# Project Proposal Notes

**Cedrick Argueta**
Department of Computer Science
Stanford University
Stanford, CA 94305
cedrick@cs.stanford.edu

**Kevin Wang**
Department of Physics
Stanford University
Stanford, CA 94305
kwang98@stanford.edu

## Abstract

## 1    What is the problem that you will be investigating? Why is it interesting?

We are investigating unpaired image-to-image translation with generative adversarial networks. This work was demonstrated in [4] and [2]. The pix2pix paper demonstrates how we can use paired images in training to create an image-to-image mapping from one domain to another. With CycleGAN, we can shift to using unpaired images for each domain.

## 2    What reading will you examine to provide context and background?

CycleGAN
pix2pix
SketchyGAN
SPADE
Google Drawing Guessing Game

## 3    What method or algorithm are you proposing? If there are existing implementations, will you use them and how? How do you plan to improve or modify such implementations? You don't have to have an exact answer at this point, but you should have a general sense of how you will approach the problem you are working on.

We will use CycleGAN to generate the image-to-image translations. We use the sketch as input, and then the GAN produces an image that it believes is in the target domain. The discriminator judges this output, deciding whether the generated image was real or fake. When a threshold has been reached, we have a GAN capable of producing fake images that belong to the target domain that corresponds to the input image, because of the cyclic property of CycleGANs.

Simple CycleGAN

## 4    What data will you use? If you are collecting new data, how will you do it?

We will combine labeled data from TU-Berlin sketch dataset [1] and ImageNet [3] to create our domains.

## 5   How will you evaluate your results? Qualitatively, what kind of results do you expect (e.g. plots or figures)? Quantitatively, what kind of analysis will you use to evaluate and/or compare your results (e.g. what performance metrics or statistical tests)?

Qualitative: whether the picture we draw can be translated to a realistic image.
Quantitative: Inception Score, Fréchet Inception Distance

## References

[1] Mathias Eitz, James Hays, and Marc Alexa. How do humans sketch objects? *ACM Trans. Graph. (Proc. SIGGRAPH)*, 31(4):44:1–44:10, 2012.

[2] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. Image-to-image translation with conditional adversarial networks. *CoRR*, abs/1611.07004, 2016.

[3] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael S. Bernstein, Alexander C. Berg, and Fei-Fei Li. Imagenet large scale visual recognition challenge. *CoRR*, abs/1409.0575, 2014.

[4] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A. Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. *CoRR*, abs/1703.10593, 2017.