Forest-to-Coast Image Translation with Cycle Consistency

{ Stylianos Rousoglou, Lydia Tsiverioti }   Stanford University CS 230

Overview

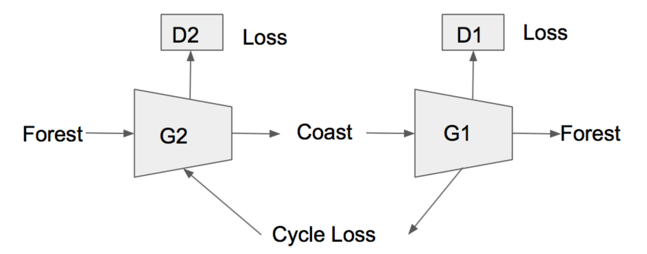
We develop models for unpaired translation from images of forests to images of coasts, and vice versa, using conditional generative adversarial networks (cGAN). We study the implications and performance of different generator and discriminator architectures, and evaluate the results in terms of aesthetic value and faith to the original, which we measure through cycle consistency loss.

Methods

We employ a 90/10 train/test split in our data. We use images of coasts and forests from MIT's LabelMe dataset. A cGAN generator is tasked to produce images the discriminator will categorize as "real." Conversely, the discriminator aims at detecting the generator's "fakes." The GAN objective is:

$$L_{GAN}(G, D_Y) = E_y[log D_Y(y)] + E_x[log(1 - D_Y(G(x)))]$$

We use two cGAN and an updated loss function as pictures Figure 1 to maintain cycle consistency between original and generated image.



$$L_{cycle} = -\frac{1}{m_{gen}}\sum_{i}^{m_{gen}} ||G_2(G_1(x_{c2}^{(i)})) - x_{c2}^{(i)}|| - \frac{1}{m_{gen}}\sum_{i}^{m_{gen}} ||G_1(G_2(x_{c1}^{(i)})) - x_{c1}^{(i)}||$$

$$L_{overall} = L_{GAN}(G_1, D_1) + L_{GAN}(G_2, D_2) + \lambda L_{cycle}$$

Figure 1: CycleGAN architecture

We explore generators U-Net128/256 and ResNet-6/9, both pictured in Figure 2. Additionally we use PatchGAN with 3 layers and then 6 layers.

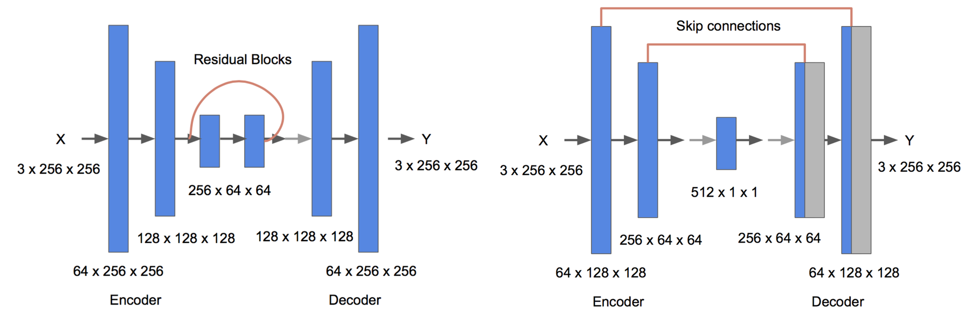


Figure 2: Left to right: ResNet architecture, U-Net architecture

Results/Discussion (I)

During epoch transition, as shown in Figure 3, we notice shapes, colors and borders becoming sharper and growing increasingly closer to resembling a coast while maintaining some of the original image's structure.

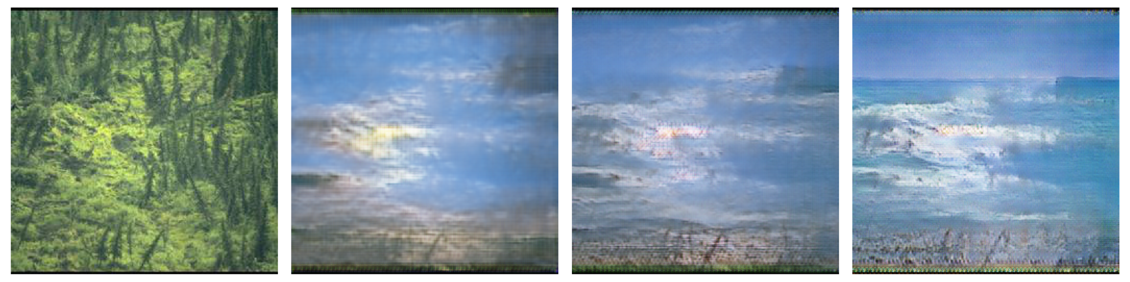


Figure 3: Left to right: original image, generated image at epoch: 50, 100, 200

Generator evaluation

Identity loss pushes the source-target to maintain the previous identity source distribution. In Figure 5, U-Net 128/256 decay more than ResNet-6/9. We observe that ResNet-6/9 images in Figure 4 exhibit "blue-er" colors that are more common in images of coasts than in those of forests. Conversely, U-Net 128/256 images have yellow/green undertones and remain more faithful to the original image of a forest. Our quantitative results validate our visual analysis.

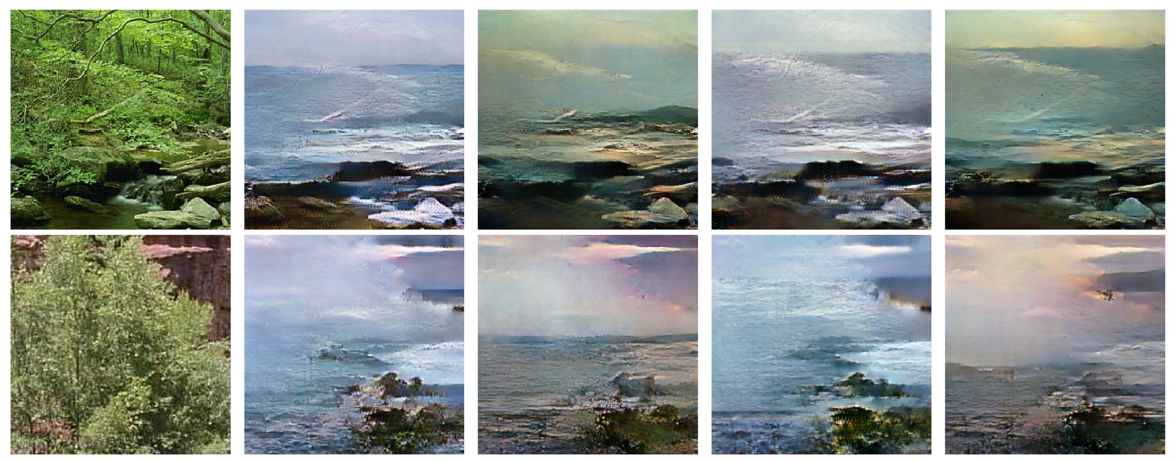


Figure 4: Left to right: original image, generated image using: ResNet9, U-net256, ResNet6, U-net128

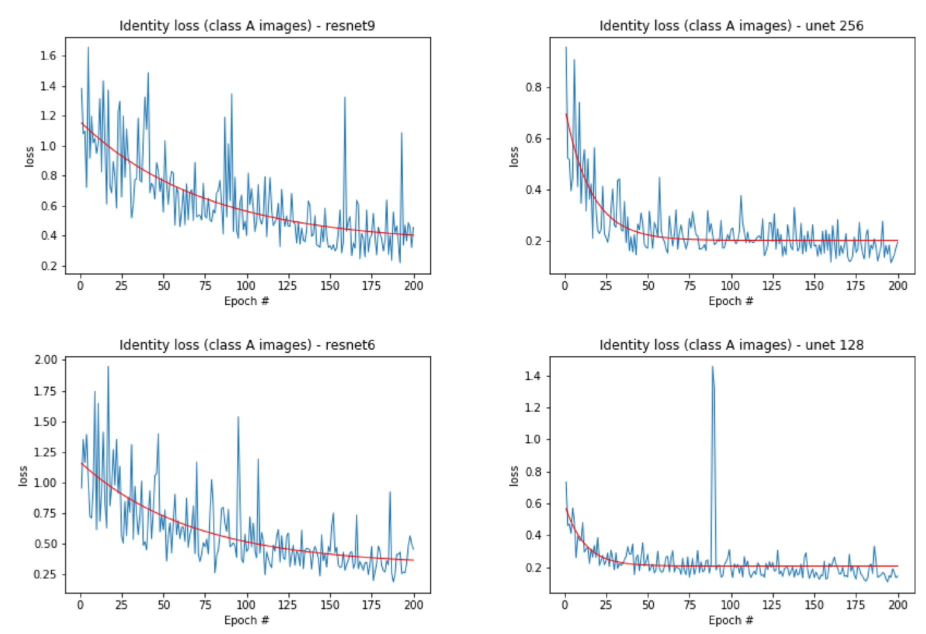


Figure 5: Plots of Identity loss vs epoch for each generator

Results/Discussion (II)

Cycle consistency evaluation

In cGAN we introduce a cycle consistency loss to enforce F(G(X)) = X. In Figure 6 U-net128 seems to be the least blurry and in general has similar vibrant colors to the original image. Conversely, ResNet-9 is the blurriest with the most "washed out" colors. Our Cycle Loss plots in Figure 7 validate our visual analysis. Even though ResNet6/9 produce more "realistic" images of coasts, they have lower Cycle consistency and Identity scores.Alternatively, U-net128/256 have better scores, but they produce less "realistic" images. There is no free lunch

Figure 6: Left to right: original image, recovered image using: ResNet9, U-net256, ResNet6, U-net128

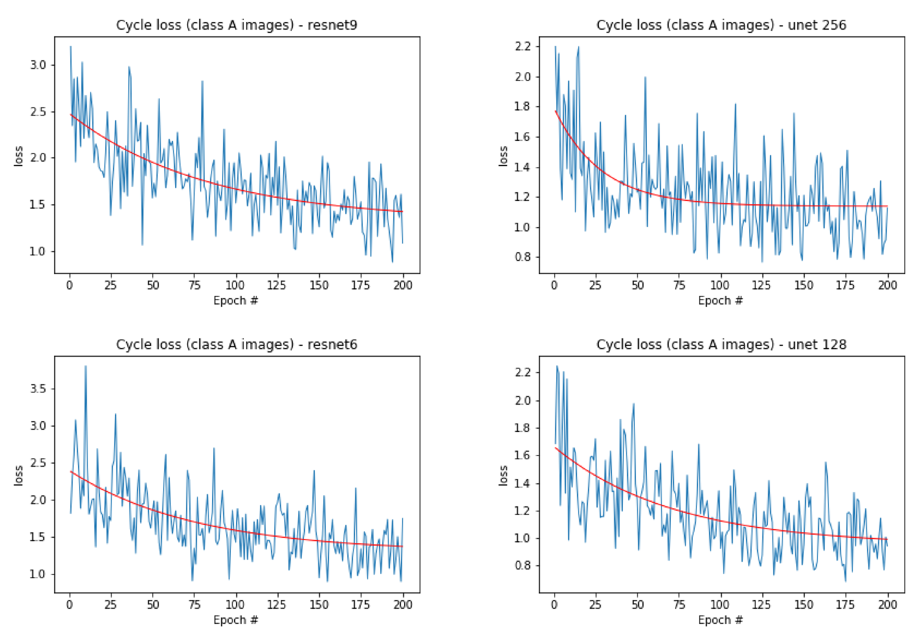


Figure 7: Left to right: Plots of cycle loss vs epoch number

Figure 8: original forest/coast - fake coast/forest - recovered forest/coast

Results/Discussion (III)

Discriminator evaluation

We examine discriminator complexity by comparing a 3 layer to a 6 layer PatchGAN. A 6 layer PatchGAN fails to produce a coast image as seen in Figure 9. The standard to trick the discriminator is much higher with increasing number of layers and therefore, the generated image has to look much more like forest in order succeed. Cyclic loss is predictably lower for a 6 layer discriminator as seen in Figure 10

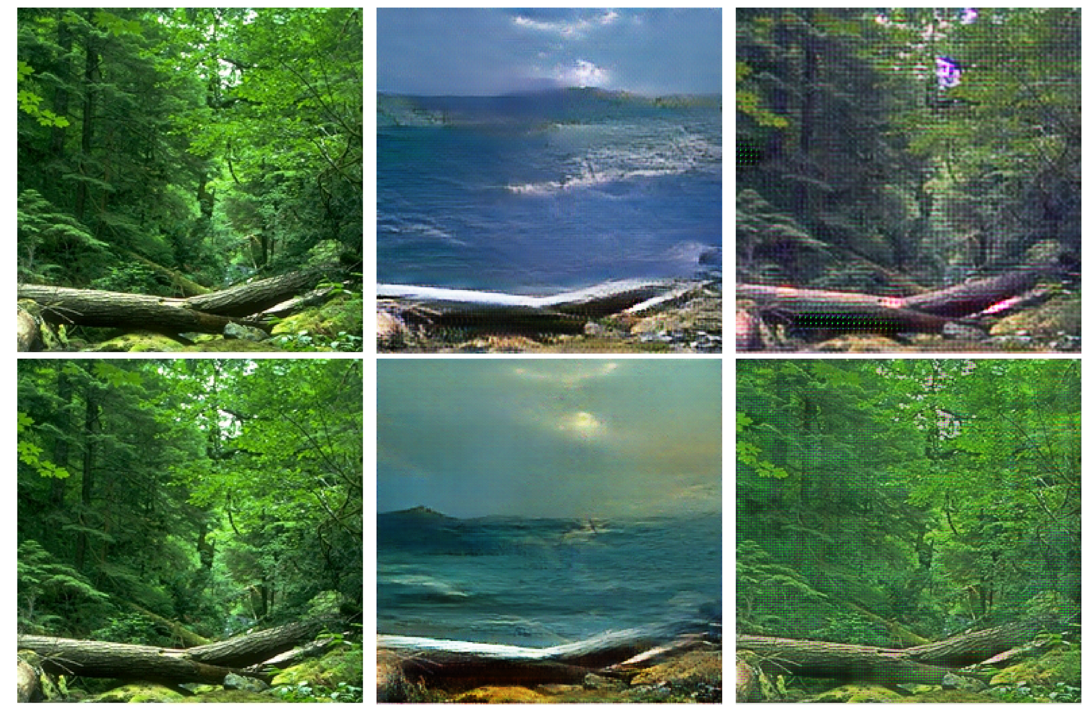


Figure 9: Top row: original image, ResNet-9 + PatchGAN with 3 layers, ResNet-9 + PatchGAN with 6 layers,Botton row: original image, U-net256 + PatchGAN with 3 layers, U-net256 + PatchGAN with 6 layers

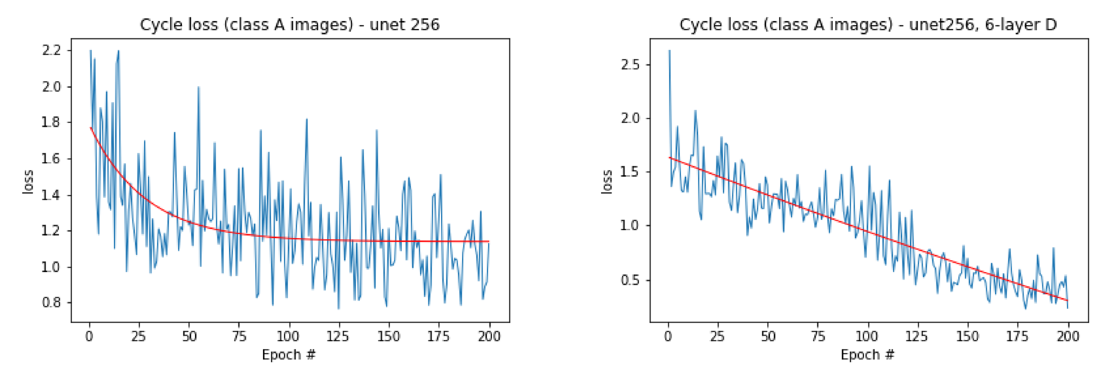


Figure 10: Cycle loss plots for U-Net256 + 3 layer PatchGAN (left), U-Net256 + 6 layer PatchGAN (right)

Conclusion

Even though the Identity and Cycle objective are crucial in our study we observe that putting too much weight on them produces less aesthetically pleasing images of coasts. We witnessed this when we added more complexity to our discriminator and or used a U-Net generator. Both succeeded in lowering these objectives but generatedx images inferior to the those of ResNet generators and shallower discriminators