Motivation: Producing fully colored drawings from sketches is a large, costly bottleneck in the anime industry.

Problem: The project aims to automate line art coloring with deep learning. The inputs are sketch drawings of anime characters. The outputs are high-quality colored images.

Solution: Neural Style Transfer, C-GAN and CycleGAN.

Dataset: Anime Sketch Colorization Pair (Kaggle).

Input pipeline:
- The training example and the ground truth in the same image are separated.
- All images are resized to 256 x 256 resolution and normalized to the range [-1, 1].

Model definitions:
- Neural Style Transfer:
  - U-Net architecture (left), Generator (middle), Discriminator (right).
- CycleGAN:
  - A generator G generates a colorized image from a sketch image, while a discriminator D takes as input a sketch image and a color image and determines whether the color image is real or fake.
  - Architecture: The generator is a U-Net. The discriminator is a PatchGAN.
  - Modeling: A generator G generates a colorized image from an input sketch image, while a discriminator D takes as input a sketch image and a color image and determines whether the color image is real or fake.
  - Architecture: The generator is a U-Net. The discriminator is a PatchGAN.

Experiment results:
- Visually, this model performs slightly better than the baseline, but far from perfect. The model focuses on two colors (black and brown), but doesn’t learn to produce different colors.
- C-GAN (Pix2Pix):
  - Experiments and Results: cont.
  - CycleGAN:
    - Results of CycleGAN. Ground truth (left), generated image (right).
  - C-GAN:
    - Results of C-GAN. epoch 10, ground truth (middle), generated image (right).

Conclusion and Future Work:
- Summary: Our models produce decent outcomes on this anime colorization task, with C-GAN yielding the best performance as it improved past its texture and color problems.

Next Steps:
- Fine-tuning hyperparameters.
- Designing a real-fake experiment to test models’ performance based on human perception.
- Using S2A X 512 resolution to generate high-quality output.
- Experimenting with alternative models (GANs and conditional VAEs) and modifying network architectures (ResNet, ImageGAN).
- Doing transfer learning with pre-trained weights for Pix2Pix.
- Utilizing different color spaces.
- Adding total variation loss to remove high frequency artifacts.
- Conditioning on certain colors to give the user more control.

References: