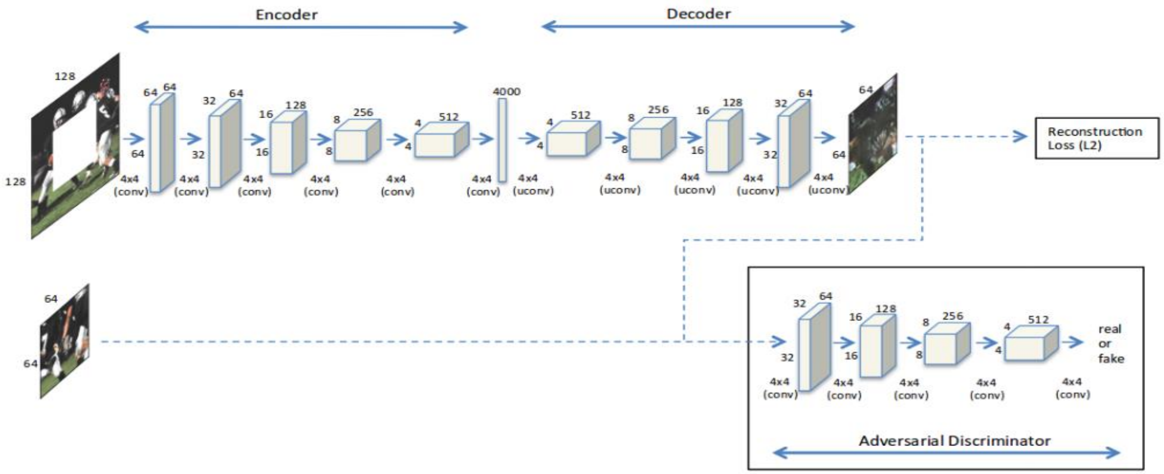
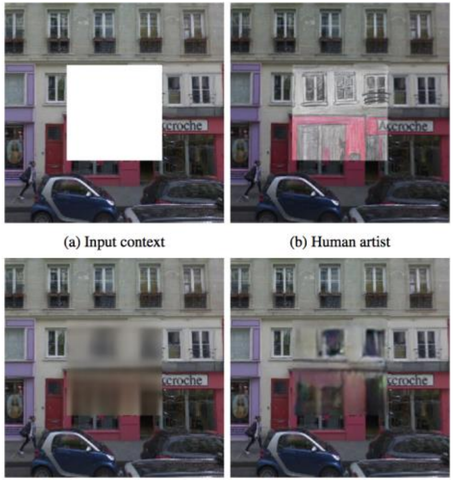
Free-Form Image Inpainting with Gated Convolution (ICCV 2019)

EdgeConnect: Structure Guided Image Inpainting using Edge Prediction (ICCVW 2019)

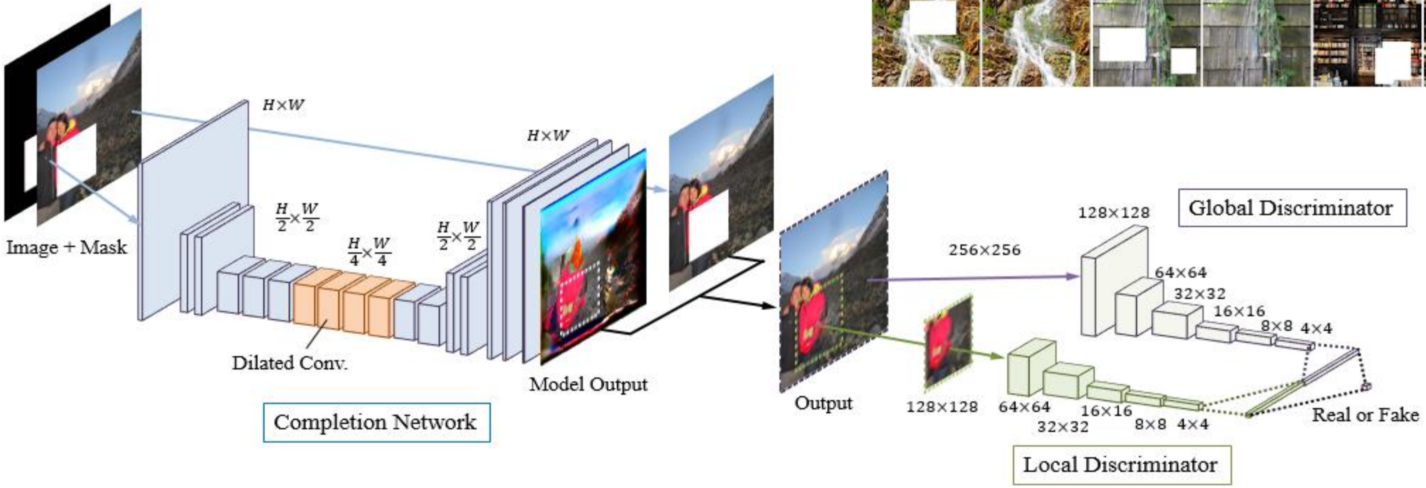
Context Encoders: Feature Learning by Inpainting (CVPR 2016)





(a) Input context
(b) Human artist
(c) Context Encoder ($L2$ loss)
(d) Context Encoder ($L2$ + Adversarial loss)

Globally and Locally Consistent Image Completion (SIGGRAPH 2017)

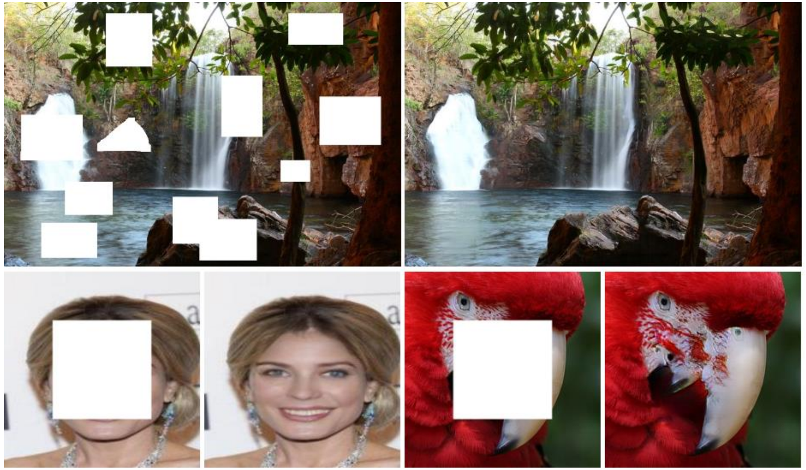


Generative Image Inpainting with Contextual Attention (CVPR 2018)

Generative Image Inpainting with Contextual Attention

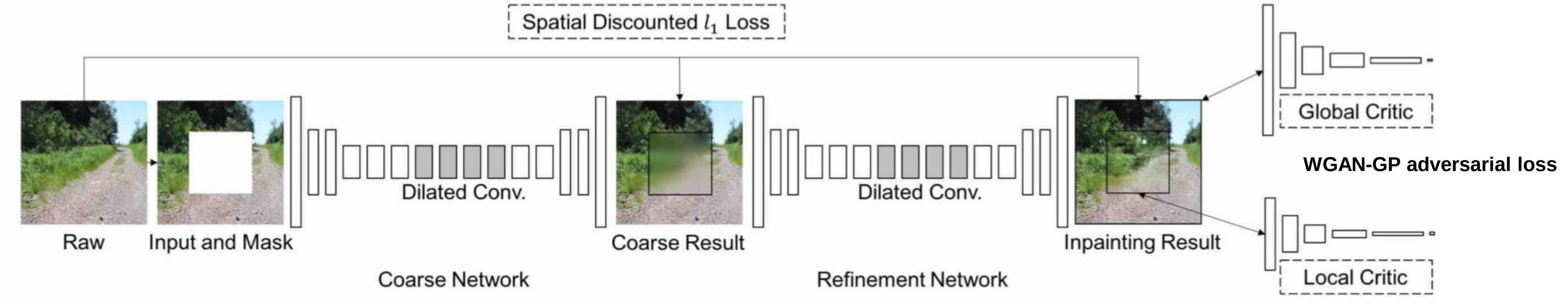
Jiahui Yu[1] Zhe Lin[2] Jimei Yang[2] Xiaohui Shen[2] Xin Lu[2] Thomas S. Huang[1]

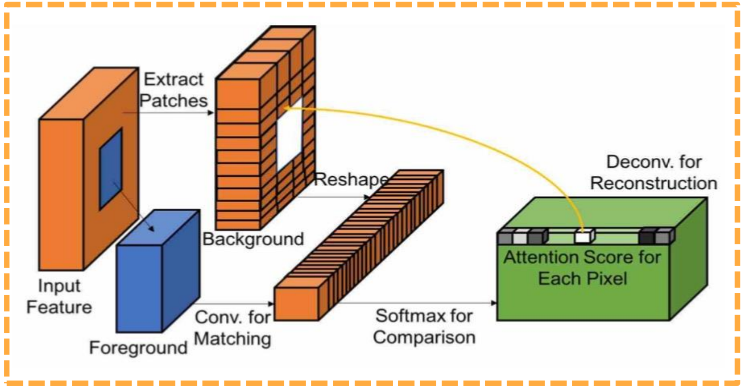
[1]University of Illinois at Urbana-Champaign
[2]Adobe Research

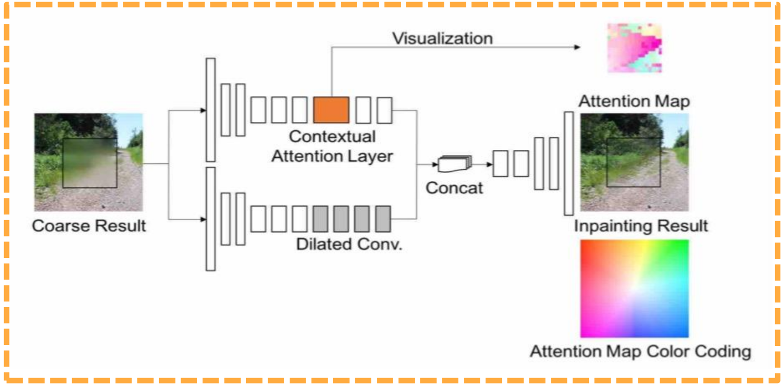


Generative Image Inpainting with Contextual Attention (CVPR 2018)

Deepfill v1 / Contextual Attention





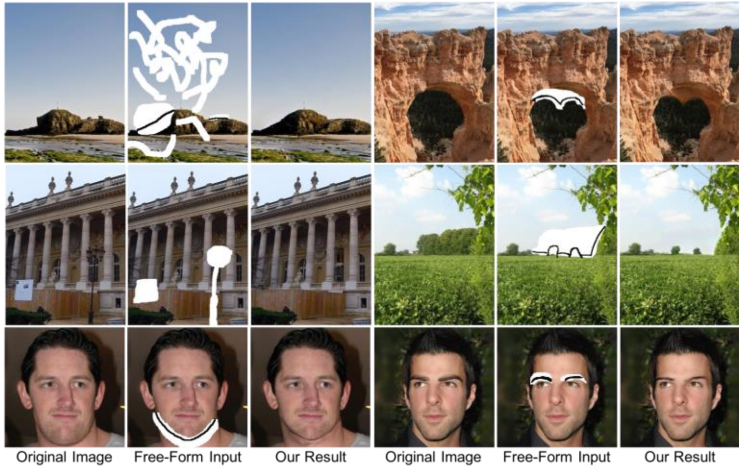


Free-Form Image Inpainting with Gated Convolution (ICCV 2019)

Free-Form Image Inpainting with Gated Convolution

Jiahui Yu[1] Zhe Lin[2] Jimei Yang[2] Xiaohui Shen[3] Xin Lu[2] Thomas Huang[1]

[1]University of Illinois at Urbana-Champaign
[2]Adobe Research
[3]ByteDance AI Lab



Free-Form Image Inpainting with Gated Convolution (ICCV 2019)

Deepfill v2 / Gated Convolution

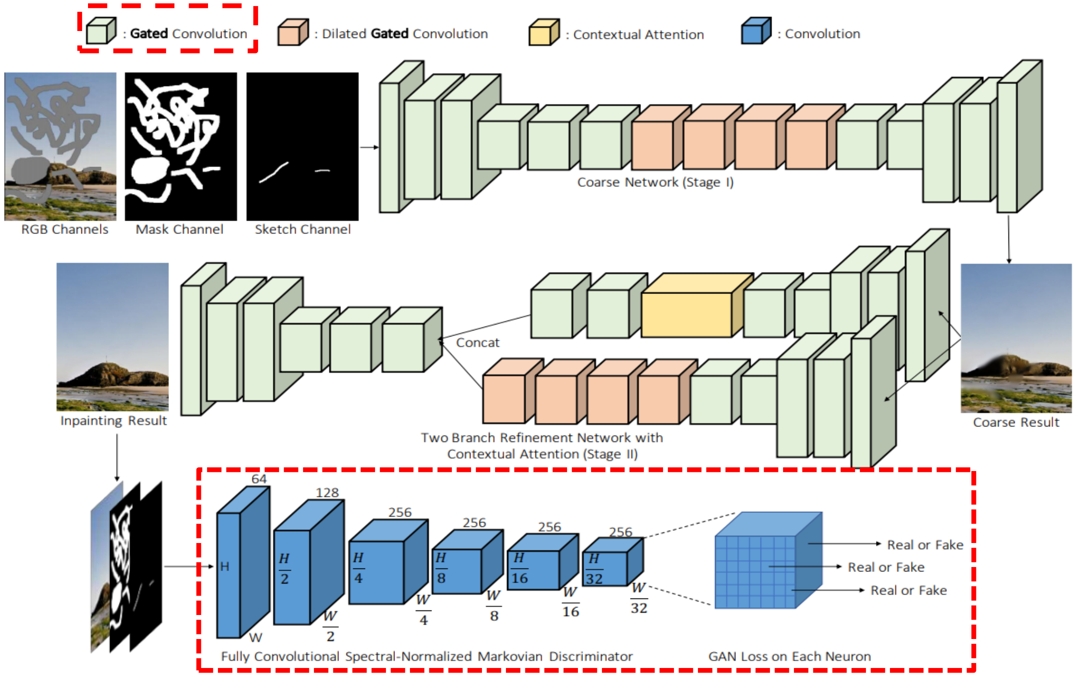


Figure 3: Overview of our framework with gated convolution and SN-PatchGAN for free-form image inpainting.

Free-Form Image Inpainting with Gated Convolution (ICCV 2019)

$$O_{y,x} = \sum_{i=-k'_h}^{k'_h} \sum_{j=-k'_w}^{k'_w} W_{k'_h+i,k'_w+j} \cdot I_{y+i,x+j},$$

vanilla conv

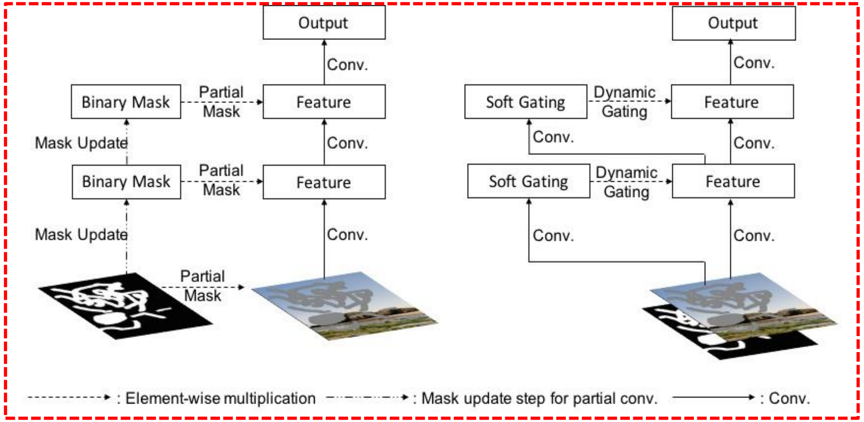
$$O_{y,x} = \begin{cases} \sum\sum W \cdot (I \odot \frac{M}{sum(M)}), & \text{if sum(M)} > 0 \\ 0, & \text{otherwise} \end{cases}$$

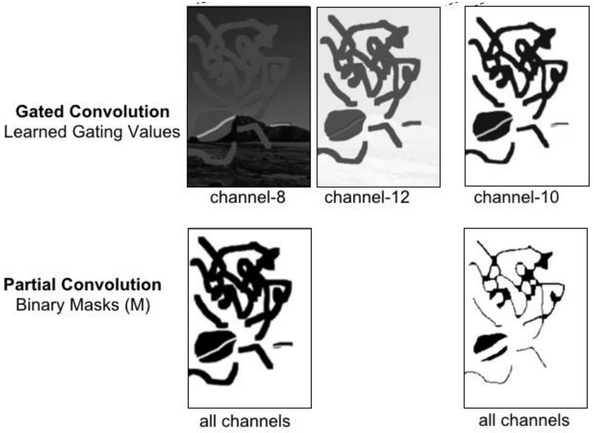
$$m'_{y,x} = \begin{cases} 1, & \text{if sum(M)} > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Partial conv (ECCV 2018)

$$Gating_{y,x} = \sum\sum W_g \cdot I$$
$$Feature_{y,x} = \sum\sum W_f \cdot I$$
$$O_{y,x} = \phi(Feature_{y,x}) \odot \sigma(Gating_{y,x})$$

Gated conv





Free-Form Image Inpainting with Gated Convolution (ICCV 2019)

SN-PatchGAN (Spectral-Normalized Markovian Discriminator)

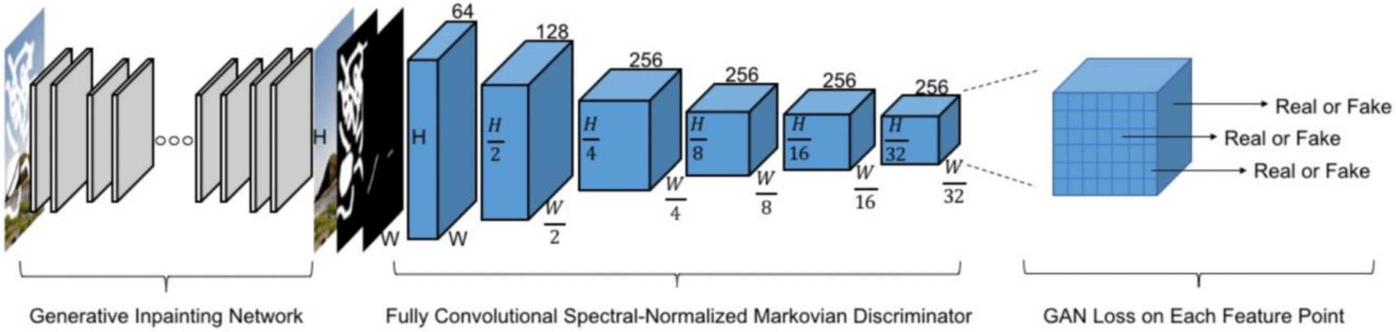


Fig. 4. Overview of SN-PatchGAN and our architecture for learning free-form image inpainting network. Details of generative inpainting network is shown in Figure 3. We aim at free-form image inpainting for which the masks may appear anywhere in images with any shapes. Previous global and local GANs [Iizuka et al. 2017] designed for a single rectangular mask are not suitable. Thus we introduce SN-PatchGAN that directly applies GAN loss for each point in output feature map of convolutional discriminator. It is simple in formulation, fast and stable for training and produces high-quality inpainting results. Note that we use convolutional kernel size 5×5 and the receptive fields of each point in output map can still cover entire input image in our training setting thus a global GAN is not used.

Free-Form Image Inpainting with Gated Convolution (ICCV 2019)

	rectangular mask		free-form mask	
Method	ℓ_1 err.	ℓ_2 err.	ℓ_1 err.	ℓ_2 err.
PatchMatch [3]	16.1%	3.9%	11.3%	2.4%
Global&Local [15]	9.3%	2.2%	21.6%	7.1%
ContextAttention [49]	**8.6%**	2.1%	17.2%	4.7%
PartialConv* [23]	9.8%	2.3%	10.4%	1.9%
Ours	**8.6%**	**2.0%**	**9.1%**	**1.6%**

Free-Form Image Inpainting with Gated Convolution (ICCV 2019)

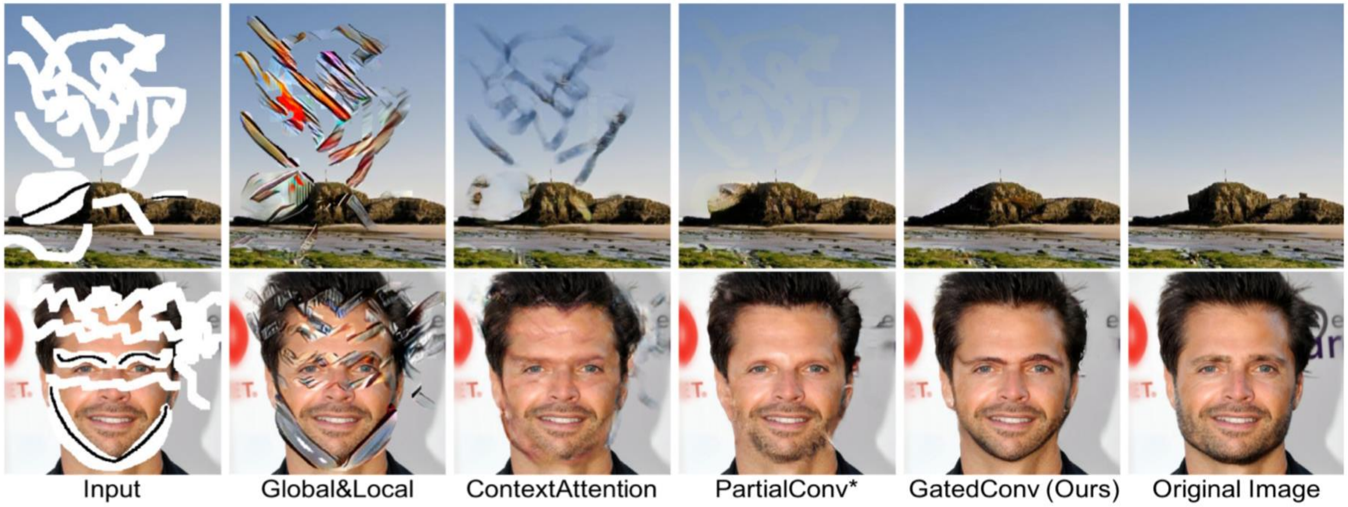
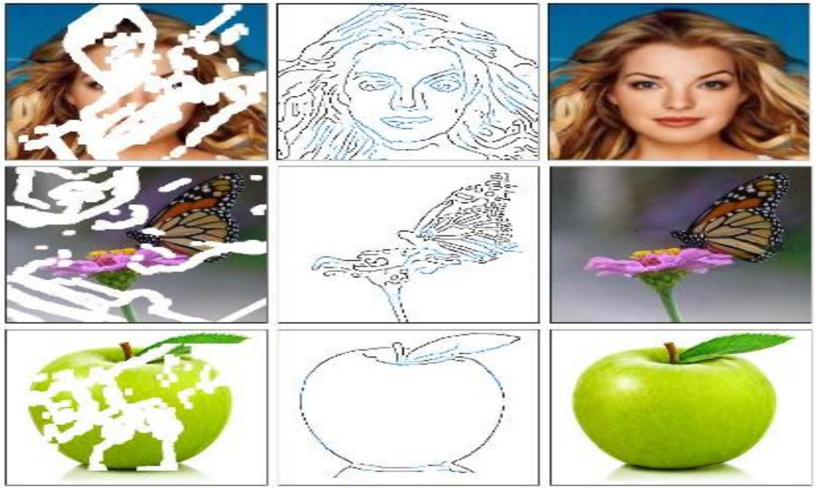


Fig. 7. Qualitative Comparisons on the Places2 and CelebA-HQ validation sets.

EdgeConnect (ICCVW 2019)

EdgeConnect: Structure Guided Image Inpainting using Edge Prediction

Kamyar Nazeri, Eric Ng, Tony Joseph, Faisal Z. Qureshi, and Mehran Ebrahimi
University of Ontario Institute of Technology, Canada
{kamyar.nazeri, eric.ng, tony.joseph, faisal.qureshi, mehran.ebrahimi}@uoit.ca
http://www.ImagingLab.ca http://www.VCLab.ca



EdgeConnect (ICCVW 2019)

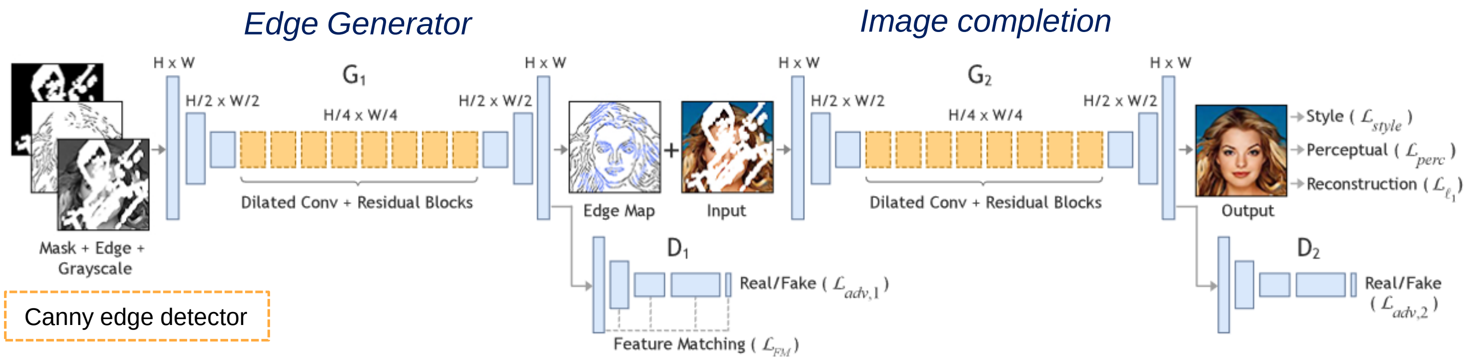


Figure 2: Summary of our proposed method. Incomplete grayscale image and edge map, and mask are the inputs of G_1 to predict the full edge map. Predicted edge map and incomplete color image are passed to G_2 to perform the inpainting task.

EdgeConnect (ICCVW 2019)

Edge Generator

Image completion

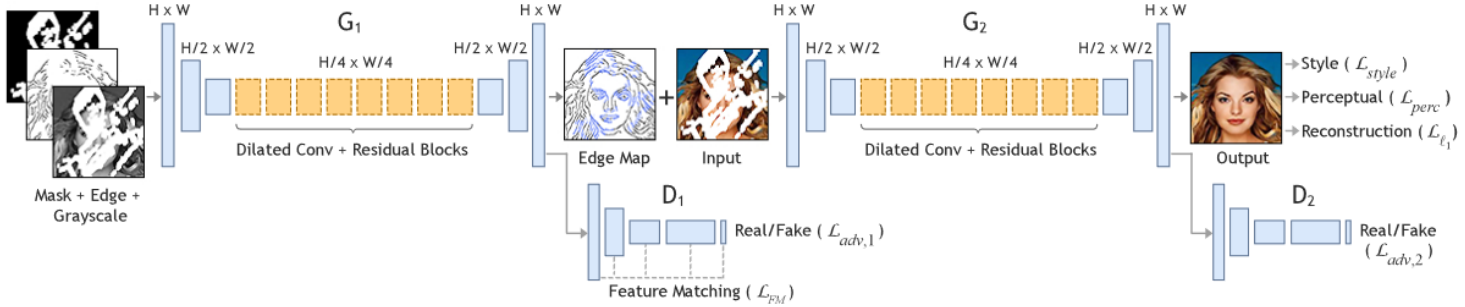


Figure 2: Summary of our proposed method. Incomplete grayscale image and edge map, and mask are the inputs of G_1 to predict the full edge map. Predicted edge map and incomplete color image are passed to G_2 to perform the inpainting task.

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

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

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

$$\mathcal{L}_{G_1} = -\mathbb{E}_{\mathbf{I}_{gray}}\left[D_1(\mathbf{C}_{pred}, \mathbf{I}_{gray})\right],$$

$$\mathcal{L}_{D_1} = \mathbb{E}_{(\mathbf{C}_{gt}, \mathbf{I}_{gray})}\left[\max(0, 1 - D_1(\mathbf{C}_{gt}, \mathbf{I}_{gray}))\right] + \mathbb{E}_{\mathbf{I}_{gray}}\left[\max(0, 1 + D_1(\mathbf{C}_{pred}, \mathbf{I}_{gray}))\right].$$

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

$$\boxed{\mathcal{J}_{G_1} = \lambda_{G_1}\mathcal{L}_{G_1} + \lambda_{FM}\mathcal{L}_{FM}}$$

EdgeConnect (ICCVW 2019)

Edge Generator

Image completion

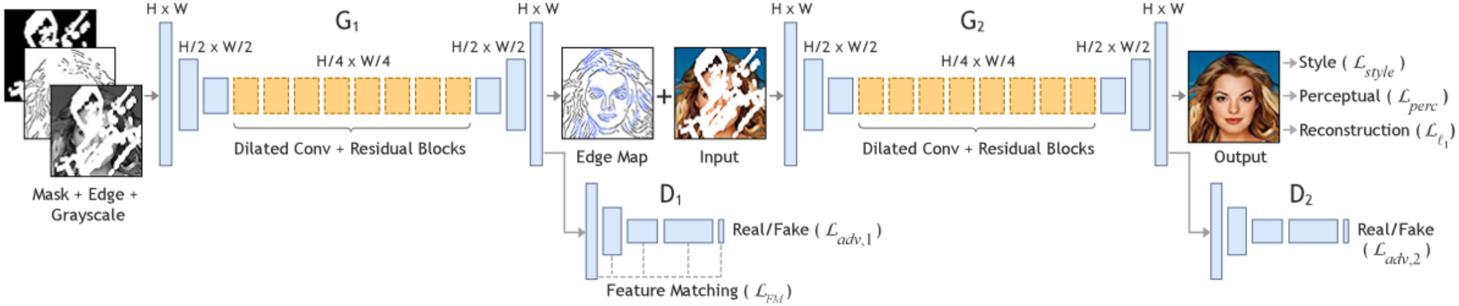


Figure 2: Summary of our proposed method. Incomplete grayscale image and edge map, and mask are the inputs of G_1 to predict the full edge map. Predicted edge map and incomplete color image are passed to G_2 to perform the inpainting task.

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

$$\mathcal{L}_{G_2} = -\mathbb{E}_{\mathbf{C}_{comp}}\left[D_2(\mathbf{I}_{pred}, \mathbf{C}_{comp})\right],$$

$$\mathcal{L}_{D_2} = \mathbb{E}_{(\mathbf{I}_{gt}, \mathbf{C}_{comp})}\left[\max(0, 1 - D_2(\mathbf{I}_{gt}, \mathbf{C}_{comp}))\right] + \mathbb{E}_{\mathbf{C}_{comp}}\left[\max(0, 1 + D_2(\mathbf{I}_{pred}, \mathbf{C}_{comp}))\right].$$

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

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

$$\boxed{\mathcal{J}_{G_2} = \lambda_{\ell_1}\mathcal{L}_{\ell_1} + \lambda_{G_2}\mathcal{L}_{G_2} + \lambda_p\mathcal{L}_{perc} + \lambda_s\mathcal{L}_{style}.}$$

EdgeConnect (ICCVW 2019)

	Mask	CA	GLCIC	PConv	Ours
ℓ_1 (%)†	10-20%	2.41	2.66	1.55	**1.50**
	20-30%	4.23	4.70	2.71	**2.59**
	30-40%	6.15	6.78	3.94	**3.77**
	40-50%	8.03	8.85	5.35	**5.14**
	Fixed	4.37	4.12	3.95	**3.86**
SSIM*	10-20%	0.893	0.862	0.916	**0.920**
	20-30%	0.815	0.771	0.854	**0.861**
	30-40%	0.739	0.686	0.789	**0.799**
	40-50%	0.662	0.603	0.720	**0.731**
	Fixed	0.818	0.814	0.818	**0.823**
PSNR*	10-20%	24.36	23.49	27.54	**27.95**
	20-30%	21.19	20.45	24.47	**24.92**
	30-40%	19.13	18.50	22.42	**22.84**
	40-50%	17.75	17.17	20.77	**21.16**
	Fixed	20.65	21.34	21.54	**21.75**
FID†	10-20%	6.16	11.84	**2.26**	2.32
	20-30%	14.17	25.11	**4.88**	4.91
	30-40%	24.16	39.88	**8.84**	8.91
	40-50%	35.78	54.30	15.18	**14.98**
	Fixed	8.31	8.42	10.53	**8.16**

	Mask	CA	GLCIC	PConv	Ours
JND (%)	10-20%	20.98	16.91	36.04	**39.69**
	20-30%	15.45	14.27	30.09	**36.99**
	30-40%	12.86	12.29	20.60	**27.53**
	40-50%	12.74	10.91	18.31	**25.44**
Y-N (%)	10-20%	38.71	22.46	79.72	**88.66**
	20-30%	23.44	12.09	64.11	**77.59**
	30-40%	13.49	4.32	52.50	**66.44**
	40-50%	9.89	2.77	37.73	**58.02**

Edges	CelebA		Places2	
	No	Yes	No	Yes
ℓ_1 (%)	4.11	3.03	6.69	5.14
SSIM	0.802	0.846	0.682	0.731
PSNR	23.33	25.28	19.59	21.16
FID	6.16	2.82	32.18	14.98

StructureFlow: Image Inpainting via ***Structure-aware Appearance Flow***(ICCV 2019)

Image Inpainting with Learnable Bidirectional ***Attention*** Maps(ICCV 2019)

Coherent Semantic ***Attention*** for Image Inpainting(ICCV 2019)

MUSICAL: Multi-Scale Image Contextual ***Attention*** Learning for Inpainting(IJCAI 2019)

Coarse-to-Fine Image Inpainting via Region-wise Convolutions and ***Non-Local Correlation***(In IJCAI 2019)

Question ?