The Architectural Illustrator

David Heffren[1]*, Yizhou Tan[2]*
[1]Department of Computer Science and Department of Applied Mathematics, Brown University
[2]Department of Architecture, Rhode Island School of Design

CSCI-1430
Computer Vision
Fall 2022
Instructor: Srinath Sridhar
Project TA: Zachary Zhu

Motivation

Emerging technologies are increasingly affecting a lot of fields, but it has a relatively low impact on the design field right now because of the creative nature of design. We want to come up with tools that use technology and facilitate the design process.

Define the Problem

We want to generate realistic renderings of architectural designs from a sketch of them. This helps architects quickly visualize their ideas and improve design efficiency.

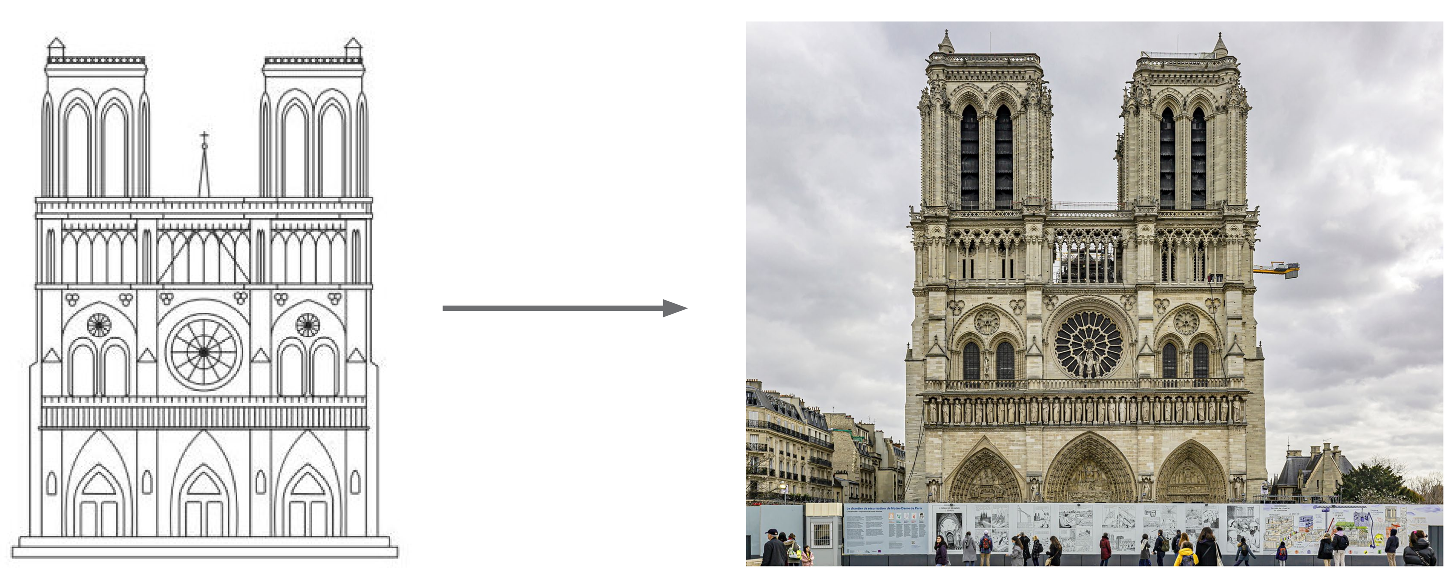

Figure 1. Conceptual Problem

Goal

We want to approach this problem with a conditional GAN. Since this model already exists as Pix2Pix, we also want to test out different combinations of model setups and compare the performances. We use edge detection on a dataset of architecture images to approximate the sketches of these images. The edges are then the conditions, or the input, to the GAN. The output is expected to mimic the original image.

Figure 2. Example Data and Preprocessing

Methodology

Our Generator U-Net I has the structure:
Encoder: C64 > CB128 > CB256 > CB512 > CB512 > CB512 > CB512 > CB512
Decoder: CBD512 > CBD1024 > CBD1024 > CB1024> CB1024 > CB512 > CB256 > CB128
Our Generator U-Net II has the structure:
Encoder: C64 > CB128 > CB256 > CB512 > CB512 > CB512 > CB512 > CB512
Decoder: CBD512 > CBD512> CBD512 > CB512 > CB512 > CB256 > CB128 > CB128

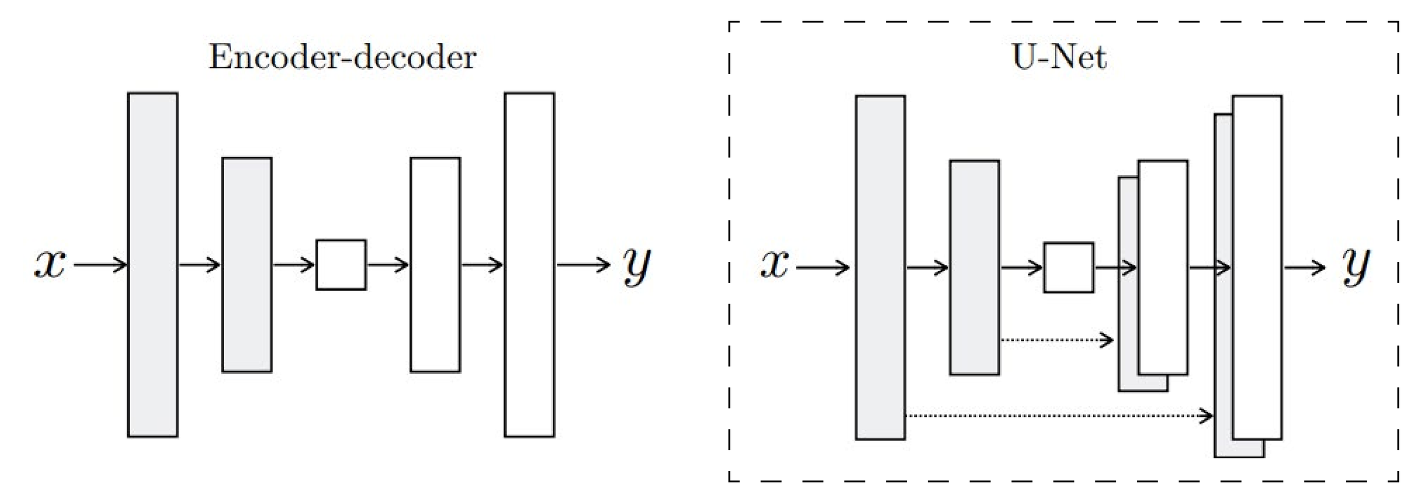

Figure 3-4. Difference between U-Net and normal Encoder-Decoder structure [2]

Results

Figure 5. U-Net I

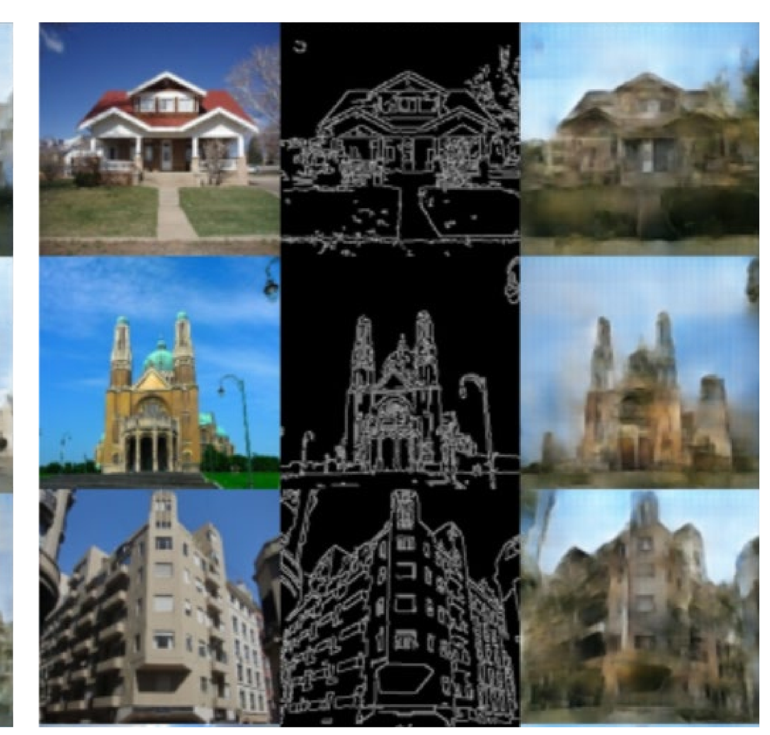

Figure 6. U-Net II

Figure 7. U-Net II without L1 Loss

Figure 8. U-Net II without PatchGAN

Performances Comparison

We plotted the losses for the four models that we tested.

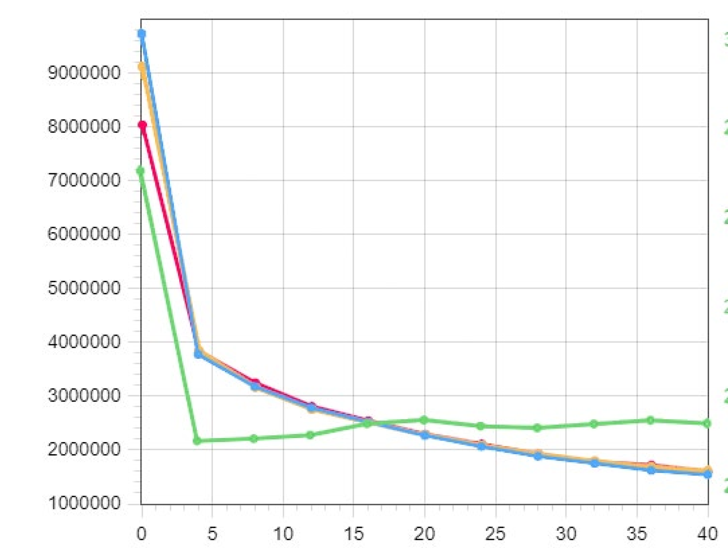

Figure 9. Train Sum Loss

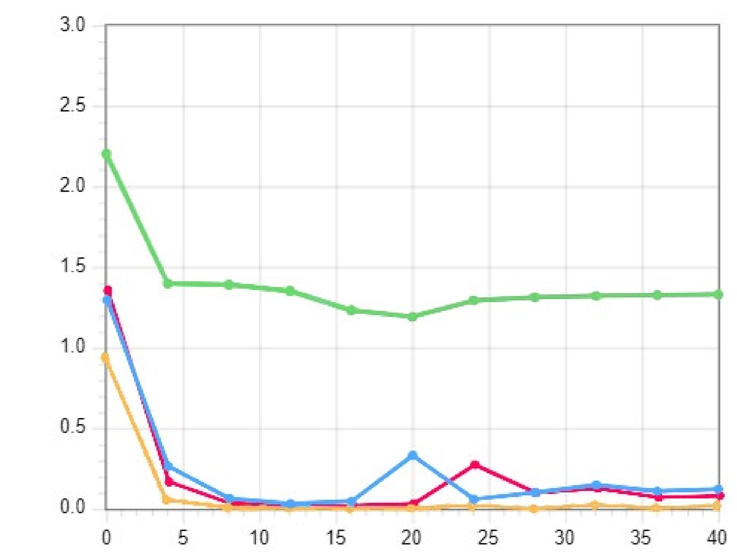

Figure 10. Train Discriminator Loss

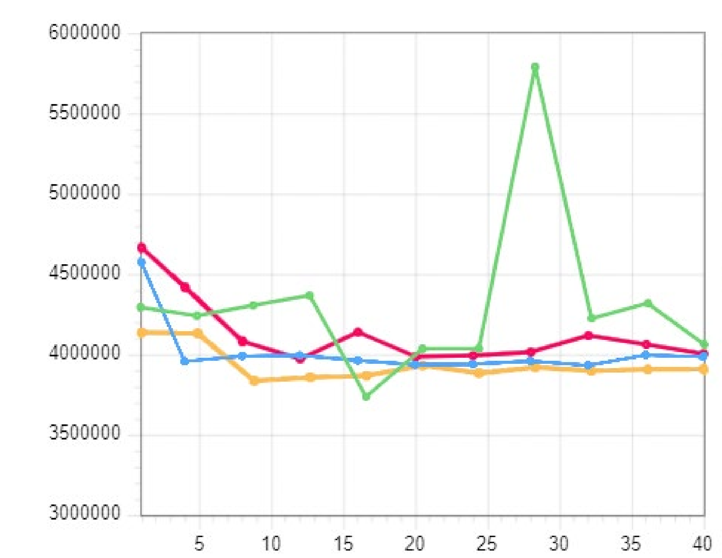

Figure 11. Validation Sum Loss

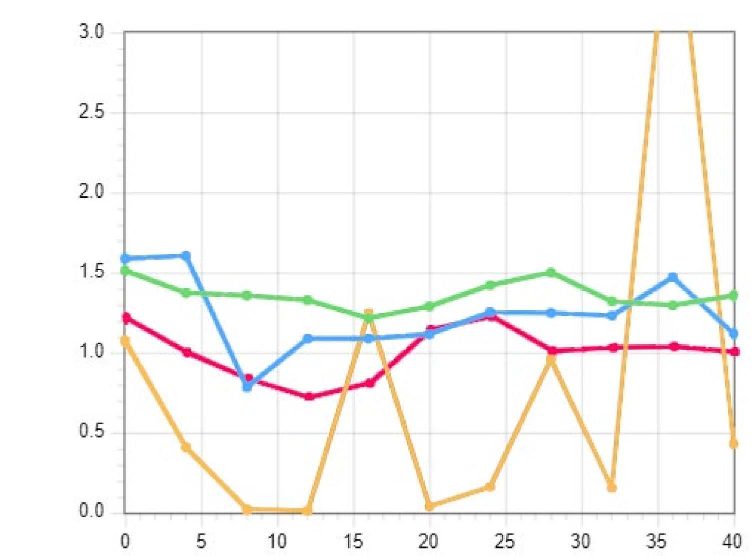

Figure 12. Validation Discriminator Loss

U-Net I
U-Net II
U-Net II without L1 Loss
U-Net II without PatchGAN

Acknowledgements

- Professor Srinath Sridhar for this course
- Our TA, Zachary Zhu, for providing help throughout the way

References

[1] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks, 2014.
[2] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. Image-to-image translation with conditional adversarial networks, 2016.
[3] Mehdi Mirza and Simon Osindero. Conditional generative adversarial nets, 2014.
[4] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation, 2015.
[5] Zhe Xu, Dacheng Tao, Ya Zhang, Junjie Wu, and Ah Chung Tsoi. Architectural style classification using multinomial latent logistic regression. In David Fleet, Tomas Pajdla, Bernt Schiele, and Tinne Tuytelaars, editors, Computer Vision – ECCV 2014, pages 600–615, Cham, 2014. Springer International Publishing.