



Segmentation of CycleGAN (SecleGAN)

---Improvement on CycleGAN

By: Jin Wu
Vince Chen
Jiayi Luo



Background (Original CycleGAN Algorithm)



The result of the CycleGAN Algorithm

- Image-to-Image transformation technique
- Useful in terms of transforming the input image to the desired output
- Very useful and can be broadly generalized in different domains

Background (limitation of the original algorithm)



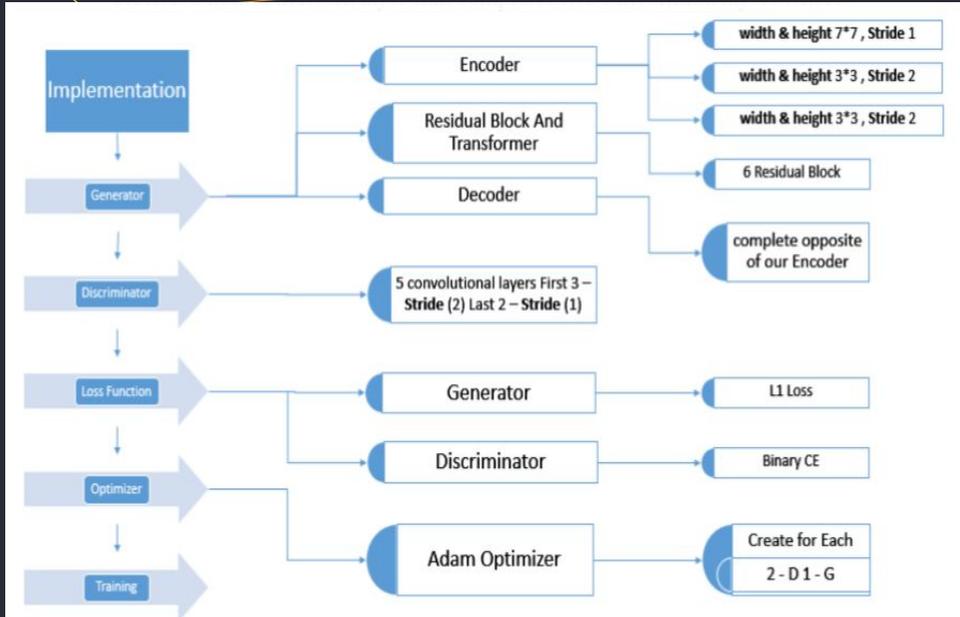
The result of the CycleGan Algorithm

The result from Original CycleGAN is not perfect in some cases.

In our project, we want to rectify this and make it right!

The Solution Is **SecleGAN!**

Literature Survey



Instead of the stochastic gradient descent algorithm that most CycleGan algorithms have been using, some researchers have used the Adam Optimizer and L1 loss function to optimize the CycleGan.

Result:

Even the Adam Optimizer and L1 loss function cannot resolve the above problem.

Literature Survey

Moreover, in the paper “CycleGan with better cycles” proposed by Wang. et al , it is mentioned that three changes were applied to improve the cycleGan model from the perspective of better cycle consistency:

1. Including a L1 loss on the CNN features by discriminator part to learn more features.
2. Gradual decay on the weight of cycle consistency loss λ as the training progresses.
3. Weight cycle consistency loss by the quality of generated images, which were obtained using decriminator's outputs.

$$\tilde{\mathcal{L}}_{\text{cyc}}(G, F, D_X, X, \gamma) = \mathbb{E}_{x \sim p_{\text{data}}(x)} \left[D_X(x) \left(\gamma \|f_{D_X}(F(G(x))) - f_{D_X}(x)\|_1 + (1 - \gamma) \|F(G(x)) - x\|_1 \right) \right]$$

New Loss function after changes

This method might work but we will propose a more straightforward improvement for the original CycleGan algorithm.



Why ML/DL approach?

Our method consists of two key parts: The CycleGan layer and the instance segmentation layer. There are many state-of-art methods for the instance segmentation. For example, photoshopping technique to take the instance out from the background or using deep learning to achieve this. And we choose the second option rather than the first one.

The reason is that deep learning method can be trained to achieve high accuracy yet remains high efficiency, in which the trained segmentation model can take many inputs at the same time and produce the desired results in very short time. However, if we were to choose the photoshopping technique, it would require a lot more labor and time. Overall, deep learning technique is a better choice for our project than other techniques.

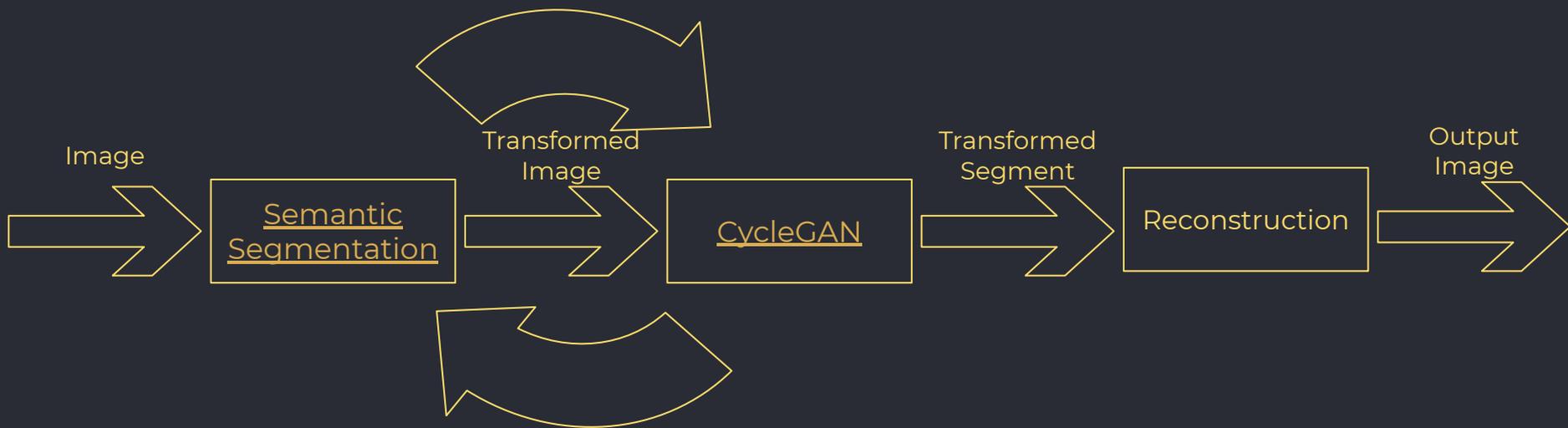


Dataset

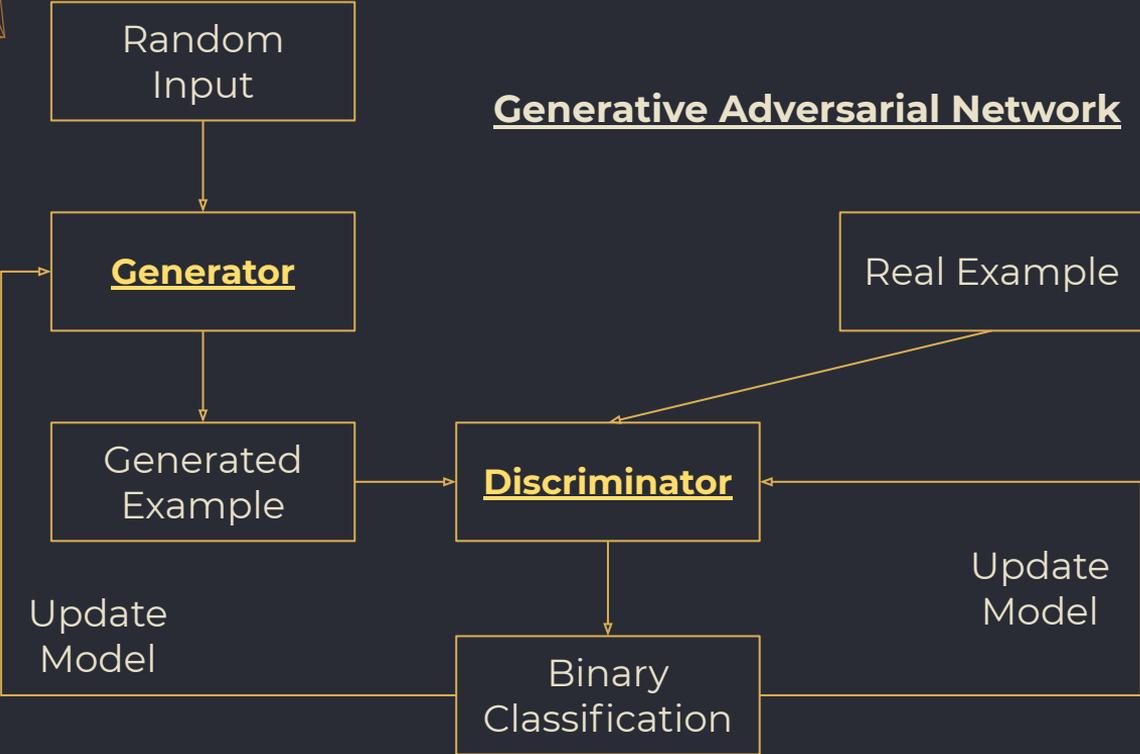


- The dataset we use in the project is Caltech Birds (2011) dataset which can be downloaded in Kaggle. (<https://www.kaggle.com/tarunkr/caltech-birds-2011-dataset>)
- This dataset contains 11788 images of birds in 200 different categories.
- In our project, we use two classes from this dataset:
056.Pine_Grosbeak and
043.Yellow_bellied_Flycatcher.

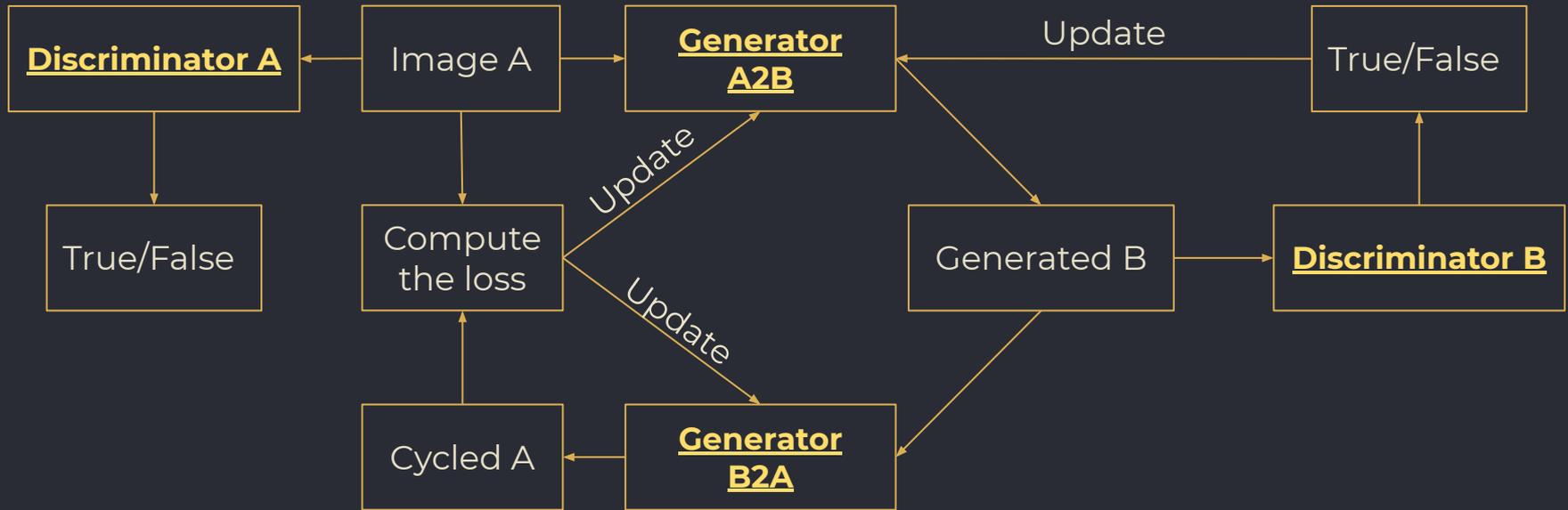
Overall Structure of SecleGAN



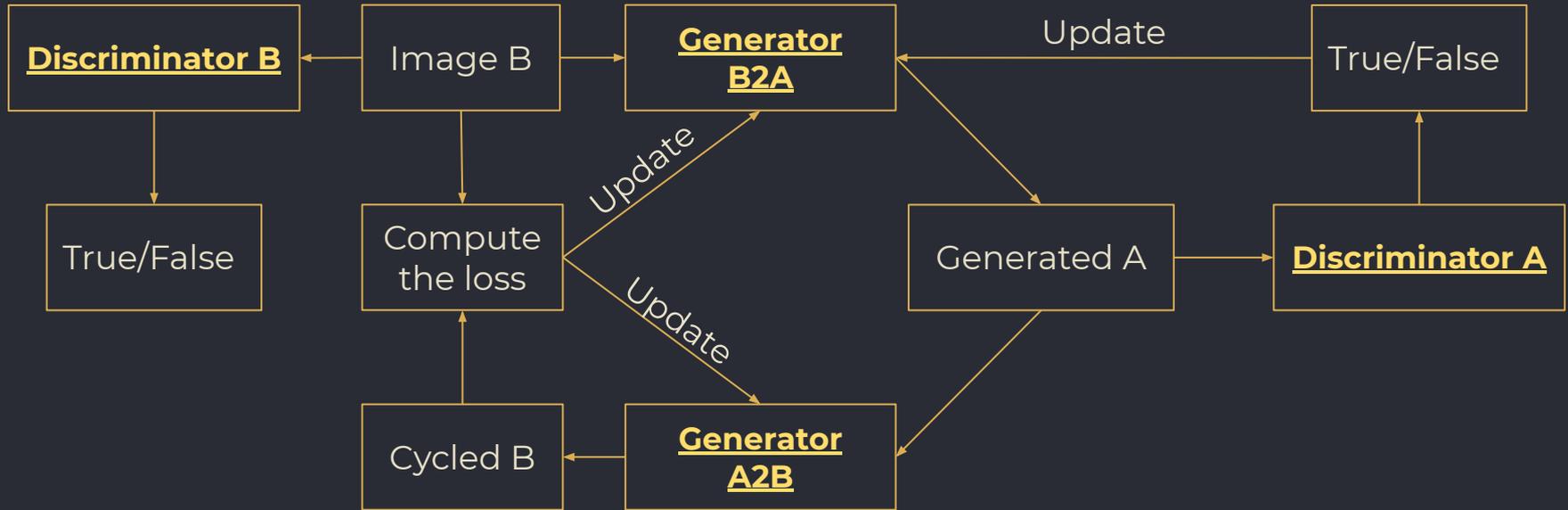
GAN Architecture



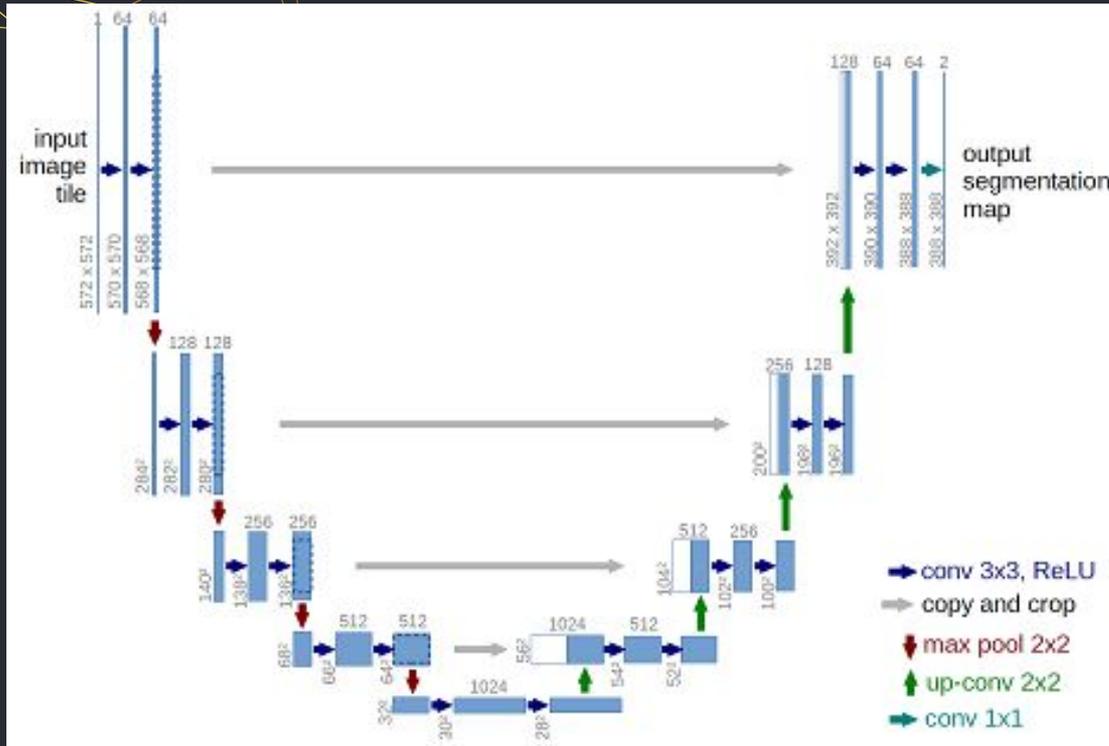
Models (CycleGAN-forward)



Models (CycleGAN-backward)



Models (Unet)



- Regular CNN cannot achieve the output of semantic segmentation since the output is a reconstructed image rather than a single label.
- Not only the convolution layer is included in U-net, but also the up-sample layer which is similar to the process of reconstructing images.
- U-net is the most adequate neural network architecture for semantic segmentation.

Results-forward path



Original

CycleGAN

SecleGAN

- The left images are the results of our project using forward path.
- The first column includes the original images.
- The second column includes the reconstructed images from original CycleGAN.
- The third column includes the reconstructed images from our SecleGAN.

Results-backward path



Original

CycleGAN

SecleGAN

- The left images are the results of our project using backward path.
- The first column includes the original images.
- The second column includes the reconstructed images from original CycleGAN.
- The third column includes the reconstructed images from our SecleGAN.



Further Work

1. We did not save the training loss curve and the accuracy curve yet. We are hoping that we could attach the curves in our final report.
 2. Try other semantic segmentation algorithms other than U-net and compare the loss curve and the results to see if other techniques can result in better transformations than using the U-net.
- 

Reference

1. Jun-Yan Zhu, Tacsung Park, Phillip Isola, and Alexei A. Efros. *Unpaired image-to-image translation using cycle-consistent adversarial networks*. *CoRR*, abs/1703.10593, 2017.
2. Tongzhou Wang, and Yihan Lin. *CycleGAN with Better Cycles*. https://ssnl.github.io/better_cycles/report.pdf



THANK YOU

