General Idea:

Translating real images to **Realism** art style images. This art style dates back to the 17th century and there is no image-to-image pair data available to train traditional networks, So I am implementing ideas expressed in the paper **Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks** (https://arxiv.org/pdf/1703.10593v7.pdf) to implement a GAN to transfer real Image to realism art style image, a bi-product of this implementation will also be capable to convert Realism art style images to somewhat real photo.

Project proposal:

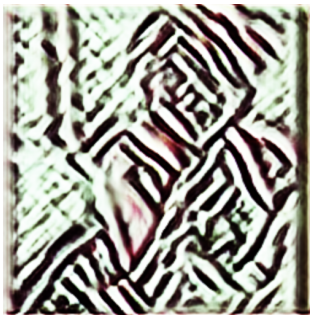
Usually, art style transfer is done with a neural style transfer on an existing model, but this has a drawback where we would require a style image and a real image. This type of transfer has drawbacks where human intervention to check if the style is suitable for a given image example given below.

Examples generated from https://reiinakano.com/arbitrary-image-stylization-tfjs/

(A portrait)

(Style)

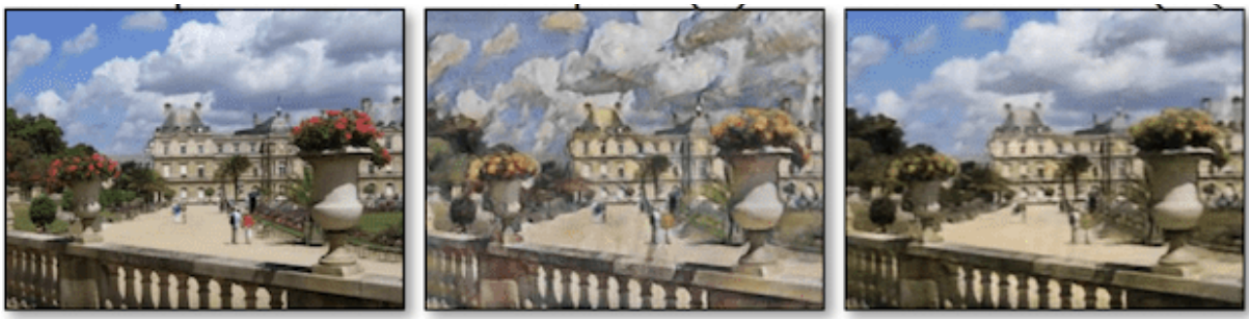
(Resulting image)



Above we can see that it is crucial to have a human intervention in such a network, else we can end up with images which don't transfer style or don't retain features from real images.
One other idea to do such translation is using Real Image - Art Image pair to train the network. The major drawback is this kind of dataset is not available for many cases/styles (for example Ukiyo_e - A genre of Japanese art which flourished from the 17th through 19th centuries, and Old Realism arts - the accurate, detailed, unembellished depiction of nature or of contemporary life.)

So I am using ideas in the paper - **Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks** (https://arxiv.org/pdf/1703.10593v7.pdf) to create a model to generate Realism art style models.

Example from the trained model taken from https://machinelearningmastery.com/impressive-applications-of-generative-adversarial-networks/



(Real) Output (Generated g(x)) (Reconstructed f(g(x))

Dataset:

I'm using Realism image set from https://github.com/cs-chan/ArtGAN/tree/master/WikiArt%20Dataset
It contains 10,734 images belonging to this art style and I'll split this into 70:20:10 ratio for train:test:val respectively.

Examples from dataset:

Defining baseline and evaluation metrics:

The paper Cycle gan paper **Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks**(https://arxiv.org/pdf/1703.10593v7.pdf) uses traditional methods of human turkers labeling and comparison of the loss for Per-pixel acc, Per-class acc with other models. There are few issues with such metrics, Per-pixel acc is an evaluation on f(g(x)) on which we are not interested as our primary output is g(x) and we have to evaluate this, Per-class acc depends on the prediction of the class of the object on a trained object however we can't be sure that all the object in the art/ real image is being trained with this model.

Because of the above-mentioned issues, I'll be deviating from baseline and evaluation metrics from the base paper and instead will be using a likeness score as defined in the paper **A Novel Measure to Evaluate Generative Adversarial Networks Based on Direct Analysis of Generated Images** (https://arxiv.org/pdf/2002.12345.pdf) which defines a likeness score as a combination of Creativity, Inheritance, and Diversity. I'll be using the implementation in https://github.com/ShuyueG/GAN_evaluation_LS and use the score generated here.
LS is closer to 1 if the GAN performs better.

To define the baseline, I would require a real image used to test the model, and output of this. This real image test set will be a mixture of landscapes, portraits, and indoor images. Therefore I won't be able to define a baseline on a common feature, and instead, I'll be taking a random image from my dataset and using it to define the baseline score LS. (This gives the bound on creativity: Generated images should not be the same as real, Diversity: Generated images are not similar to each other, and Inheritance: Overlap of creativity and diversity)
The same set of 200 images will be used for baseline definition and evaluation of the model. These images will be picked from
(landscape) https://www.kaggle.com/arnaud58/landscape-pictures,
(indoor) https://www.kaggle.com/itsahmad/indoor-scenes-cvpr-2019,
(portrait) https://github.com/NVlabs/ffhq-dataset

LS score for this set is: 6.529e-05
Dataset used:
https://drive.google.com/drive/folders/19S3l5pbouar6db8NDmAsnsyH4CQCdU-u?usp=sharing
Results and code:
https://colab.research.google.com/drive/15mYNIkiycrvI884d3D7Xon12eTpYKt51?usp=sharing
After training, our model should have an LS score more than what was defined in the baseline. As LS score approaches 1, it indicates our GAN performs well.