Individual Final Report - Image Colorization with GAN

Ruijin Jia

Introduction

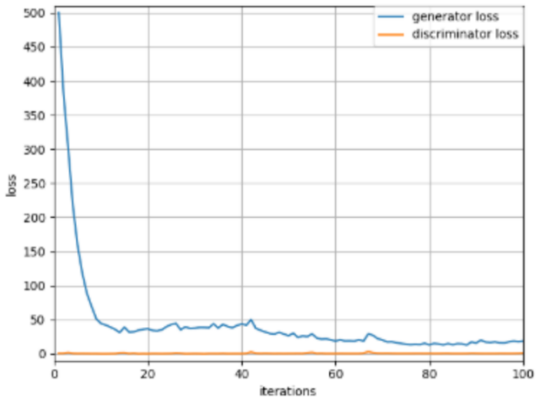
The project aims to help black and white pictures to get colors. The application of this technology includes restoring old images, help 2D animation get mutual colors in a short time. There are three members in total in our group, and shared work is the final presentation, final report, literature review and data preprocessing, determining which model to use.

Individual Work

GANs is efficient in understanding texture and symmetry in photos, and in our project, we used two methods in image colorization: GANs model based on L*A*B color space, and GANs models based on the image-to-image translation. My first part is conducting the model evaluation. For the first method, the Generator loss is the cross-entropy loss of generated images plus MSE of ab channels, and the Discrimination loss is the average of the loss of generated images and the loss of real images. For the second method, the Generator loss is the MSE loss of generated images and the loss for the pixelwise is the L1 loss. Therefore the total generator loss is the MSE loss plus the pixlwise loss time 100 (which is referred as lambda pixel). And the Discirmation loss is the weighted sum of the real loss and fake loss. These two losses can be calculated in the same way with GAN loss. The second part of the work is doing the loss function and visualize the results. Also, I trained the model, and recorded how the images got trained by time, which is helpful in further model improvement. Last but not least, I conducted the paperwork, including summarizing the literature work, and methodology summarization.

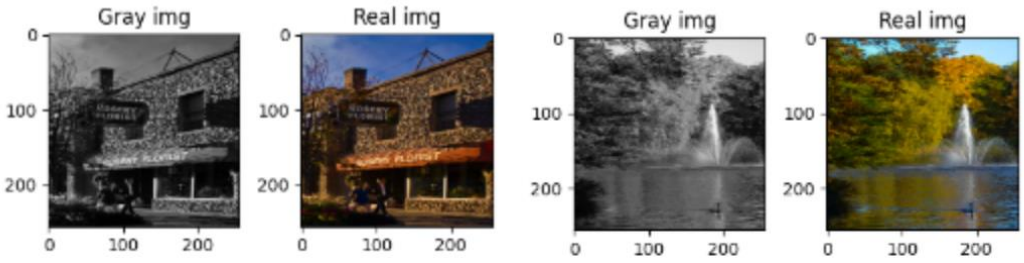
Results

The loss function of L*a*b* GAN is shown below. The training processing is stable. When training after 10 epochs, the loss goes down to below 50 and then keeps decreasing.

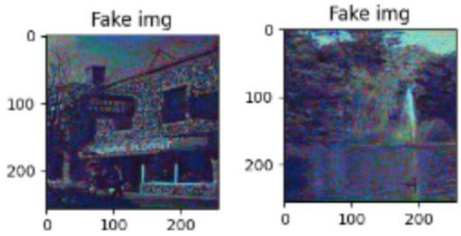


GAN based on L*a*b* color space:

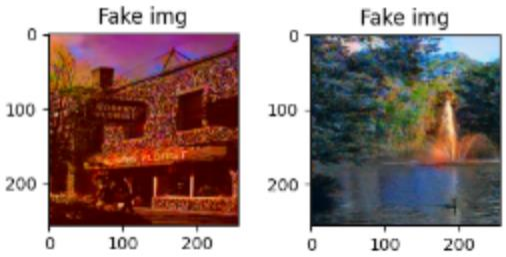
The original photos is shown below, the left is the photo with only L* channel (input of the generator) and the right are the original photo



The result of 1 epoch training is shown below. Model learn a little bit of a*b* channel but there is still a lot of noisy data inside the image.



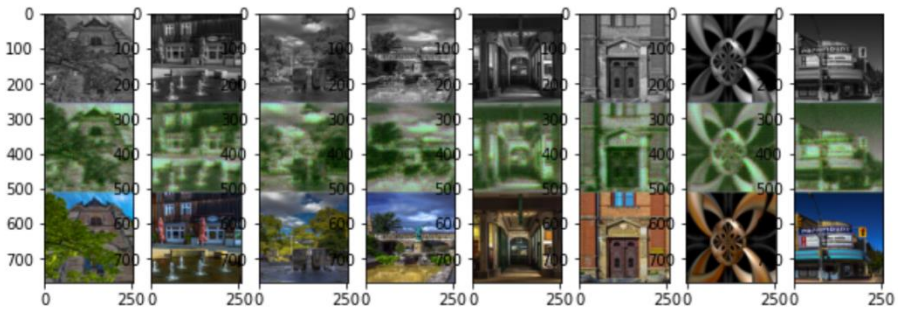
When it comes to 50 epochs, there is no more noisy data in images. But the color is still not as good as real images.



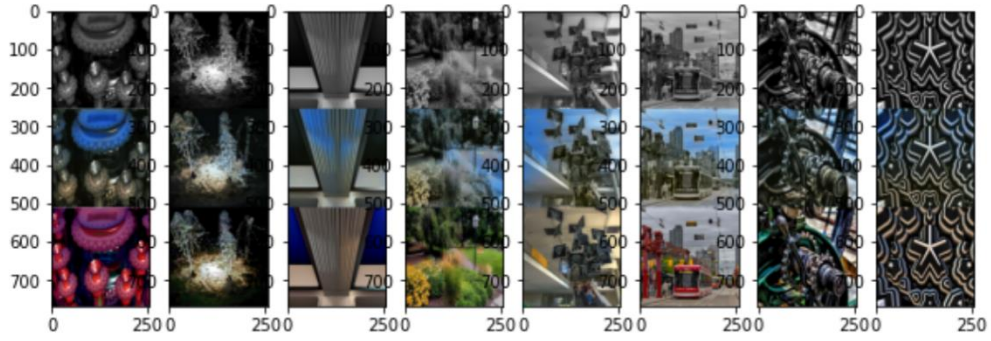
The result of 300 epochs training is shown below, images are very close to the original one.

GAN Image Translation:

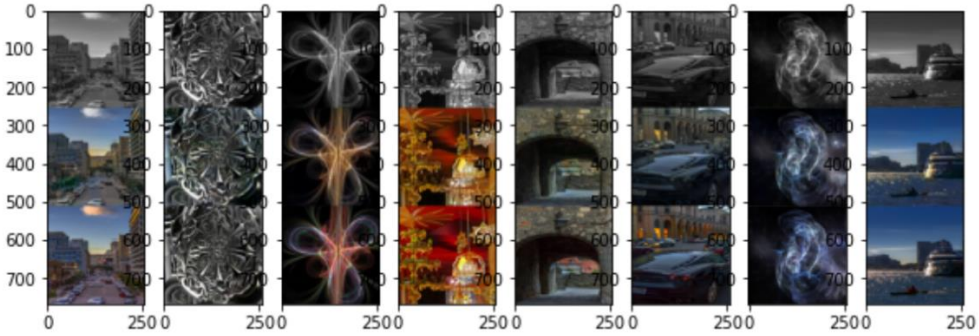
For each picture, the first row is the images we input, which are 3-channels grey images. The second one is images that the generator made. And the third row is real images. From the beginning, it is not hard to find that there is a lot of noise in those images while it can be seen as the overlap of noisy data plus gray images.



The results after training 5 times. Things became better as we can see some blue pixels can be generated. However, except for the sky, generators may color some areas blue which is not supposed to be blue.



The results after training 100 times. All fake images are good as the rarest color red can be colored well. And it is hard to distinguish the difference between real and fake images for the training set.



Summary and conclusions

Our evaluations are all based on the qualitative method, and from this scale, we can see how the model gets improved when we trained. To get more accurate results, we may introduce quantitative metrics methods. Moreover, we also got some wrong colors when images contain high levels of textured details. After analyzing the results, we think it is because we did not train enough person pictures so that model didn't learn enough information about humans.

Codes

The percentage of the code that you found or copied from the internet:

1-(60+168)/(130+45+280)=50.1%

References

Image colorization: https://arxiv.org/abs/1803.05400

Image to Image translation: https://openaccess.thecvf.com/content_cvpr_2017/papers/Isola_Image-To-Image_Translation_With_CVPR_2017_paper.pdf