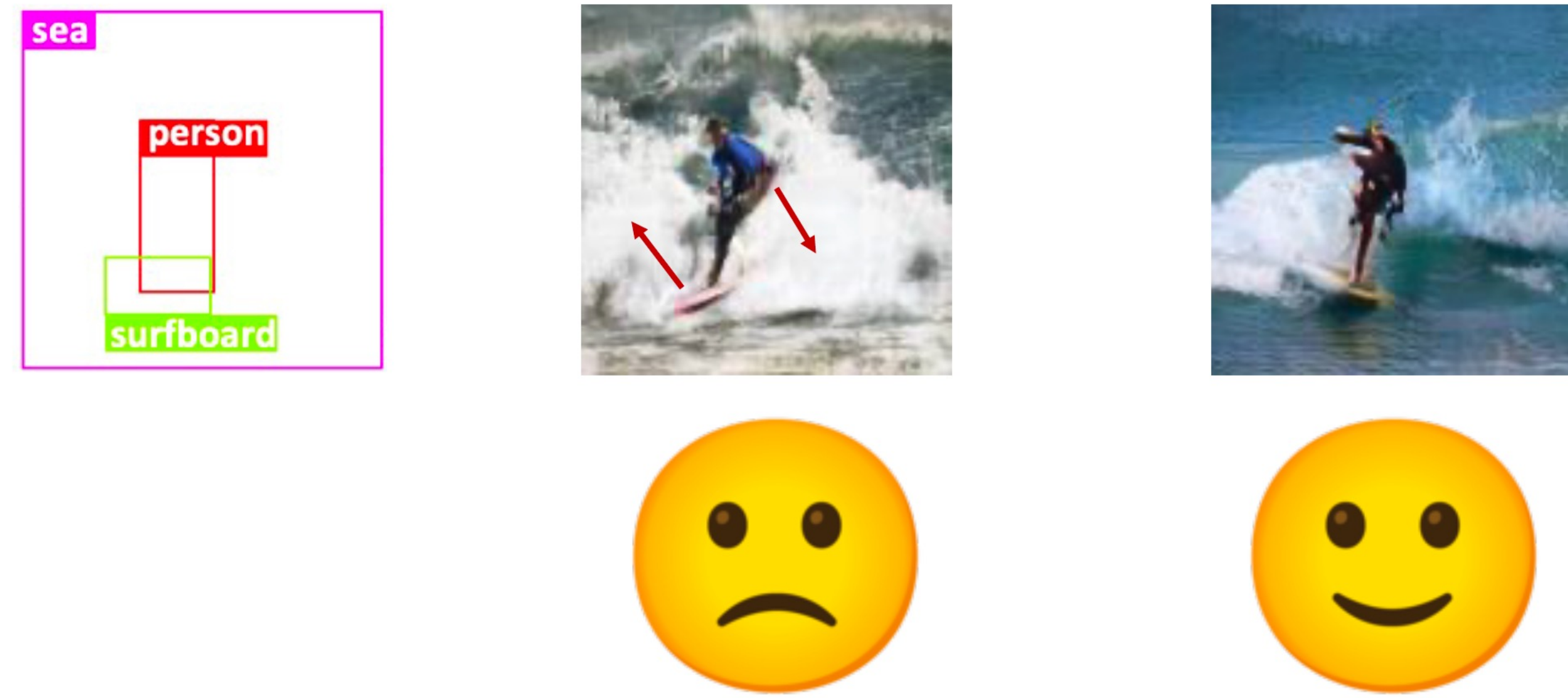
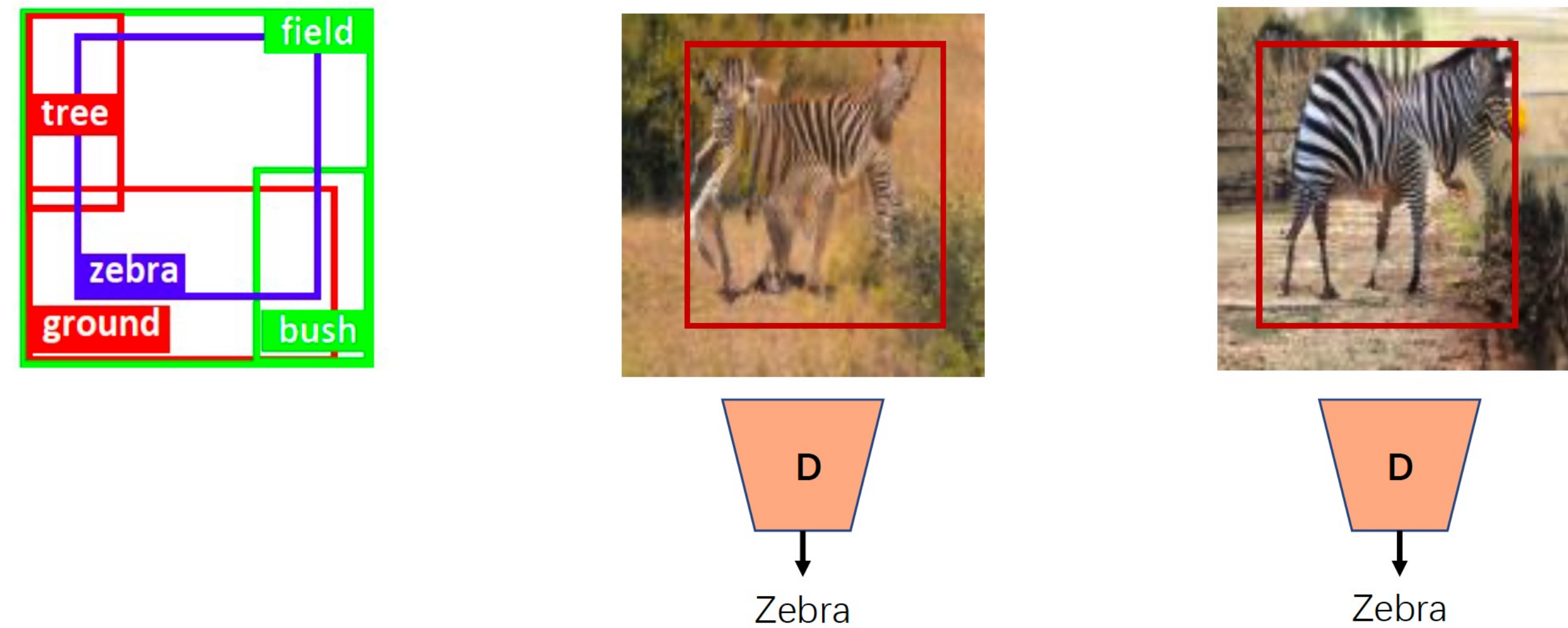


Challenges in Layout to Image Generation

1. Context

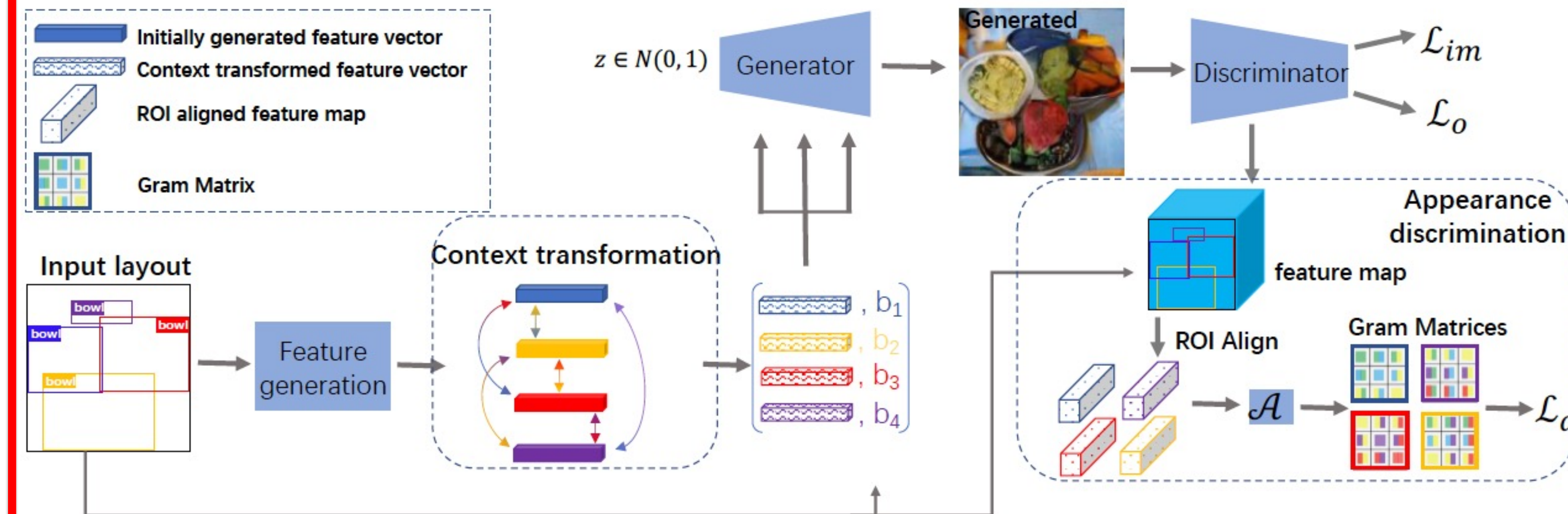


2. Location sensitive appearance penalty



This work aims to address the two challenges which haven't gain much attention in previous works.

Our Method



1. Context transformation

$$\mathbf{p}_i^c = \sum_{j=1}^N w_{i,j} \mathbf{p}_j \mathbf{W}_v,$$

$$w_{i,j} = \frac{\exp(\alpha_{i,j})}{\sum_{k=1}^N \exp(\alpha_{i,k})},$$

$$\alpha_{i,j} = (\mathbf{p}_i \mathbf{W}_q)(\mathbf{p}_j \mathbf{W}_k)^T,$$

2. Appearance discrimination

$$\mathcal{A}_i = \mathbf{s}_i \mathbf{s}_i^T / d_s$$

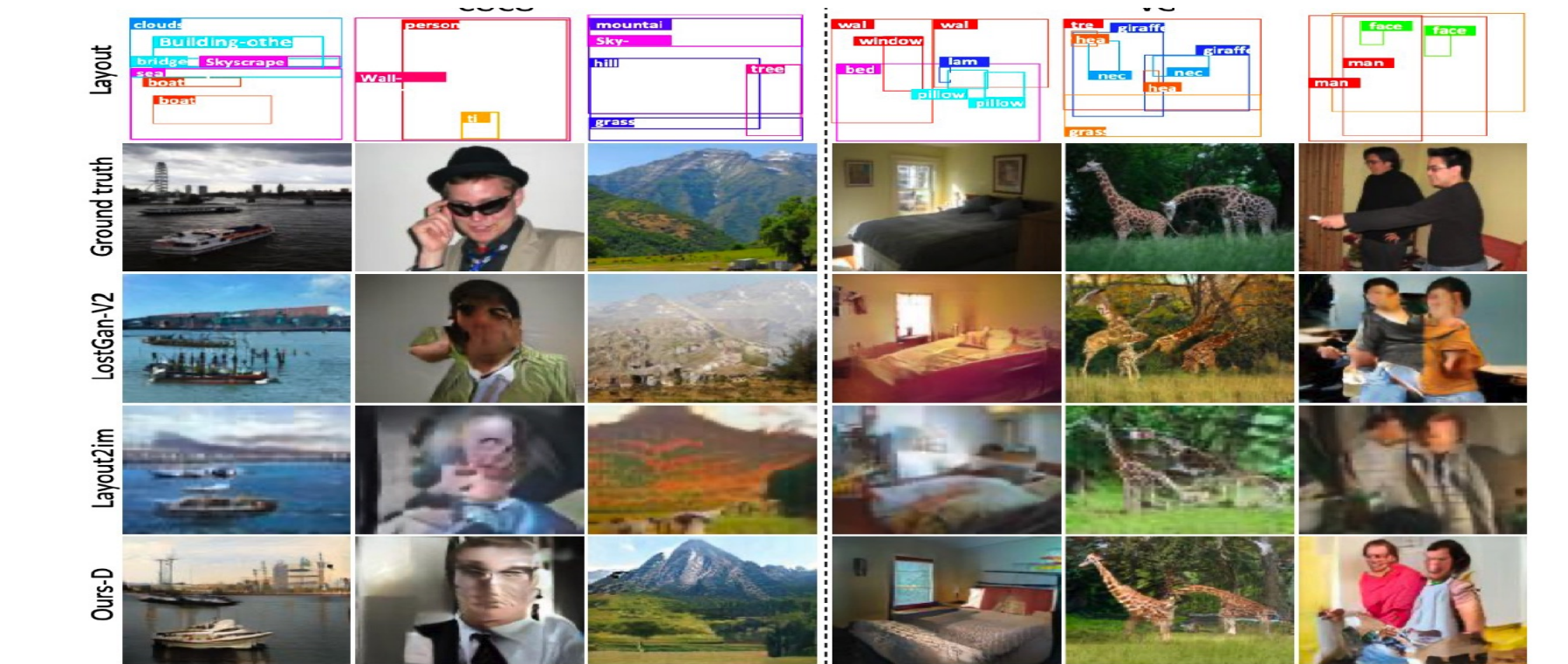
$$\mathcal{L}_a(\mathcal{G}, \mathcal{D}) = \mathbb{E}_{\mathcal{A}^r \sim p_{data}^r(\mathcal{A}^r)} [\log(\mathcal{D}(\mathcal{A}^r|y))] + \mathbb{E}_{\mathcal{A}^g \sim p_{data}^g(\mathcal{A}^g)} [1 - \log(\mathcal{D}(\mathcal{A}^g|y))]$$

Results

Quantitative result

Methods	Resolution	Generator	Inception Score \uparrow		FID \downarrow		Diversity Score \uparrow	
			COCO	VG	COCO	VG	COCO	VG
Real images	64 \times 64	-	16.3 \pm 0.4	13.9 \pm 0.5	-	-	-	-
Real images	128 \times 128	-	22.3 \pm 0.5	20.5 \pm 1.5	-	-	-	-
pix2pix [17]	64 \times 64	E-D	3.5 \pm 0.1	2.7 \pm 0.02	121.97	142.86	0	0
Layout2im [48]	64 \times 64	E-D	9.1 \pm 0.1	8.1 \pm 0.1	38.14	40.07	0.15 \pm 0.06	0.17 \pm 0.09
Ours-ED	64 \times 64	E-D	10.27 \pm 0.25	8.53 \pm 0.13	31.32	33.91	0.39 \pm 0.09	0.4 \pm 0.09
Grid2im [3]	128 \times 128	E-D	11.22 \pm 0.15	-	63.44	-	0.28 \pm 0.11	-
LostGAN-v1 [38]	128 \times 128	D	13.8 \pm 0.4	11.1 \pm 0.6	29.65	29.36	0.40 \pm 0.09	0.43 \pm 0.09
LostGAN-v2 [49]	128 \times 128	D	14.21 \pm 0.4	10.71 \pm 0.76	24.76	29.00	0.55 \pm 0.09	0.53 \pm 0.09
OC-GAN [40]	128 \times 128	D	14.0 \pm 0.2	11.9 \pm 0.5	36.04	28.91	-	-
AG-Layout2im [27]	128 \times 128	E-D	-	8.5 \pm 0.1	-	39.12	-	0.15 \pm 0.09
Ours-D	128 \times 128	D	15.62 \pm 0.05	12.69 \pm 0.45	22.32	21.78	0.55 \pm 0.09	0.54 \pm 0.09

Qualitative result



Ablation study

Table 2. Ablation study on COCO-Thing-Stuff dataset.

baseline [38]	context	appearance	Inception Score	FID
\checkmark			13.8 \pm 0.4	29.65
\checkmark	\checkmark		14.97 \pm 0.27	24.05
\checkmark		\checkmark	15.28 \pm 0.24	21.73
\checkmark	\checkmark	\checkmark	15.62 \pm 0.05	22.32

Visualization

