

BargainNet: Background-Guided Domain Translation for Image Harmonization

Wenyan Cong, Li Niu, Jianfu Zhang, Jing Liang, Liqing Zhang

MoE Key Lab of Artificial Intelligence, Shanghai Jiao Tong University















F











Background - Image Composition









Image Composition



Background - Image Harmonization



- Foreground and background are captured under different conditions (e.g., weather, season, time of the day).
- > Matches the visual appearance of composite foreground and background.



Background

F











Flaws of Our DoveNet (CVPR'20)

Instable training caused by GAN structure and adversarial learning.
Domain representation is not so meaningful as expected.









Exemplar-guided Domain Translation



Image Harmonization















Background

F











BargainNet – Framework

Combine domain code extractor *E* with an image harmonization generator *G*.
Extract the background domain code to guide the foreground translation.





BargainNet – Domain Code Extractor



- > Extract the domain codes of the foreground/background regions.
- \succ Given an image triplet, we can obtain their domain codes: z_b , \tilde{z}_f , z_f , and \hat{z}_f





BargainNet – Domain Code Extractor



- > Extract the domain codes of the foreground/background regions.
- \succ Given an image triplet, we can obtain their domain codes: z_b , \tilde{z}_f , z_f , and \hat{z}_f





BargainNet – Triplet Losses

Employ two triplet losses $\mathcal{L}_{\hat{f}b}$ and $\mathcal{L}_{f\hat{f}}$ to enforce domain code contains domain information instead of other domain-irrelevant information (e.g., semantic layout).



Background

F











Experiment – Quantitative Comparison on iHarmony4

Sub-dataset	Sub-dataset HCOCO		HAd	HAdobe5k		HFlickr		Hday2night		All	
Evaluation metric	MSE↓	PSNR	↑ MSE↓	PSNR ↑	MSE↓	PSNR ↑	MSE↓	PSNR↑	MSE↓	PSNR ↑	
Input composite	69.37	33.94	345.54	28.16	264.35	28.32	109.65	34.01	172.47	31.63	
Lalonde and Efros[4]	110.10	31.14	158.90	29.66	329.87	26.43	199.93	29.80	150.53	30.16	
Xue et al.[5]	77.04	33.32	2 274.15	28.79	249.54	28.32	190.51	31.24	155.87	31.40	
Zhu et al.[20]	79.82	33.04	414.31	27.26	315.42	27.52	136.71	32.32	204.77	30.72	
DIH [1]	51.85	34.69	92.65	32.28	163.38	29.55	82.34	34.62	76.77	33.41	
DoveNet [3]	36.72	35.83	3 52.32	34.34	133.14	30.21	54.05	35.18	52.36	34.75	
S ² AM [2]	33.07	36.09	48.22	35.34	124.53	31.00	48.78	35.60	48.00	35.29	
Ours	24.84	37.03	39.94	35.34	97.32	31.34	50.98	35.67	37.82	35.88	
Foreground ratios		$0\%\sim5\%$		$5\% \sim 15\%$		$15\% \sim 100\%$		0%	$0\% \sim 100\%$		
Evaluation m	etric	MSE↓	fMSE↓	MSE↓	fMSE↓	MSE↓	fMSE↓	MSE	fMSI	E↓	
Input compo	site	28.51	1208.86	119.19	1323.23	577.58	1887.05	5 172.47	7 1387.	30	
Lalonde and Ef	ros[6]	41.52	1481.59	120.62	1309.79	444.65	1467.98	3 150.53	3 1433.	21	
Xue et al.[7]	31.24	1325.96	132.12	1459.28	479.53	1555.69	9 155.87	7 1411.	40	
Zhu et al.	8]	33.30	1297.65	145.14	1577.70	682.69	2251.76	5 204.77	7 1580.	17	
DIH [9]		18.92	799.17	64.23	725.86	228.86	768.89	76.77	773.	18	
DoveNet []	1]	14.03	591.88	44.90	504.42	152.07	505.82	52.36	549.9	96	
S ² AM [10		13.51	509.41	41.79	454.21	137.12	449.81	48.00	481.7	79	
Ours		10.55	450.33	32.13	359.49	109.23	353.84	37.82	405.2	23	



Experiment – Domain Code Analyses

Compared to our DoveNet (CVPR'20), the domain code extracted by Domain Code Extractor contains domain information as expected.

		$d_{b,f} < d_{b,\tilde{f}}$	$d_{b,\hat{f}} < d_{b,\tilde{f}}$	$d_{f,\hat{f}} < d_{f,\tilde{f}}$	$d_{\hat{f},f} < d_{\hat{f},\tilde{f}}$	$d_{\hat{f},b} < d_{\hat{f},\tilde{f}}$	$d_{f,b} < d_{f,\tilde{f}}$	All
DoveNet[3]	Train	47.08%	49.24%	72.22%	71.47%	12.01%	11.75%	5.93%
	Test	51.34%	51.58%	62.34%	54.65%	13.68%	15.64%	5.09%
Ours	Train	88.63%	97.87%	93.65%	91.92%	96.38%	87.98%	80.70%
	Test	90.28%	97.39%	91.87%	89.28%	96.26%	89.09%	81.36%



Experiment – Qualitative Comparison on iHarmony4



Input

Ground Truth

DIH

DoveNet

Net

S²AM BargainNet (w/o \mathcal{L}_{tri}) BargainNet



Experiment – Qualitative Comparison on Real Composite images



Method	B-T score↑
Input composite	0.357
DIH [9]	0.813
DoveNet [1]	0.897
S ² AM [10]	1.140
Ours	1.266

Composite

S²AM

BargainNet

Experiment – Background Harmonization

Input

DIH

DoveNet

S²AM

BargainNet

Experiment – Inharmony Level Prediction

Is.

1.5654

ME 2021 IEEE International Conference on Multimedia and Expo July 5-9, 2021 || Shenzhen, China || Virtual

Thanks for watching!

BargainNet: Background-Guided Domain Translation for Image Harmonization

Wenyan Cong, Li Niu, Jianfu Zhang, Jing Liang, Liqing Zhang

MoE Key Lab of Artificial Intelligence, Shanghai Jiao Tong University

More details