CS512 - Gated-GAN: Adversarial Gated Networks for Multi-Collection Style Transfer

Team Members

- Brandon Bennitt: A20416733
- Kajol Tanesh Shah: A20496724

Main Paper Information

Xinyuan Chen, Chang Xu, Xiaokang Yang, Li Song, and Dacheng Tao, “Gated-GAN: Adversarial Gated Networks for Multi-Collection Style Transfer,” IEEE Transactions on Image Processing, DOI: 10.1109/TIP.2018.2869695

Link: https://arxiv.org/pdf/1904.02296.pdf

Problem statement

The problem we wish to solve is a variant of what is commonly known as style transfer. Style transfer describes the rendering of an image's semantic content as different artistic styles. Manually transferring the image style by a professional artist usually takes considerable time. Recently, with the emergence and power of generative adversarial networks (GANs), researchers have proved that they are extremely effective in style transfer as we are generating an image as an output naturally. Although general GANs can be used to solve this problem, there are issues related to ensuring that quality is guaranteed. In addition, general GAN generators are only compatible with one style i.e. the style they are trained on. To provide a series of possible style transfer options, different models and networks must be trained separately. In this project, we specifically wish to solve the problem of needing multiple models to generate images in multiple different styles. Therefore, we seek to build a single gated adversarial network that can perform this task.

Approach

In this approach, the proposed method uses gated adversarial networks (Gated-GAN) for the transfer of multiple artist or genre styles in a single network. Here, two sets of unpaired training samples are considered: one set of input images and a target set of collections of artists and genres. The Generator is trained to generate images in the style of a target artist/genre using adversarial loss and the Discriminator is trained to differentiate between the real and the generated images. Along with the conventional encoder-decoder architecture, it also considers additionally a gated-transformer network between the encoder and decoder consisting of multiple gates, each corresponding to one style so that the users can choose between different styles. The encoder and decoder are made up of a series of convolutional layers and the training is stabilized using auto-encoder reconstruction loss. To facilitate generating new styles, the gated transformer uses weighted connections between the transformer branches. The residual branch is used as the transfer module for each branch. Also, Label GAN is used as the model tends to get confused and mixes multiple styles if only the adversarial loss is used. Unfortunately, Label GAN also suffers from the overlaid-gradient problem so an auxiliary classifier is used to overcome this. The proposed method focuses on two components - first to differentiate between synthesized and genuine images, and second, to identify styles of the images. The model will be created in Python using Pytorch or TensorFlow.

Data

The dataset consists of a set of input images on which style transfer is to be performed and another set of images which is a collection of images of multiple artists/genre styles. The Generator will then generate images in the style of artists/genres. The dataset which we will be using contains 4 styles of different artists/genres which are Cezanne, Monet, Ukiyoe, and Vangogh containing 1000 images each. The set of test data also has 1000 images which will be used as input. In order to improve the performance of the model and reduce overfitting, data augmentation techniques can be used. Using these techniques, more new data can be generated from the original dataset and can create a more robust model. Later on, we can also use data from the ImageNet database to test our model.

Team Member Responsibilities

	Tasks	Assigned
1	Data Preprocessing	Brandon
2	Data Augmentation	Kajol
3	Building the original network from the paper	Brandon
4	Training the original network and test with K-Fold Cross Validation	Kajol
5	Modifications to try and increase accuracy	Kajol and Brandon
6	Prepare the report and presentation	Kajol and Brandon

References

- Controllable Artistic Text Style Transfer via Shape-Matching GAN: https://openaccess.thecvf.com/content_ICCV_2019/papers/Yang_Controllable_Artistic_Text_Style_Transfer_via_Shape-Matching_GAN_ICCV_2019_paper.pdf
- DRB-GAN: A Dynamic ResBlock Generative Adversarial Network for Artistic Style Transfer: https://arxiv.org/abs/2108.07379
- Encoding in Style: a StyleGAN Encoder for Image-to-Image Translation: https://openaccess.thecvf.com/content/CVPR2021/papers/Richardson_Encoding_in_Style_A_StyleGAN_Encoder_for_Image-to-Image_Translation_CVPR_2021_paper.pdf
- Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks: https://arxiv.org/abs/1703.10593
- Style Transfer Using Generative Adversarial Networks for Multi-site MRI Harmonization: https://link.springer.com/epdf/10.1007/978-3-030-87199-4_30?sharing_token=CVl0dCRLZm2qRehY46dXkPe4RwlQNchNByi7wbcMAY77-APbzlXwOT5RhkQVJUpAJ2M8Qw8VdiUcZKvGj628Ov6g9rrqbckOxwF73l1oKp2XmMpxWBKTnY7RHkACYRR8mUp0GCGvGqiwsCh4RupbRliDTbs3kVv_VTyxrPUj-a4%3D
- Deep Fakes" using Generative Adversarial Networks (GAN): http://noiselab.ucsd.edu/ECE228_2018/Reports/Report16.pdf
- Dataset: https://github.com/colemiller94/gatedgan/tree/master/photo2fourcollection image-net.org