Summary of the project proposal

Team members: Huong Doan and Yunting Chiu

Basically, image inpainting is to recover the damaged image or fill out a missing part of an image. There are a lot of existing image inpainting methods which are published with or without codes. Those methods will have their own advantages and disadvantages. In addition, those methods usually work well on the datasets which were used for the experiments on the paper but it does not guarantee that those methods can work well for other datasets as well. Several related readings are Texture Memory-Augmented Deep Patch-Based Image Inpainting, Image Fine-grained Inpainting and Generative Adversarial Networks.

In our study, we choose an existing method which we are interested in and apply it to the interesting datasets we choose to observe its performance as well as its limitation if there is any. We choose the method proposed in the **Pluralistic Image Completion** paper to investigate on the two datasets taken from Flickr-Faces-HQ Dataset (FFHQ) and City Space dataset. The main challenging parts are to figure out how to run the github code using our datasets and how to evaluate the performance of our experiments. For the experiments, we are planning to change the number of iterations in the train file to observe whether the results are different. We will visualize the recovered images and the original images without missing parts. The expected results are recovered images but the methodology of the paper does not guarantee that those images exactly look like the original images without missing parts. We are also planning to use loss plots for the performance evaluation. However, this is not the final decision yet.

In the Pluralistic Image Completion paper, the author proposed the methodology of generating multiple and diverse plausible solutions for image completion. The architecture of the method is shown below.

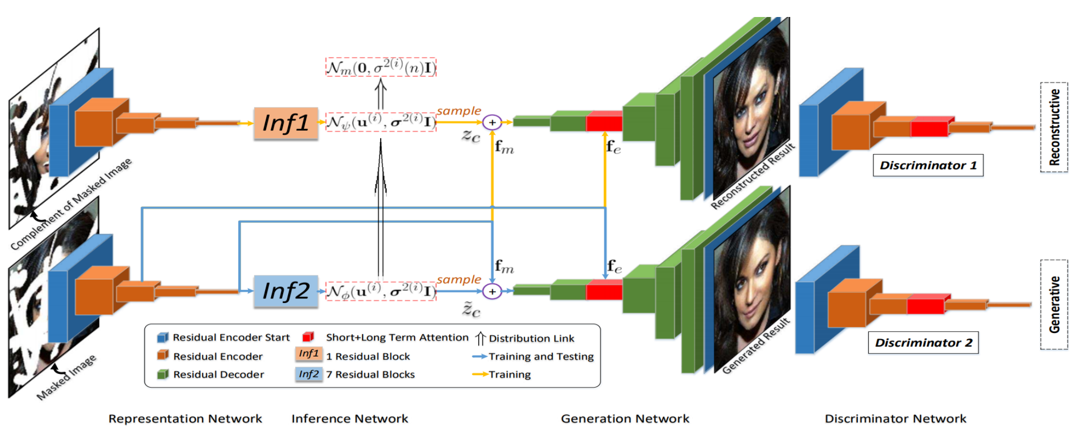

Figure 3. Overview of our architecture with two parallel pipelines. The **reconstructive** pipeline (yellow line) combines information from $\mathbf{I}_m$ and $\mathbf{I}_c$, which is used only for training. The **generative** pipeline (blue line) infers the conditional distribution of hidden regions, that can be sampled during testing. Both representation and generation networks share identical weights.

This neural network consists of two parallel pipelines. The yellow line (reconstructive) merges data from Im and Ic, which are only used for training purposes. The blue (generative) pipeline estimates the conditional distribution of hidden regions, which can then be sampled during testing, that is, Ig = {Ic, Im}.

- Ig is the original image, and Im is the masked image. The method is mapping Ig to Im.
- Define Ic as the converse of Im, which is constructed from the masked image.
- This paper final goal is to take sample from p(Ic|Im) to recover images.

Training Loss

$$\mathcal{L} = \alpha_{\mathrm{KL}}(\mathcal{L}^{r}_{\mathrm{KL}} + \mathcal{L}^{g}_{\mathrm{KL}}) + \alpha_{\mathrm{app}}(\mathcal{L}^{r}_{\mathrm{app}} + \mathcal{L}^{g}_{\mathrm{app}}) + \alpha_{\mathrm{ad}}(\mathcal{L}^{r}_{\mathrm{ad}} + \mathcal{L}^{g}_{\mathrm{ad}})$$

References

1. Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial networks. *arXiv preprint arXiv:1406.2661.*
2. Hui, Z., Li, J., Wang, X., & Gao, X. (2020). Image fine-grained inpainting. *arXiv preprint arXiv:2002.02609.*
3. Xu, R., Guo, M., Wang, J., Li, X., Zhou, B., & Loy, C. C. (2020). Texture Memory-Augmented Deep Patch-Based Image Inpainting. *arXiv preprint arXiv:2009.13240.*
4. Zheng, C., Cham, T. J., & Cai, J. (2019). Pluralistic image completion. *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 1438-1447).*
5. Flickr-Faces-HQ Dataset (FFHQ) dataset (https://github.com/NVlabs/ffhq-dataset)
6. City Space dataset (https://www.cityscapes-dataset.com/)