Image Inpainting

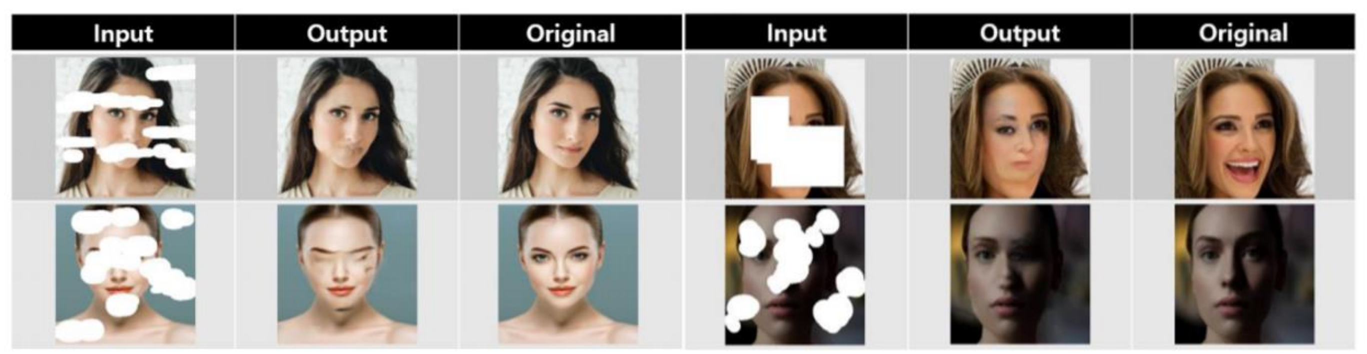
Jin Woo, Baik
Stanford University

Motivation

- **Image Inpainting**

- A task which fills missing pixels with semantically and perceptually plausible contents.

- **Problem**



- Many inpainting methods gives blurry or awkward images. From this project, I will generate clear and plausible outputs to improve image quality.

New Technique

- **Recurrent Convolution**

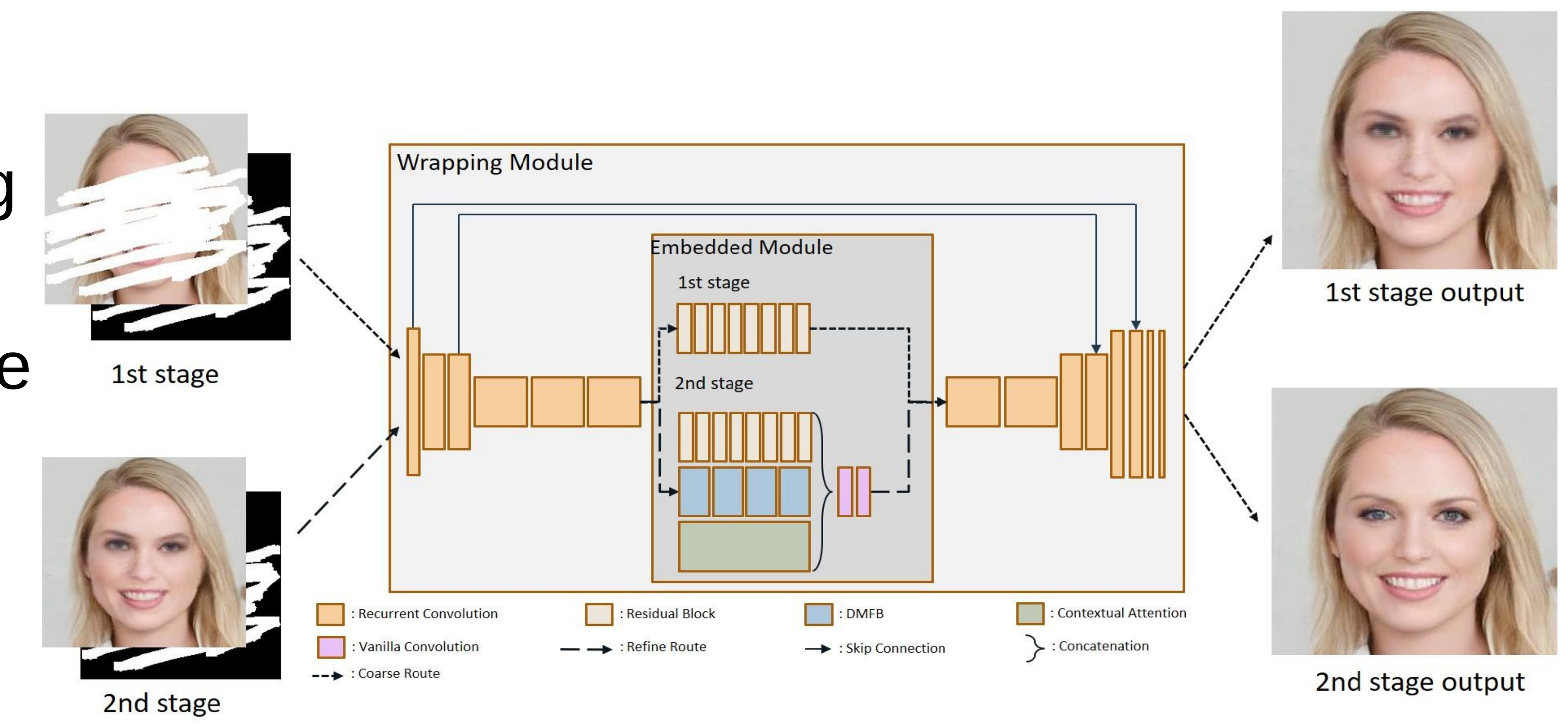
- Implement RNN structure by using buffer
- Reduce the network size & solve gradient vanishing

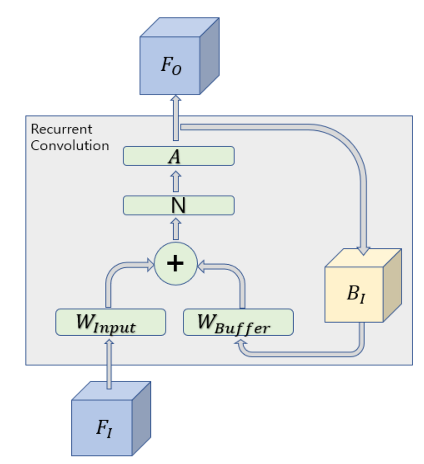
- **Sobel Edge Discriminator**

- Sobel edge discriminator distinguishes real and fake edge map of generated image and the ground truth

- **Frequency Separation Loss**

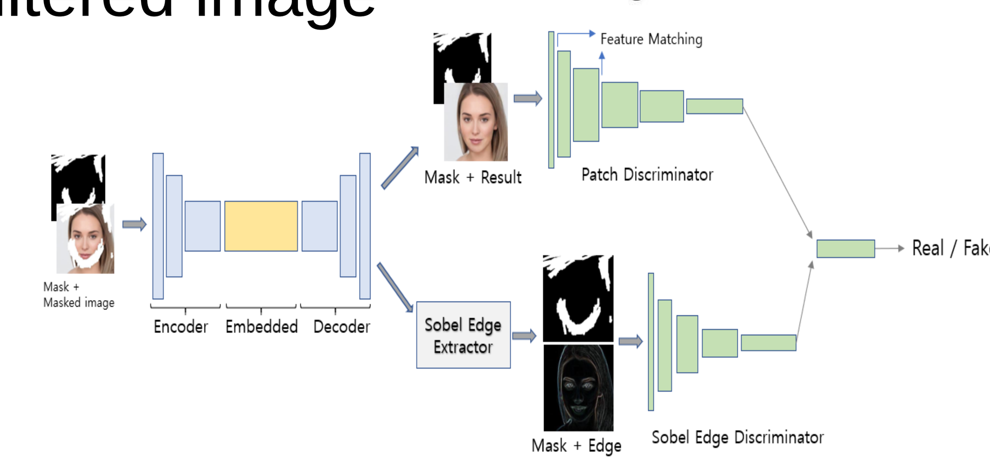
- Coarse Route : L1 loss of low pass filtered image
- Refine Route : L1 loss of high pass filtered image

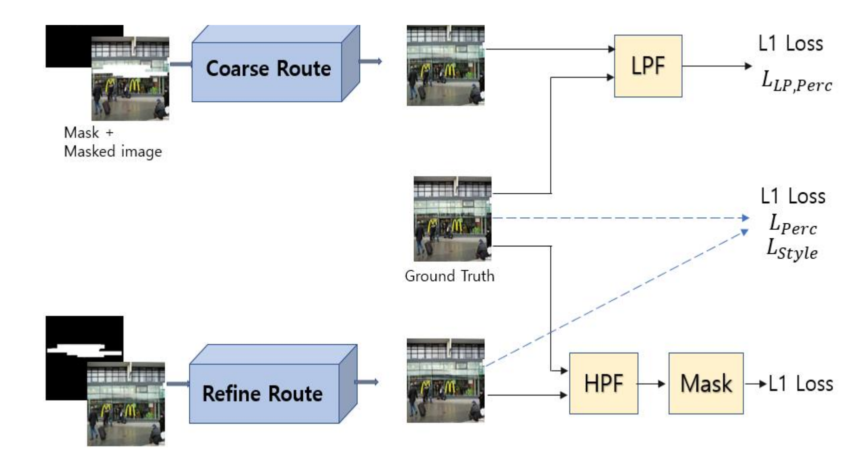




F_I: Input Features
F_O: Output Features
B_I: Input Buffer
W_{Input}: Conv for Input Features
W_{Buffer}: Conv for Buffer
N: Instance Normalization
A: ELU activation function

$F_O = B_I$
$F_O = A[(F_I * W_{Input}) + (B_I * W_{Buffer})]$





Related Work

[1] Coarse to Refine

Coarse to Refine network first produces coarse output and use this information to produce refined image.

[2] RNN

Recurrent Neural Network(RNN) is designed to deal with sequential data such as music, movie, and natural language. The main difference with other neural networks is that RNN uses past information by adding hidden states.

[3] Edgeconnect

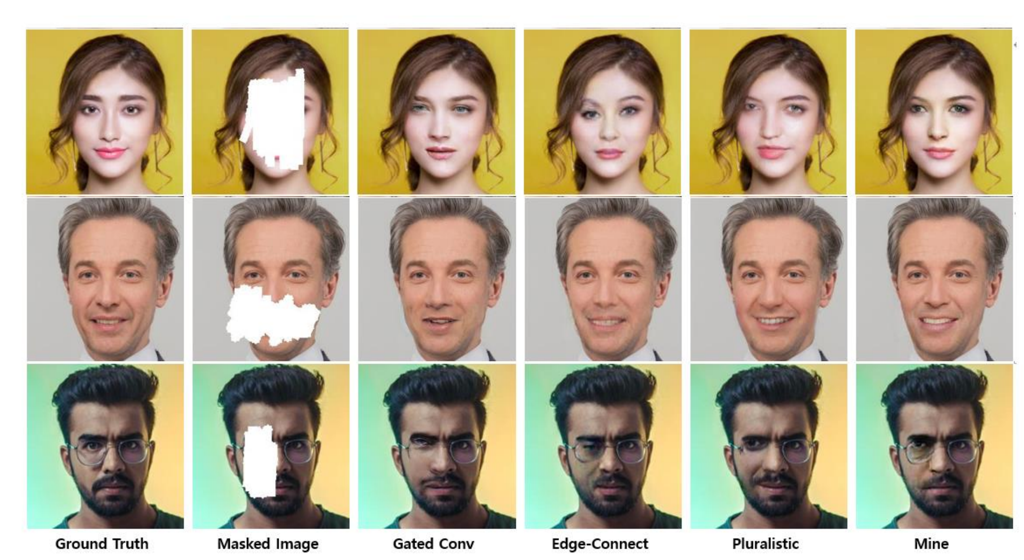
Edgeconnect uses edge information for inpainting. Edgeconnect first produces edge map of the masked region and use this information to complete the image.

References

[1] Yuqing Ma, Xianglong Liu, Shihao Bai, Lei Wang, "Coarse-to-Fine-Image Inpainting via Region- wise Convolution and Non-Local Correlation" in Twenty-Eighth International Joint Conference on Artificial Intelligence, 2019
[2] Sherstinsky, Alex. "Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network." Physica D: Nonlinear Phenomena 404 (2020): 132306.
[3] K. Nazeri, E. Ng, T. Joseph, F. Qureshi, and M. Ebrahimi, "Edgeconnect: Generative image inpainting with adversarial edge learning. arxiv 2019," arXiv preprint arXiv:1901.00212.

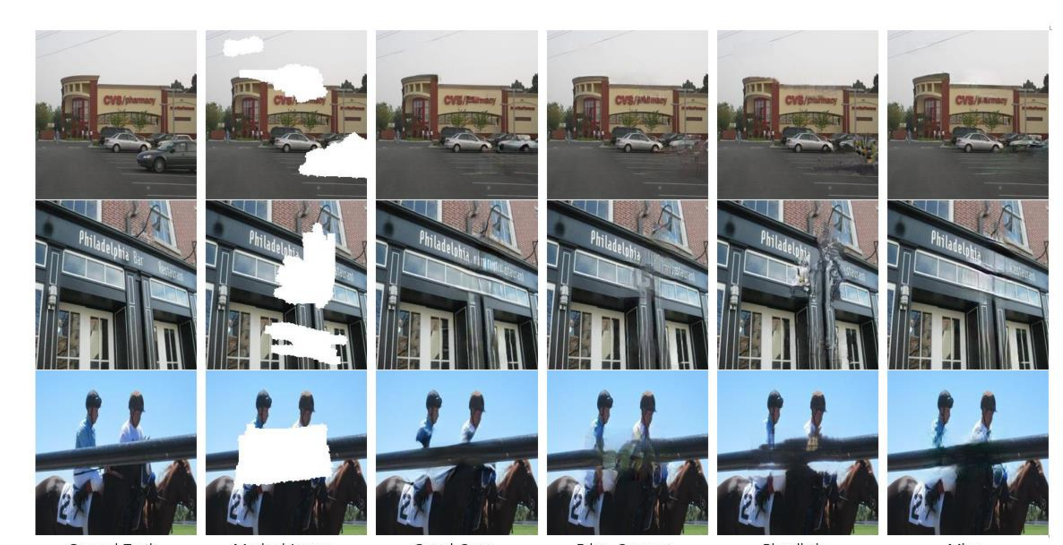
Experimental Results

CelebaHQ



Mask rate (%)	10~20				20~30				30~40			
Type	Gated	Edge	Plural	Ours	Gated	Edge	Plural	Ours	Gated	Edge	Plural	Ours
L1 error (%) ▼	4.95	5.79	4.11	5.76	6.54	6.70	6.32	6.54	9.86	9.19	10.12	8.76
L2 error (%) ▼	1.47	1.32	1.47	1.28	2.03	1.54	2.72	1.43	4.38	2.79	4.81	2.58
Perceptual Loss▼	0.431	0.554	0.358	0.542	0.592	0.632	0.512	0.604	0.813	0.794	0.739	0.762
PSNR ▲	29.575	29.803	29.693	29.920	27.853	28.748	27.847	29.152	24.026	25.888	24.468	26.012
SSIM ▲	0.933	0.917	0.941	0.920	0.904	0.898	0.912	0.905	0.849	0.854	0.854	0.863

Places 365



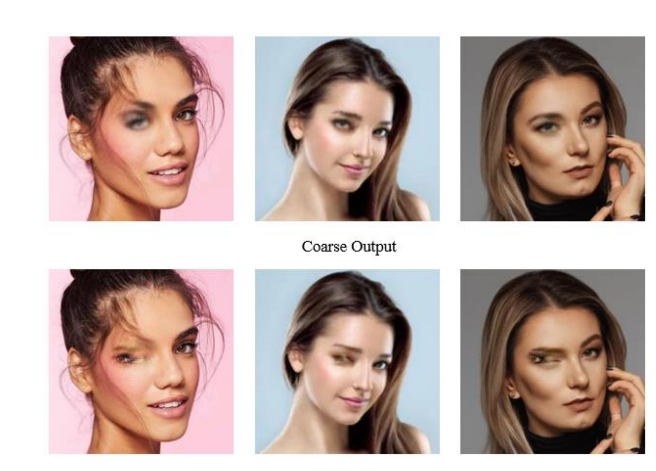
Mask rate (%)	10~20				20~30				30~40			
Type	Gated	Edge	Plural	Ours	Gated	Edge	Plural	Ours	Gated	Edge	Plural	Ours
L1 error (%) ▼	14.25	12.70	14.60	11.82	16.47	15.52	18.89	14. 02	28.24	27.55	32.087	24.840
L2 error (%) ▼	12.25	10.03	10.53	9.99	12.80	11.17	13.89	10.53	25.62	22.06	26.217	19.999
Perceptual Loss▼	0.682	0.592	0.718	0.546	0.927	0.847	1.001	0.775	1.311	1.205	1.413	1.118
PSNR ▲	23.311	24.112	22.835	24.237	21.617	22.270	21.313	22.572	19.114	20.104	18.890	20.308
SSIM ▲	0.899	0.898	0.880	0.907	0.858	0.856	0.832	0.871	0.769	0.768	0.729	0.789

Conclusion

- **Analysis**

My model produces better outputs than other methods for both face and place dataset. The outputs are more natural and give better PSNR, SSIM, and Loss

- **Future Work**



However, there are some cases which coarse outputs are better than refine outputs. There might be a problem with refine network. Hence, the refine module can be modified to solve this issue.