Image inpainting

EdgeConnect : Generative Image Inpainting with Adversarial Edge Learning

컴퓨터소프트웨어학과 석사 2기 김 경 아

CONTENTS

01. Concept of Image Inpainting

01 What is image inpainting?

[Image Restoration]

01 What is image inpainting?

[Image Restoration]

01 What is image inpainting?

[Object Removal]

Original Original + Masking inpainting

01 Traditional Image inpainting method

1. Diffusion-based image synthesis

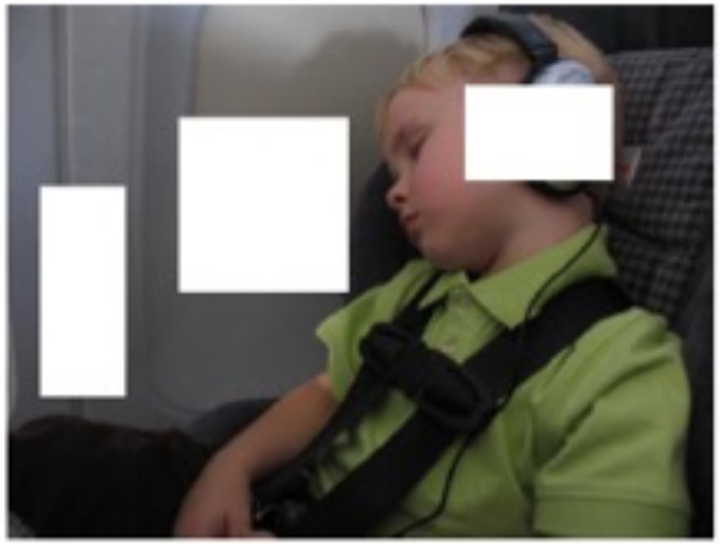
비어있는 부분의 바로 근처에 있는 픽셀들을 그대로 가져오는 방식.

01 Traditional Image inpainting method

2. Patch-based image synthesis

소스 이미지에서 유사한 부분을 찾아 메꾸는 방식.

Input



PatchMatch

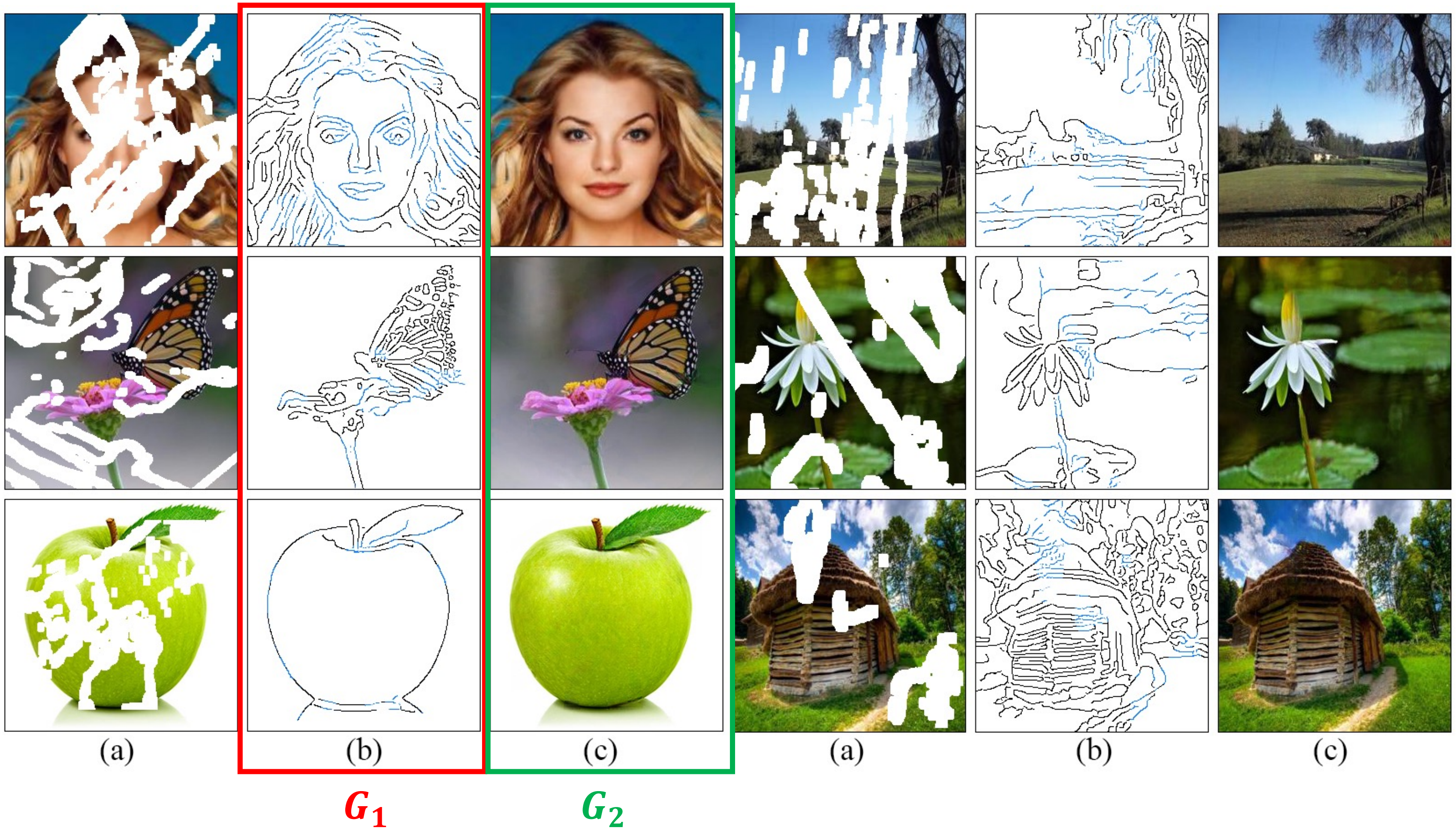
limitation : 이러한 기존의 방법들은 누락된 영역에 위치할 수 있는 복잡한(중요한) 세부 정보를 재구성하는 데는 한계가 있음.

02. Edge Connect Architecture

02 Proposed Approach : EdgeConnect

Two Stage Process

Edge-generation (G_1) + Image Completion (G_2)



02 Proposed Approach : EdgeConnect

Overview

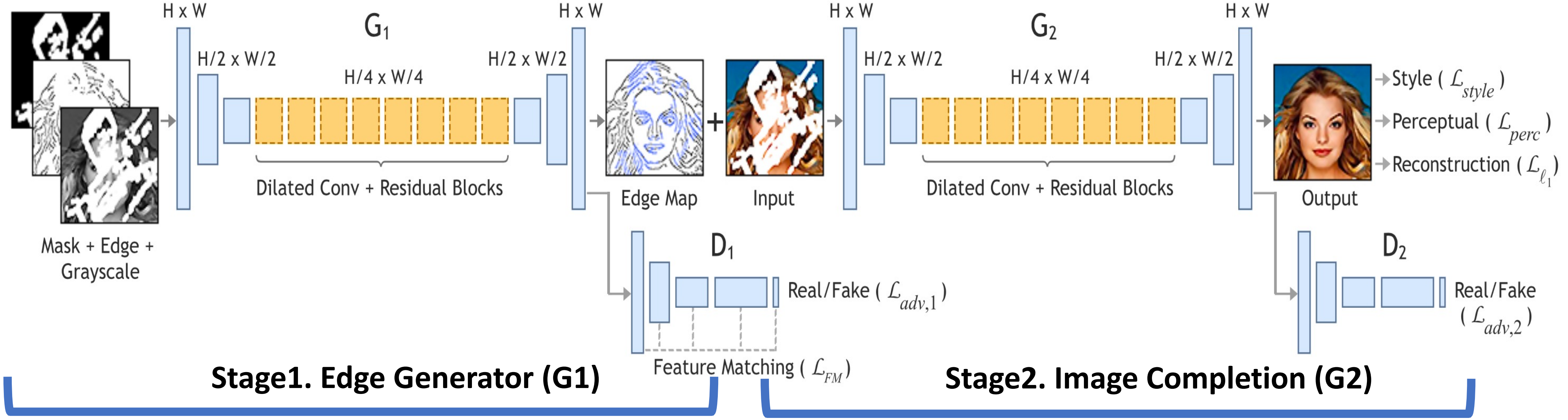
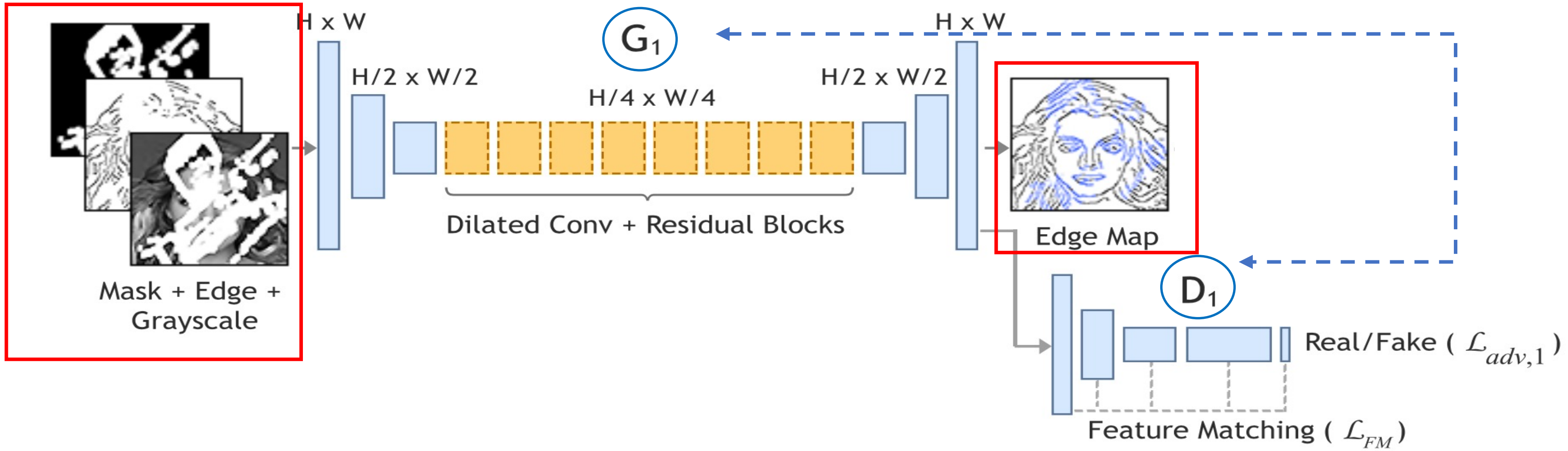


Figure 2: Summary of our proposed method. Incomplete grayscale image and edge map, and mask are the inputs of G_1 to predict the full edge map. Predicted edge map and incomplete color image are passed to G_2 to perform the inpainting task.

02 Proposed Approach : EdgeConnect

Stage 1. Edge-Generator



[Input]

- Mask : 0 (background) / 1 (missing region)
- Masked Grayscale

$$\tilde{\mathbf{I}}_{gray} = \mathbf{I}_{gray} \odot (\mathbf{1} - \mathbf{M})$$

- Masked Edge map

$$\tilde{\mathbf{C}}_{gt} = \mathbf{C}_{gt} \odot (\mathbf{1} - \mathbf{M})$$

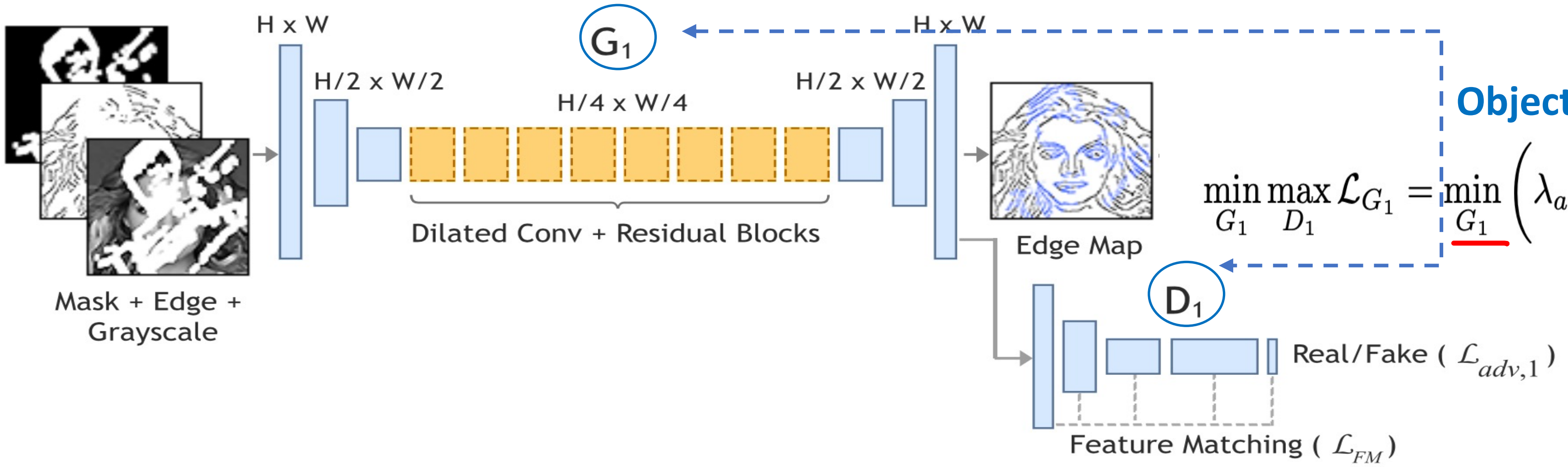
[Output]

- Edge Map

$$\mathbf{C}_{pred} = G_1\left(\tilde{\mathbf{I}}_{gray}, \tilde{\mathbf{C}}_{gt}, \mathbf{M}\right)$$

02 Proposed Approach : EdgeConnect

Training loss

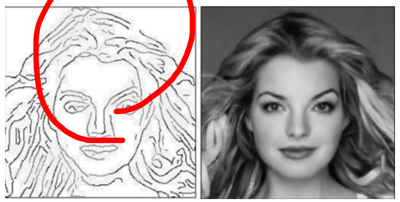


Objective Function

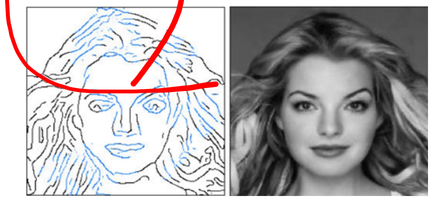
$$\min_{G_1} \max_{D_1} \mathcal{L}_{G_1} = \min_{G_1} \left(\lambda_{adv,1} \max_{D_1} (\mathcal{L}_{adv,1}) + \lambda_{FM} \mathcal{L}_{FM} \right)$$

- Adversarial loss

$$\mathcal{L}_{adv,1} = \mathbb{E}_{(\mathbf{C}_{gt}, \mathbf{I}_{gray})} [\log D_1(\mathbf{C}_{gt}, \mathbf{I}_{gray})] + \mathbb{E}_{\mathbf{I}_{gray}} \log [1 - D_1(\mathbf{C}_{pred}, \mathbf{I}_{gray})]$$



C_{gt} I_{gray}



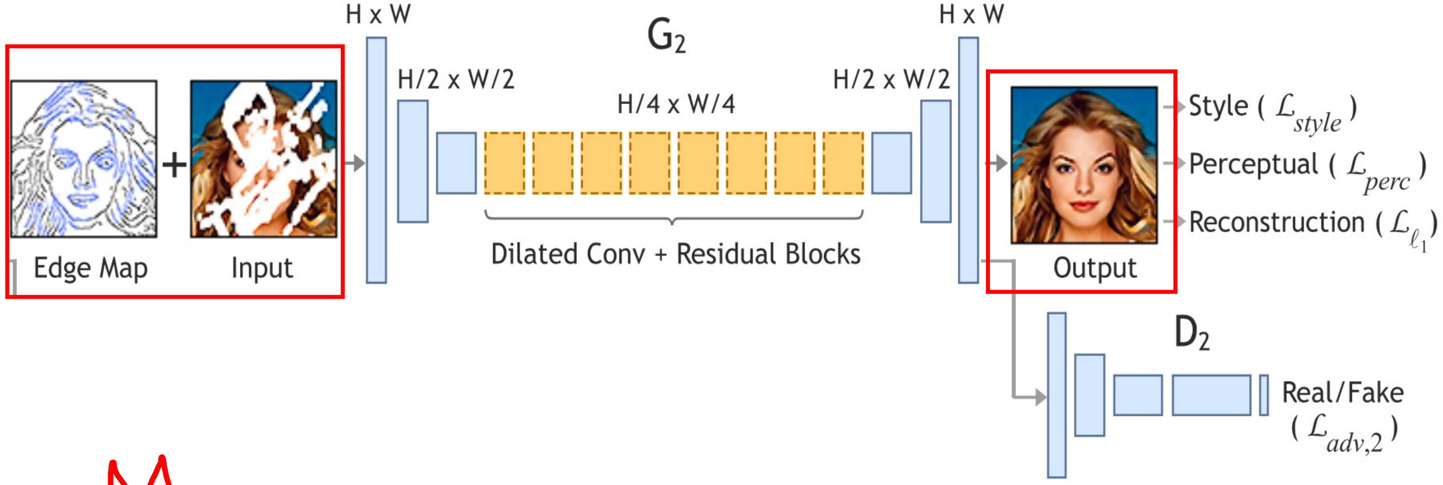
C_{pred}

- Feature-Matching loss

$$\mathcal{L}_{FM} = \mathbb{E} \left[\sum_{i=1}^{L} \frac{1}{N_i} \left\| D_1^{(i)}(\mathbf{C}_{gt}) - D_1^{(i)}(\mathbf{C}_{pred}) \right\|_1 \right]$$

02 Proposed Approach : EdgeConnect

Stage 2. Image Completion



[Input]

M

- Incomplete color image (masked)

$$\tilde{\mathbf{I}}_{gt} = \mathbf{I}_{gt} \odot (\mathbf{1} - \mathbf{M})$$

- Combing background region of GT edge map + generated edge map

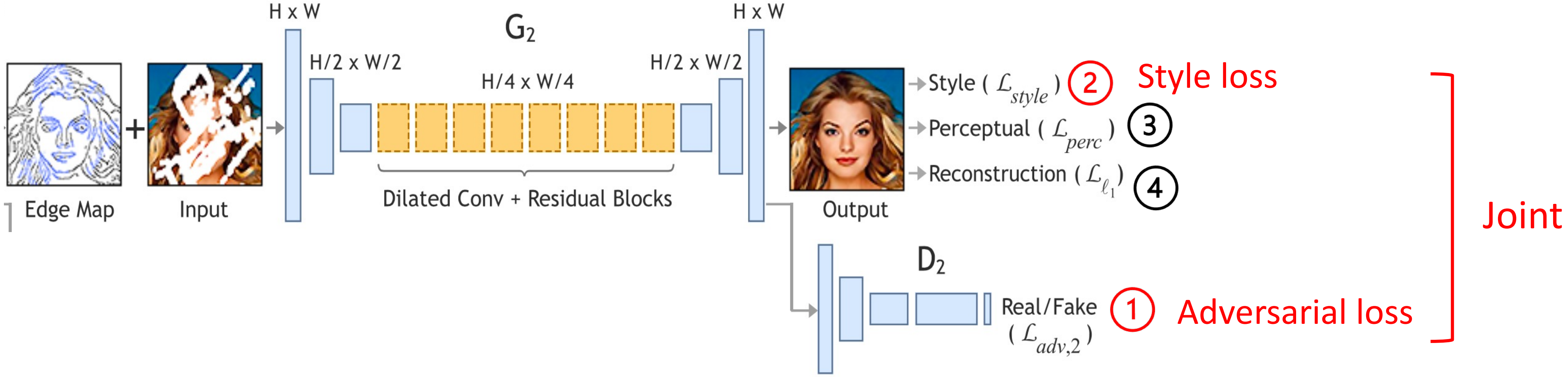
$$\mathbf{C}_{comp} = \mathbf{C}_{gt} \odot (\mathbf{1} - \mathbf{M}) + \mathbf{C}_{pred} \odot \mathbf{M}.$$

[Output]

$$\mathbf{I}_{pred} = G_2\left(\tilde{\mathbf{I}}_{gt}, \mathbf{C}_{comp}\right)$$

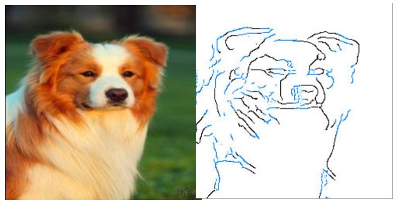
02 Proposed Approach : EdgeConnect

Training loss



- Adversarial loss

$$\mathcal{L}_{adv,2} = \mathbb{E}_{(\mathbf{I}_{gt}, \mathbf{C}_{comp})} [\log D_2(\mathbf{I}_{gt}, \mathbf{C}_{comp})] + \mathbb{E}_{\mathbf{C}_{comp}} \log [1 - D_2(\mathbf{I}_{pred}, \mathbf{C}_{comp})]$$



I_{gt} C_{comp}

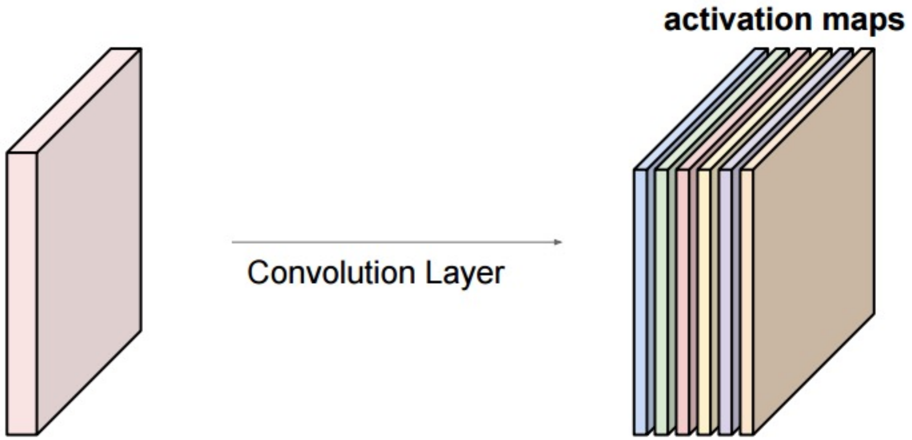
I_{pred}

- Style loss

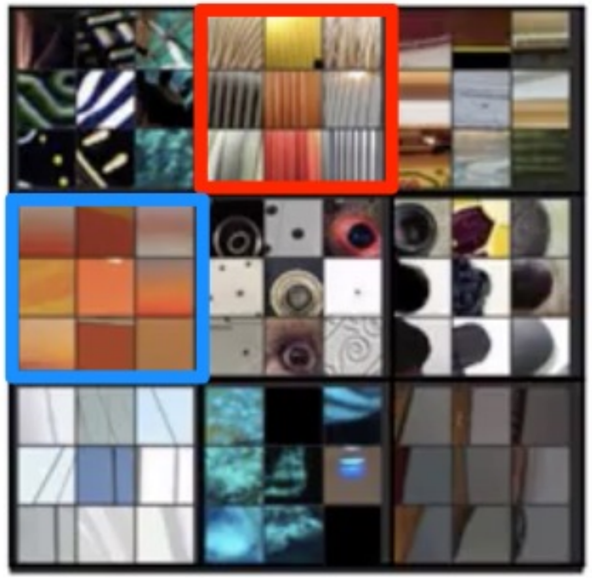
$$\mathcal{L}_{style} = \mathbb{E}_j \left[\| G_j^{\phi}(\tilde{\mathbf{I}}_{pred}) - G_j^{\phi}(\tilde{\mathbf{I}}_{gt}) \|_1 \right]$$

02 Proposed Approach : EdgeConnect

- Style loss

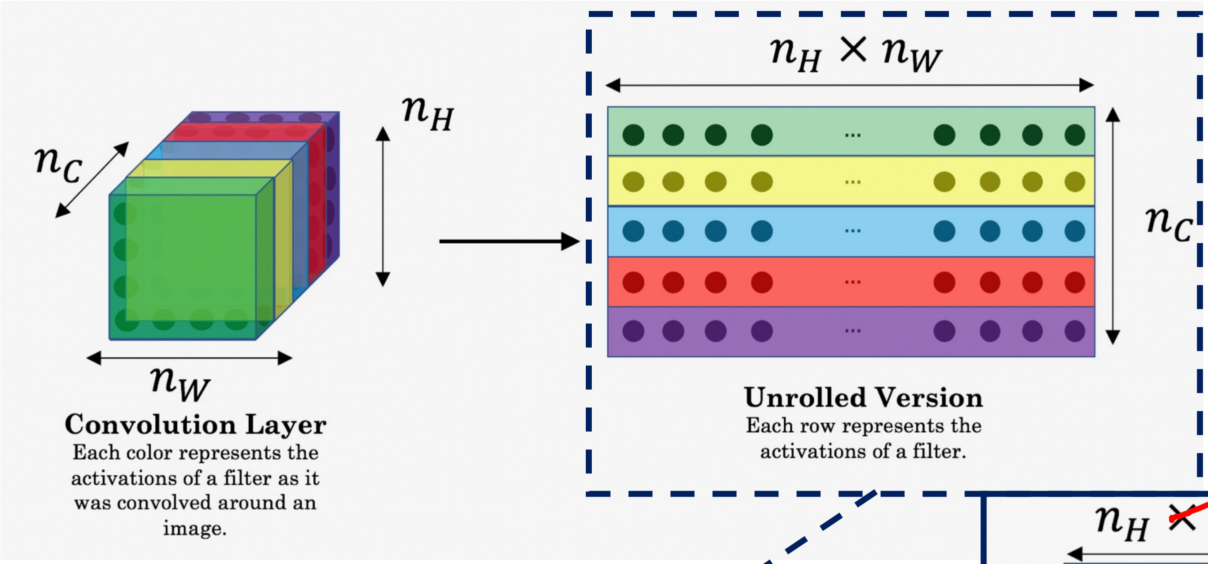


Activation map (Feature map)

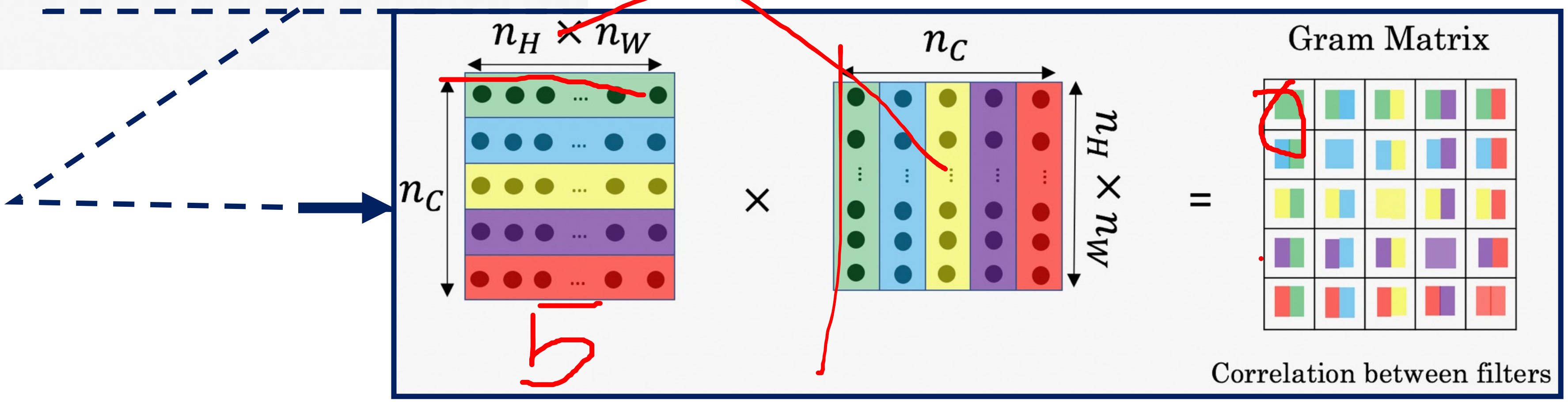


02 Proposed Approach : EdgeConnect

- Style loss

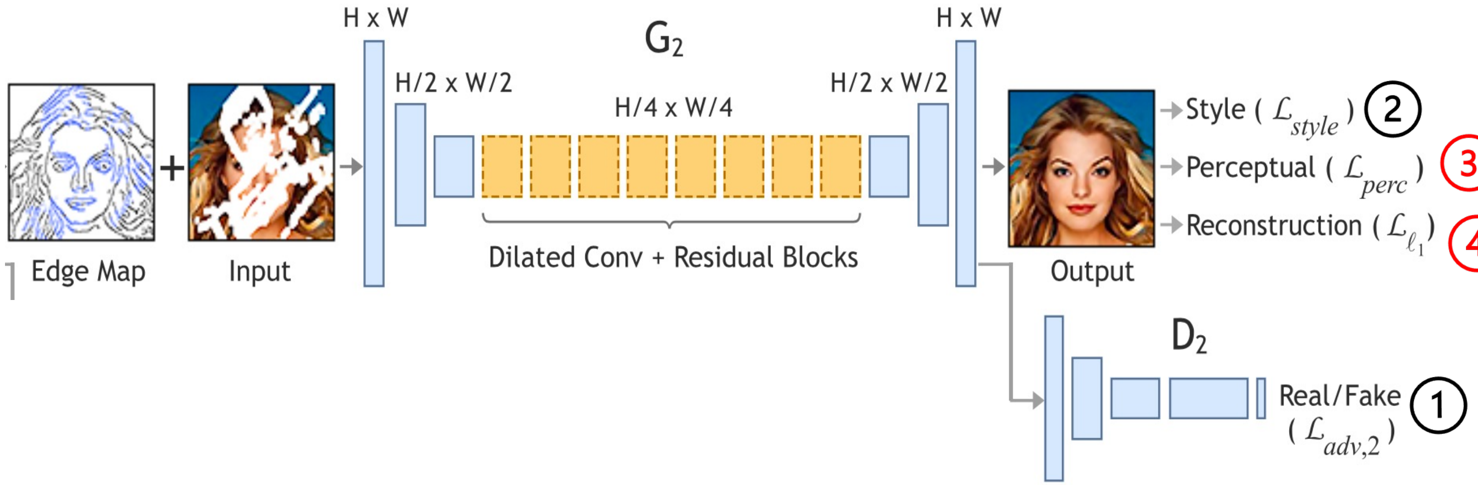


Gram Matrix notation: $G^l_{ij} = \sum_k F^l_{ik} F^l_{jk}.$



02 Proposed Approach : EdgeConnect

Training loss



3 Perceptual loss

4 Reconstruction loss

Joint

- Perceptual loss

$$\mathcal{L}_{perc} = \mathbb{E}\left[\sum_i \frac{1}{N_i} \|\phi_i(\mathbf{I}_{gt}) - \phi_i(\mathbf{I}_{pred})\|_1\right]$$

- Reconstruction loss

l_1 loss

- Final loss (Joint)

$$\mathcal{L}_{G_2} = \lambda_{\ell_1}\mathcal{L}_{\ell_1} + \lambda_{adv,2}\mathcal{L}_{adv,2} + \lambda_p\mathcal{L}_{perc} + \lambda_s\mathcal{L}_{style}$$

03. Experiment / Result

03 Experiment / Result

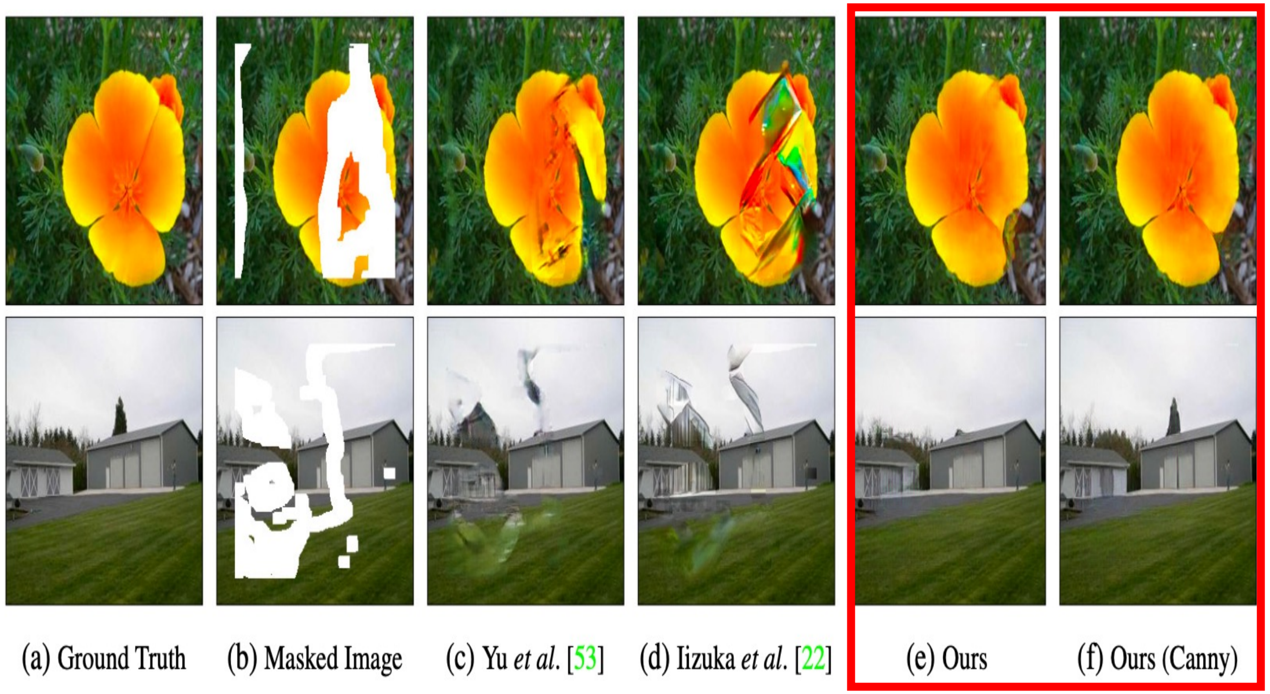


Figure 3: Comparison of qualitative results with existing models. (a) Ground Truth Image. (b) Ground Truth with Mask. (c) Yu *et al.* [53]. (d) Iizuka *et al.* [22]. (e) Ours (end-to-end). (f) Ours (G_2 only with Canny $\sigma = 2$).

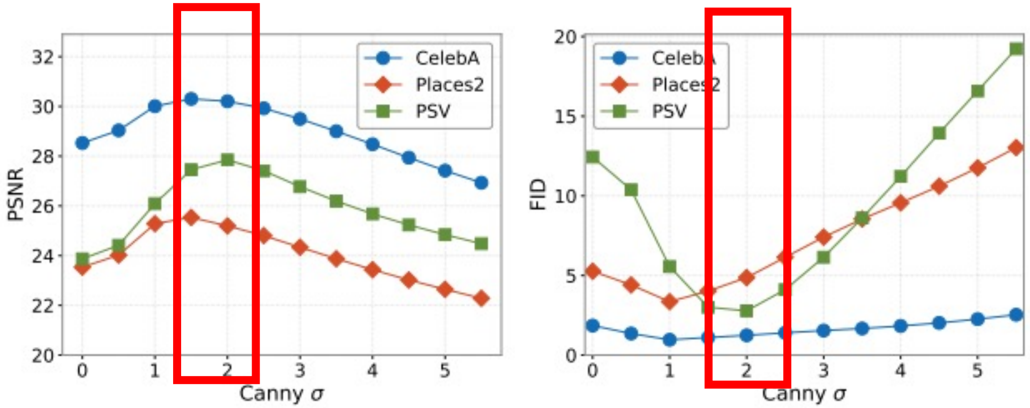


Figure 6: Effect of σ in Canny detector on PSNR and FID.

03 Experiment / Result

Quantitative Result (with Place2 dataset)

- l_1
- SSIM(**Structural Similarity Index Map**)
- PSNR(**Peak Signal-to-Noise Ratio**)
- FID(**Frechet Inception distance**)

	Mask	CA	GLCIC	PConv*	Ours	Canny
ℓ_1 (%)†	10-20%	2.41	2.66	**1.14**	1.50	1.16
	20-30%	4.23	4.70	**1.98**	2.59	1.88
	30-40%	6.15	6.78	**3.02**	3.77	2.60
	40-50%	8.03	8.85	**4.11**	5.14	3.41
	Fixed	4.37	4.12	-	**3.86**	2.22
SSIM*	10-20%	0.893	0.862	0.869	**0.920**	0.941
	20-30%	0.815	0.771	0.777	**0.861**	0.902
	30-40%	0.739	0.686	0.685	**0.799**	0.863
	40-50%	0.662	0.603	0.589	**0.731**	0.821
	Fixed	0.818	0.814	-	**0.823**	0.892
PSNR*	10-20%	24.36	23.49	**28.02**	27.95	30.85
	20-30%	21.19	20.45	24.90	**24.92**	28.35
	30-40%	19.13	18.50	22.45	**22.84**	26.66
	40-50%	17.75	17.17	20.86	**21.16**	25.20
	Fixed	20.65	21.34	-	**21.75**	26.52
FID†	10-20%	6.16	11.84	-	**2.32**	2.25
	20-30%	14.17	25.11	-	**4.91**	3.42
	30-40%	24.16	39.88	-	**8.91**	4.87
	40-50%	35.78	54.30	-	**14.98**	7.13
	Fixed	8.31	8.42	-	**8.16**	3.24

Table 1: Quantitative results over Places2 with models: Contextual Attention (CA) [53], Globally and Locally Consistent Image Completion (GLCIC) [22], Partial Convolution (PConv) [28], G_1 and G_2 (Ours), G_2 only with Canny edges (Canny). The best result of each row is boldfaced except for Canny. *Values taken from the paper [28]. †Lower is better. *Higher is better.

03 Experiment / Result

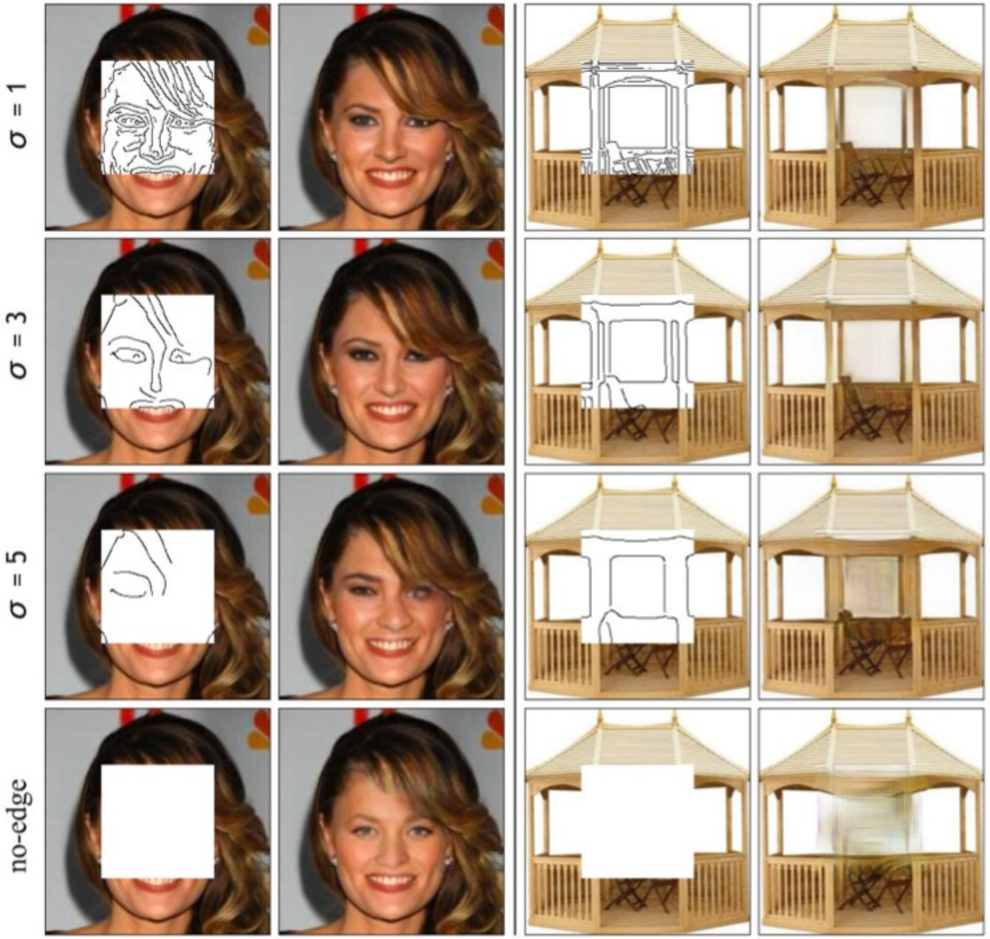


Figure 7: Effect of σ in Canny edge detector on inpainting results. Top to bottom: $\sigma = 1, 3, 5$, no edge data.

04. Discussion / Future Works

04 Discussion / Future Works

Bad Case (Limitation)

(1) Image Restoration

Figure 9: Inpainted results where edge generator fails to produce relevant edge information.

04 Discussion / Future Works

Good Case (Limitation)

(2) Image synthesis

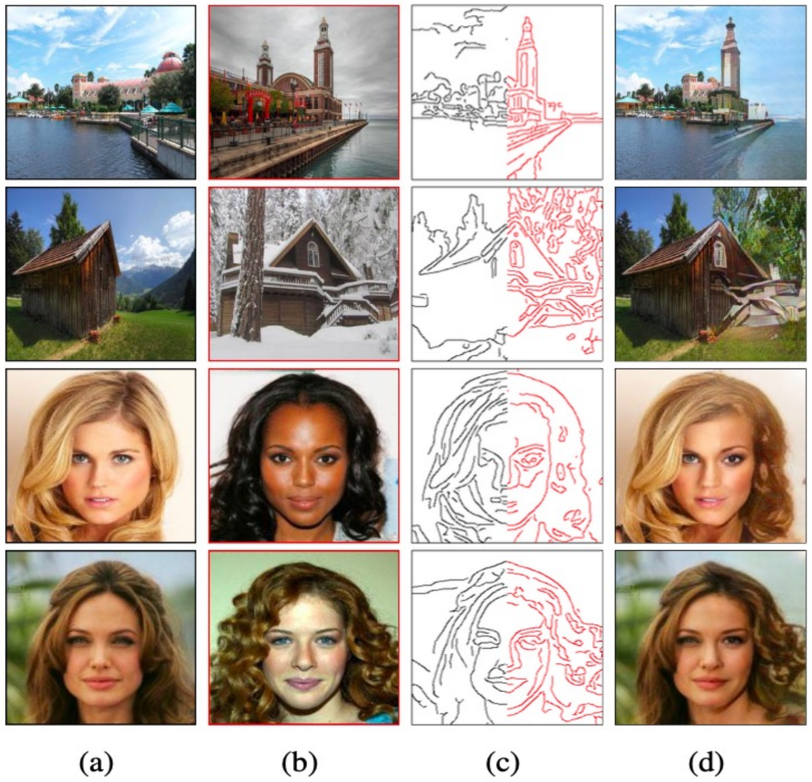


Figure 10: Edge-map (c) generated using the left-half of (a) (shown in black) and right-half of (b) (shown in red). Input is (a) with the right-half removed, producing the output (d).

(3) Object removal

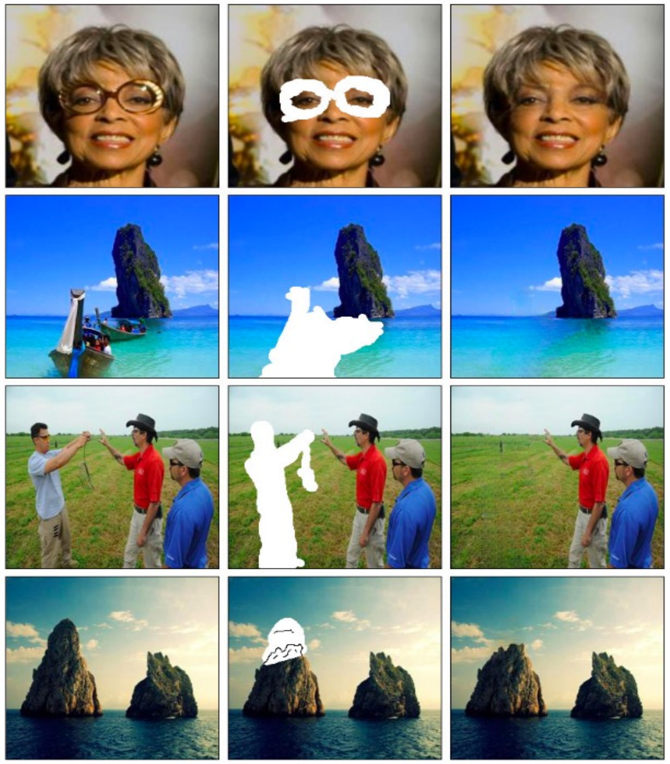


Figure 11: Examples of object removal and image editing using our EdgeConnect model. (Left) Original image. (Center) Unwanted object removed with optional edge information to guide inpainting. (Right) Generated image.

* References

- University of Ontario Institute of Technology, Canada. **Edgeconnect - Generative image inpainting with adversarial edge learning. (*CVPR 2019*)**

- Leon A. Gatys Centre for Integrative Neuroscience, University of Tubingen, Germany. **Image Style Transfer Using Convolutional Neural Networks. (*CVPR 2016*)**

- J. Johnson, A. Alahi, and L. Fei-Fei. **Perceptual losses for real-time style transfer and super-resolution. In *European Conference on Computer Vision (ECCV 2016)***

Q & A

Q&A

감사합니다.

E-mail : kka960602@gmail.com
Github :https://github.com/Kyeonga-Kim/