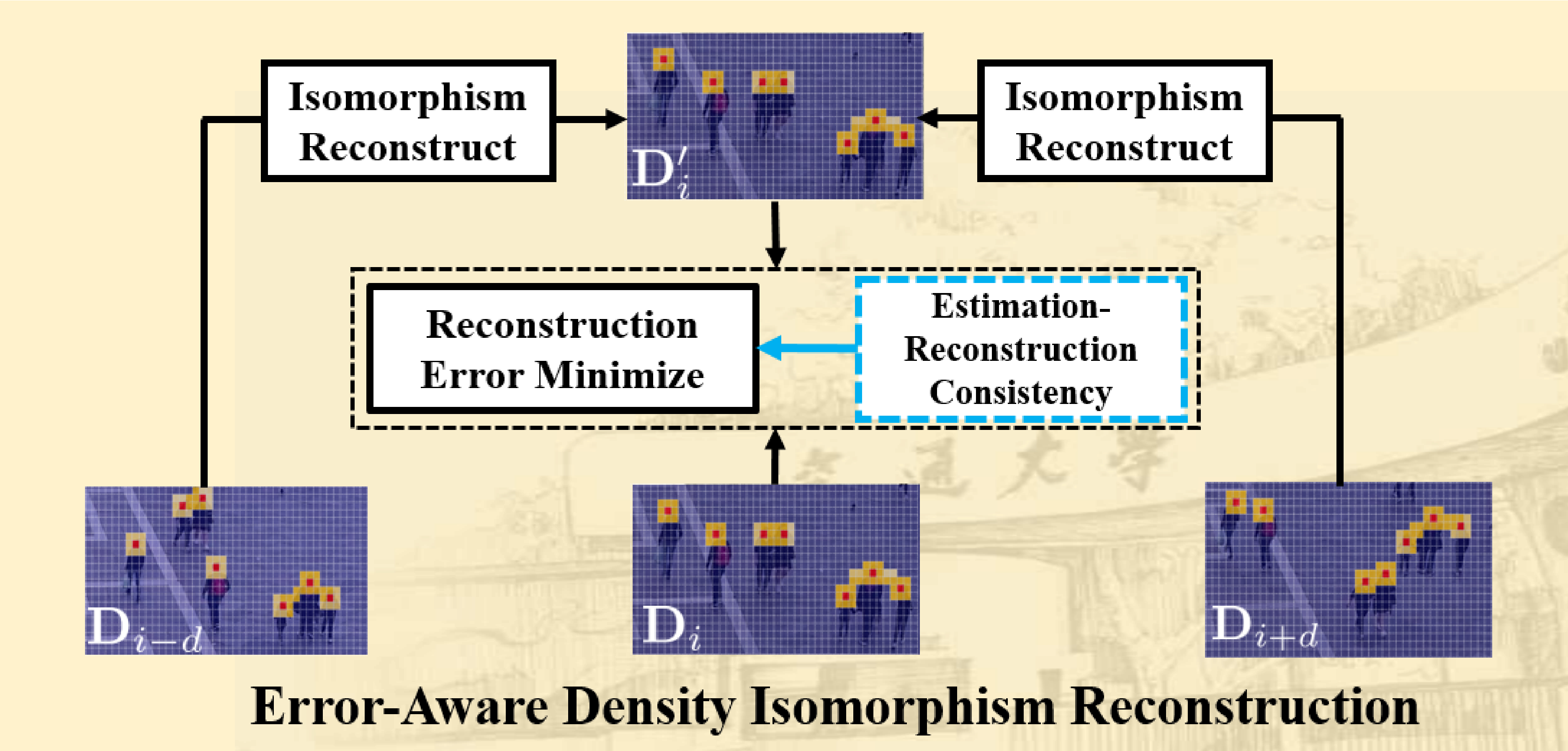
- 2.1 Algorithm I: Density Isomorphism Reconstruction



Density Isomorphism Reconstruction

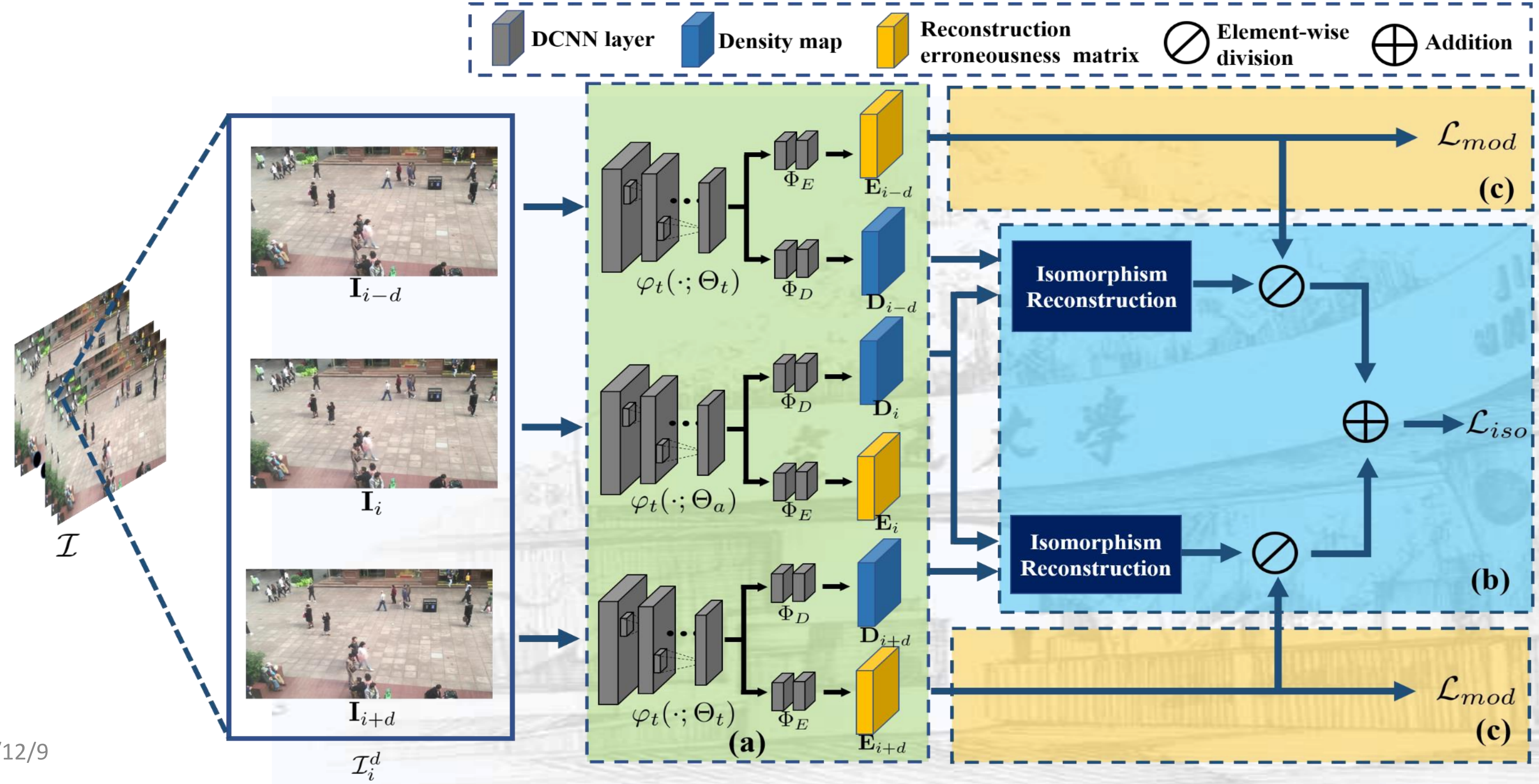
2. Error-Aware Density Isomorphism Reconstruction for Unsupervised Cross-Domain Crowd Counting

➢ 2.1 Algorithm II: Reconstruction Erroneousness Modeling



2. Error-Aware Density Isomorphism Reconstruction for Unsupervised Cross-Domain Crowd Counting

➢ 2.2 The Proposed Framework



2. Error-Aware Density Isomorphism Reconstruction for Unsupervised Cross-Domain Crowd Counting

➢ 2.2 The Proposed Framework

- **Input:**

Image tuple $\mathcal{I}_i^d = \{\mathbf{I}_{i-d}, \mathbf{I}_i, \mathbf{I}_{i+d}\}$ with time interval d, where $\mathbf{I}_i$ is the image frame at time i.

- **Density and Erroneousness Inference Module:**

Estimate a density map $\mathbf{D}_i \in \mathbb{R}_{\geq 0}^{W_D \times H_D}$ and erroneousness matrix $\mathbf{E}_i \in \mathbb{R}_{\geq 0}^{W_D \times H_D}$ for each image $\mathbf{I}_i$.

- **Isomorphism Reconstruction Module:**

Generate a reconstructed density map $\mathbf{D}_i^{i-d'}$ (or $\mathbf{D}_i^{i+d'}$) for the i-th image using $\mathbf{D}_{i-d}$ (or $\mathbf{D}_{i+d}$) .

2. Error-Aware Density Isomorphism Reconstruction for Unsupervised Cross-Domain Crowd Counting

➢ 2.2 The Proposed Framework

- **Reconstruction Erroneousness Modeling Module:**

Simultaneously minimize an density isomorphism reconstruction error and maximize an estimation-reconstruction consistency.

2. Error-Aware Density Isomorphism Reconstruction for Unsupervised Cross-Domain Crowd Counting

➢ 2.3 Key Technique I: Density Isomorphism Reconstruction

(1). Calculate image mapping matrix

$$\mathbf{M}_i^{j^*} = \underset{\mathbf{M}_i^j}{\operatorname{argmin}} \,\|\mathbf{I}_i - \rho(\mathbf{I}_j, \mathbf{M}_i^j)\|^2, \qquad (3)$$

(2). Convert image mapping matrix to density mapping mapping matrix

$$\mathbf{G}_i^{i-d}(u,v) = \mathbf{M}_i^{i-d}\left(\frac{W_I}{W_D}u, \frac{H_I}{H_D}v\right) \cdot \sqrt{\frac{W_D^2+H_D^2}{W_I^2+H_I^2}}, \qquad (4)$$

$$\mathbf{G}_i^{i+d}(u,v) = \mathbf{M}_i^{i+d}\left(\frac{W_I}{W_D}u, \frac{H_I}{H_D}v\right) \cdot \sqrt{\frac{W_D^2+H_D^2}{W_I^2+H_I^2}}. \qquad (5)$$

2. Error-Aware Density Isomorphism Reconstruction for Unsupervised Cross-Domain Crowd Counting

➢ 2.3 Key Technique I: Density Isomorphism Reconstruction

(3). Reconstruct density map according to mapping matrices

$$\mathbf{D}_i^{i-d'}(x,y) = \mathbf{D}_{i-d}(u,v), \forall(x,y) = \mathbf{G}_i^{i-d}(u,v), \quad (6)$$

$$\mathbf{D}_i^{i+d'}(x,y) = \mathbf{D}_{i+d}(u,v), \forall(x,y) = \mathbf{G}_i^{i+d}(u,v), \quad (7)$$

2. Error-Aware Density Isomorphism Reconstruction for Unsupervised Cross-Domain Crowd Counting

➢ 2.3 Key Technique II: Reconstruction Erroneousness Modeling

$$\mathcal{L}(\mathcal{I}_i^d) = \mathcal{L}_{iso}(\mathcal{I}_i^d) + \mathcal{L}_{mod}(\mathcal{I}_i^d),$$

（1） Error-aware density isomorphism reconstruction objective：

$$\begin{aligned}\mathcal{L}_{iso}(\mathcal{I}_i^d) = \left\| \left| \mathbf{D}_i - \mathbf{D}_i^{i-d'} \right|_e \oslash \mathbf{E}_{i-d} \right\|^2 + \\ \left\| \left| \mathbf{D}_i - \mathbf{D}_i^{i+d'} \right|_e \oslash \mathbf{E}_{i+d} \right\|^2,\end{aligned} \tag{8}$$

$$\mathbf{D}_i \in \mathbb{R}_{\geq 0}^{W_D \times H_D}$$

$$\mathbf{E}_i \in \mathbb{R}_{>0}^{W_D \times H_D}$$

（2） Erroneousness matrix regularization term

$$\mathcal{L}_{mod}(\mathcal{I}_i^d) = \log(\mathbf{E}_{i-d}) + \log(\mathbf{E}_{i+d}),$$

2. Error-Aware Density Isomorphism Reconstruction for Unsupervised Cross-Domain Crowd Counting

➢ 2.4 Experiment:

Table 1: Performance Evaluation on Four Benchmark Datasets.

Supervision	Method	Venice		UCSD		MALL		FDST	
		MAE↓	MSE↓	MAE↓	MSE↓	MAE↓	MSE↓	MAE↓	MSE↓
—	Baseline	33.95	39.44	7.96	8.54	4.27	5.94	4.77	8.33
Supervised	PFlow	15.00	19.60	*0.81*	*1.07*	—	—	2.84	3.57
	BL	*9.99*	*14.24*	0.84	1.08	*1.54*	*2.00*	*1.42*	*1.88*
Semi-supervised	SSR	19.84	31.13	1.68	2.07	2.69	3.38	5.41	6.13
	FSSA	<u>17.83</u>	<u>25.24</u>	<u>1.45</u>	<u>1.85</u>	<u>2.32</u>	<u>2.97</u>	<u>2.96</u>	<u>3.86</u>
Unsupervised	CSCC	18.05	22.34	8.89	9.87	4.01	4.99	5.15	7.84
	CODA	31.39	37.17	5.25	6.07	3.37	4.43	4.74	8.27
	SCP	22.79	26.52	4.55	5.71	3.03	4.04	4.28	6.74
	Ours-*w/o mod*	14.66	17.48	2.22	2.71	3.17	4.03	3.97	4.76
	Ours	**11.23**	**15.16**	**1.79**	**2.47**	**2.36**	**3.12**	**3.25**	**3.94**

2. Error-Aware Density Isomorphism Reconstruction for Unsupervised Cross-Domain Crowd Counting

- 2.4 Experiment:

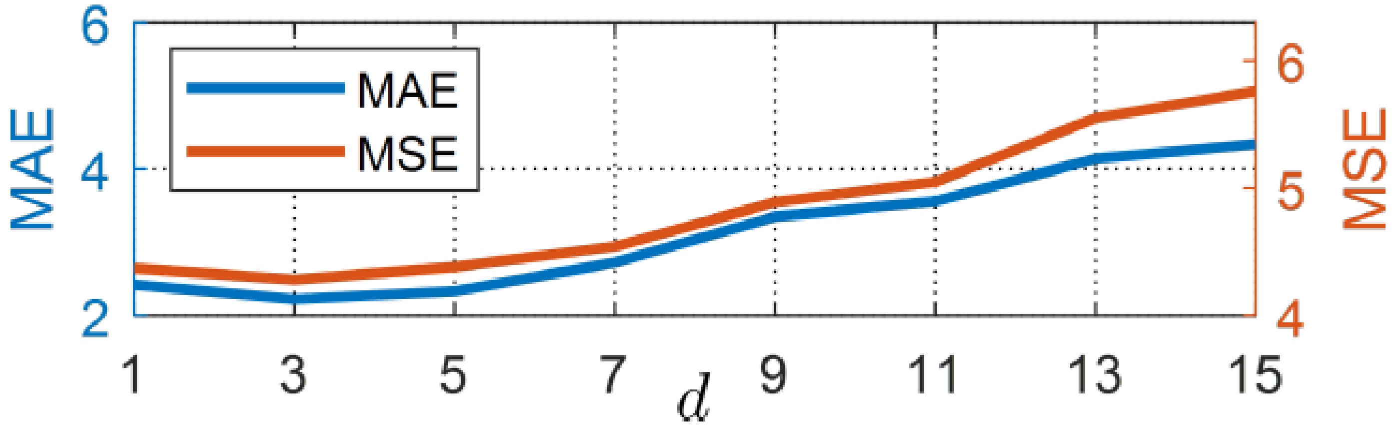


Figure 3: Influence of different d values.

2. Error-Aware Density Isomorphism Reconstruction for Unsupervised Cross-Domain Crowd Counting

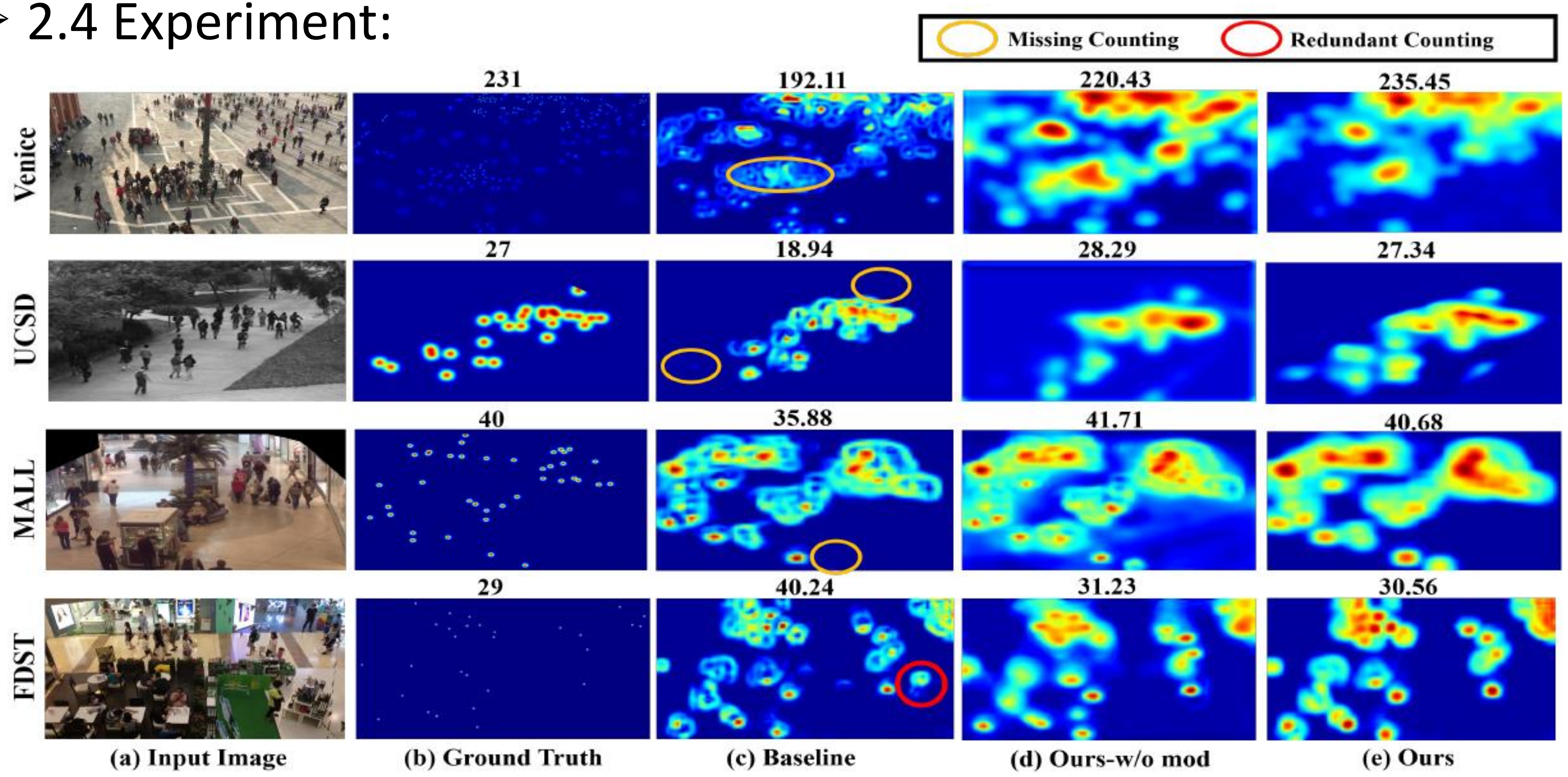
➢ 2.4 Experiment:

Table 2: Robustness to Different Pre-trained Models.

Source	Method	Venice		UCSD		MALL		FDST	
		MAE↓	MSE↓	MAE↓	MSE↓	MAE↓	MSE↓	MAE↓	MSE↓
ShanghaiTech-A	Baseline-MESA	51.57	53.68	16.80	17.81	15.67	16.75	12.80	25.59
	Ours	17.83	22.19	5.13	5.83	5.57	6.54	6.27	7.64
	Baseline-BL	40.13	51.54	15.36	16.18	12.48	12.99	5.01	8.09
	Ours	14.10	19.13	4.22	5.01	4.77	5.93	3.96	5.12
UCF-QNRF	Baseline-MESA	43.16	57.88	9.04	9.77	5.71	6.67	6.12	7.57
	Ours	13.05	15.72	2.64	3.60	4.65	6.01	4.95	6.10
	Baseline-BL	33.95	39.44	7.96	8.54	4.27	5.94	4.77	8.33
	Ours	**11.23**	**15.16**	**1.79**	**2.47**	**2.36**	**3.12**	**3.25**	**3.94**

2. Error-Aware Density Isomorphism Reconstruction for Unsupervised Cross-Domain Crowd Counting

➢ 2.4 Experiment:



Thank you