Animation Image Harmonization

2022.05.27
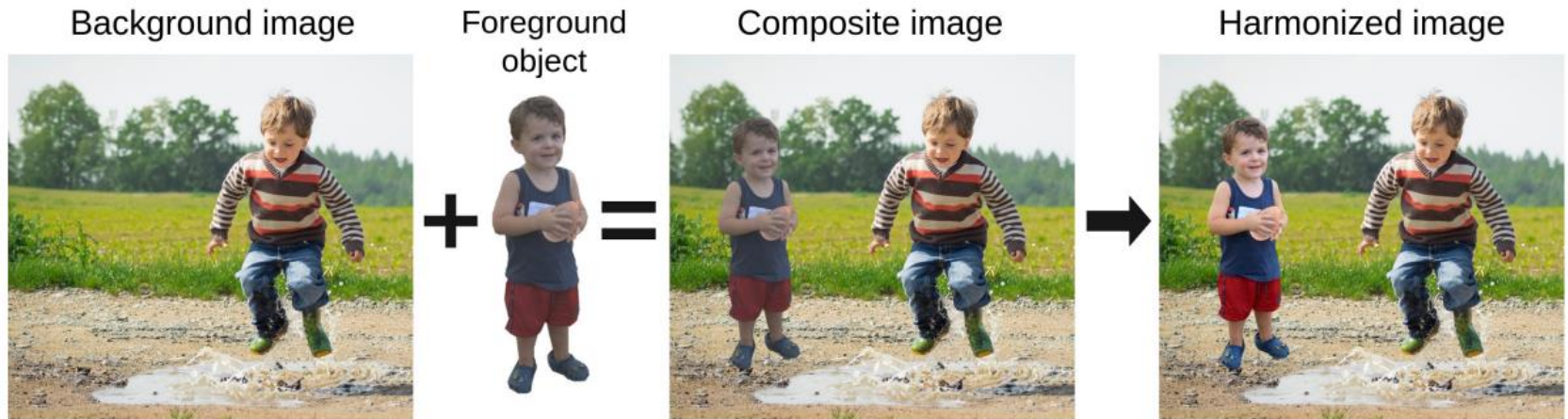
Team11 – 涂皓鈞 蔡惠芸

Outline

- Introduction for Image Harmonization

- Paper Survey
  - iS^{2}AM (2021 WACV)
  - iDIH-HRNet (2021 WACV)
  - CDTNet (2022 CVPR)
  - Experimental Results

Image Harmonization

- Photo editing: image composition
  - Composite image = foreground object + background image
  - Achieve visual consistency

Background image   Foreground object   Composite image   Harmonized image

Animation Image Harmonization

- 3D model live
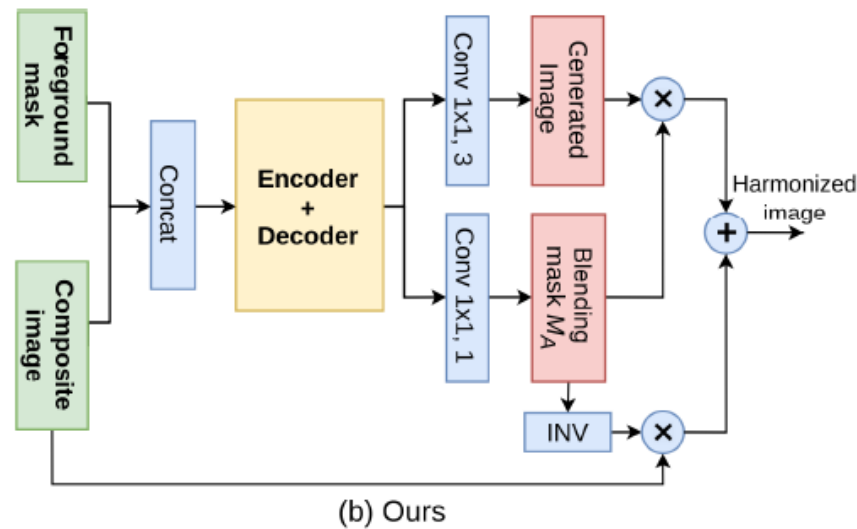
Outline

- Introduction for Image Harmonization

- Paper Survey
    - iS^{2}AM (2021 WACV)
    - iDIH-HRNet (2021 WACV)
    - CDTNet (2022 CVPR)
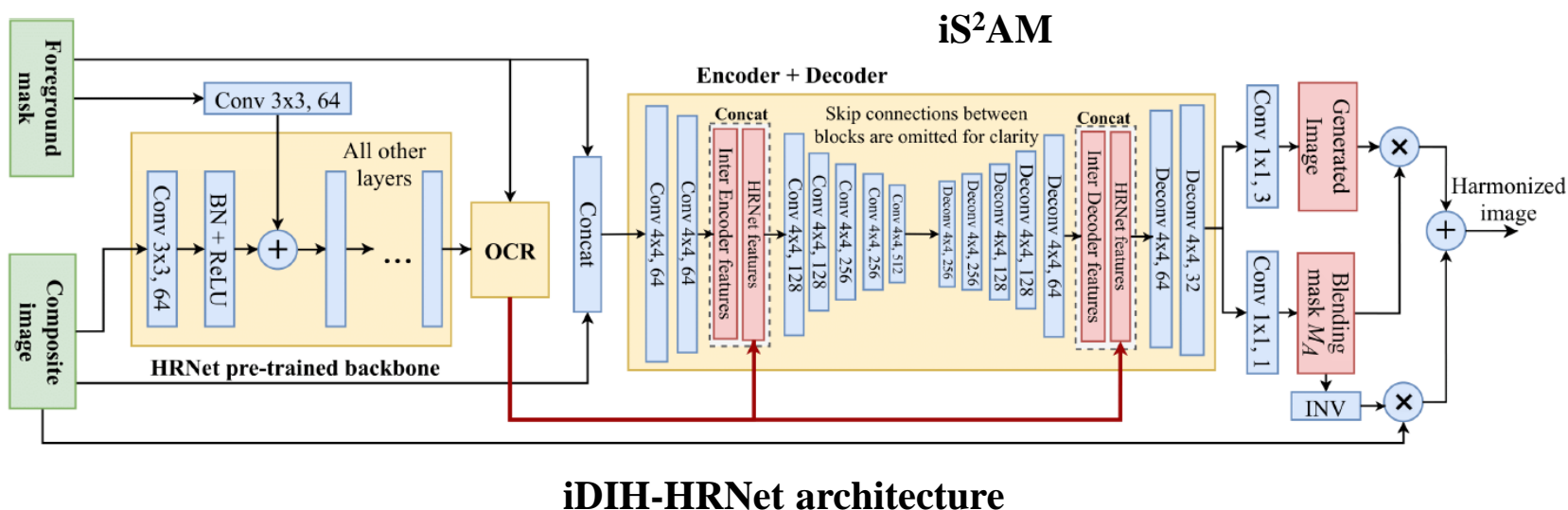   - Experimental Results

iS²AM

- Auto-encoder architecture
  - Encoder: extract **local texture** of image
  - Decoder: reconstruct harmonized output results
  - Predict blending mask M_A

(b) Ours

Source: K. Sofiiuk, "Foreground-aware Semantic Representations for Image Harmonization", 2021 IEEE WACV.

iDIH-HRNet

- Pre-trained model on ImageNet
  - Extract more **high-level semantic features**
  - Incorporate with auto-encoder architecture
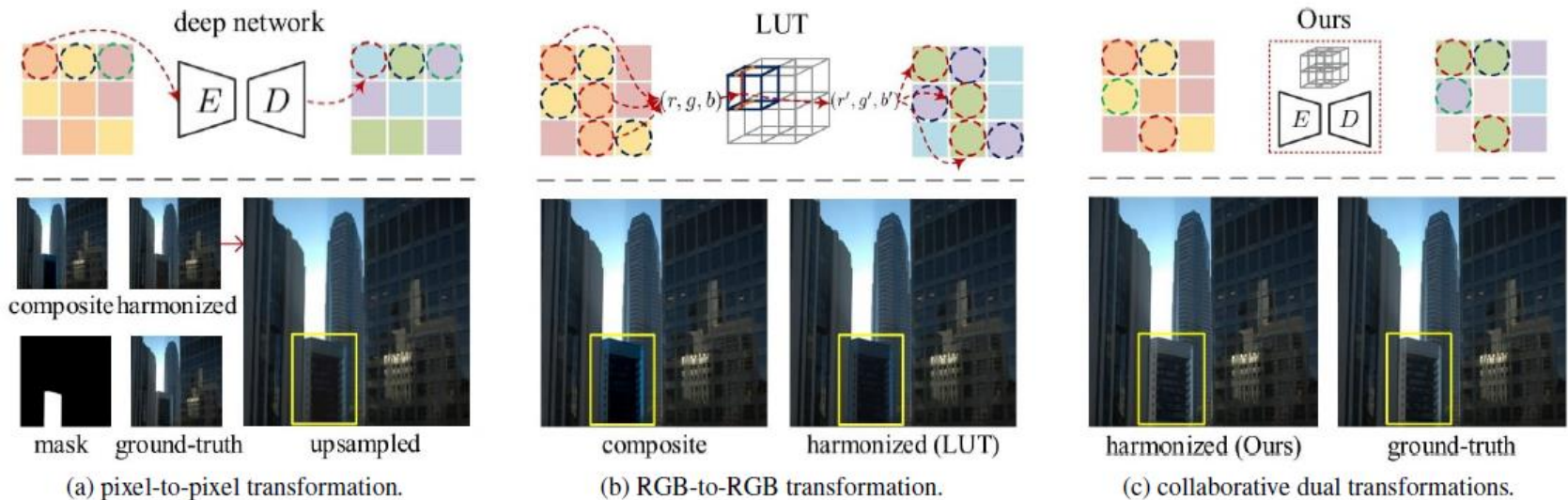

iDIH-HRNet architecture

- Loss function: FN-MSE $\mathcal{L}_{rec}(\hat{I}) = \dfrac{\sum\limits_{h,w} \left\| I_{h,w}^{pred} - I_{h,w} \right\|_2^2}{\max\left\{ A_{min}, \sum\limits_{h,w} M_{h,w} \right\}},$ **Size of foreground mask**
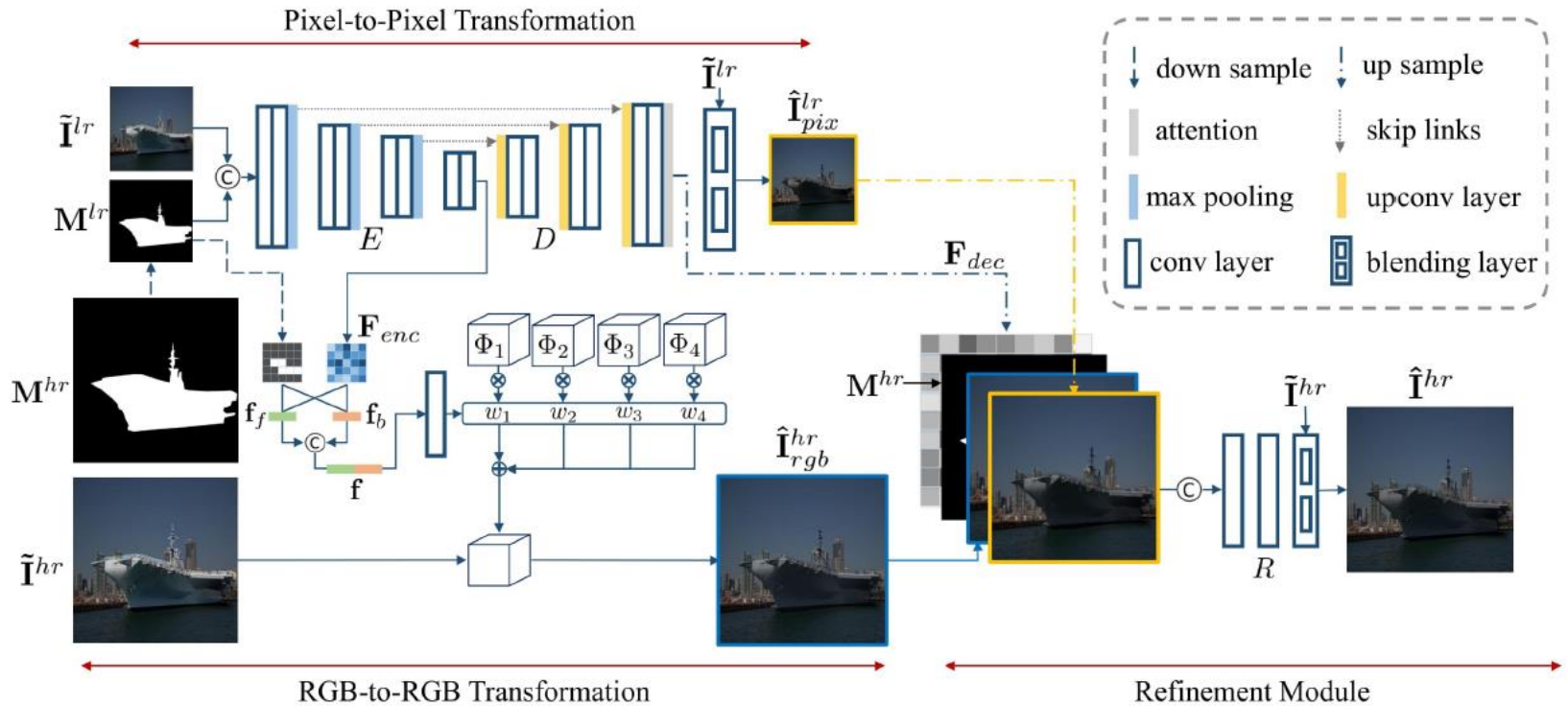
CDTNet

- Pixel-to-pixel transformation
  - Local detail consideration, low resolution
- Color transformation
  - High resolution, inharmonious local regions
- Achieve HR image harmonization

(a) pixel-to-pixel transformation.

(b) RGB-to-RGB transformation.

(c) collaborative dual transformations.

CDTNet

Experimental Results

Image Size	Method	MSE↓	PSNR↑	fMSE↓	SSIM↑
	Composite images	352.05	28.10	2122.37	0.9642
	pix2pixHD [35]	63.45	31.64	332.43	0.9135
	CRN [4]	90.11	29.77	259.28	0.8225
	HiDT [2]	265.32	29.95	1501.93	0.9628
1024	DoveNet [8]	51.00	34.81	312.88	0.9729
×	S^2AM [9]	47.01	35.68	262.39	0.9784
1024	Guo et al. [15]	56.34	34.69	417.33	0.9471
	RainNet [24]	42.56	36.61	305.17	0.9844
	iS^2AM [31]	25.03	38.29	168.85	0.9846
	CDTNet-256 (sim)	31.15	37.65	195.93	0.9841
	CDTNet-256	**21.24**	**38.77**	**152.13**	**0.9868**

Image Size	Method	MSE↓	PSNR↑	fMSE↓	SSIM↑
	Composite images	353.92	28.07	2139.97	0.9631
	iS^2AM [31]	46.37	36.57	271.59	0.9838
2048	CDTNet-256 (sim)	41.11	37.28	234.06	0.9819
×	CDTNet-256	29.02	37.66	198.85	0.9845
2048	CDTNet-512 (sim)	38.31	37.05	233.44	0.9828
	CDTNet-512	**23.35**	**38.45**	**159.13**	**0.9853**