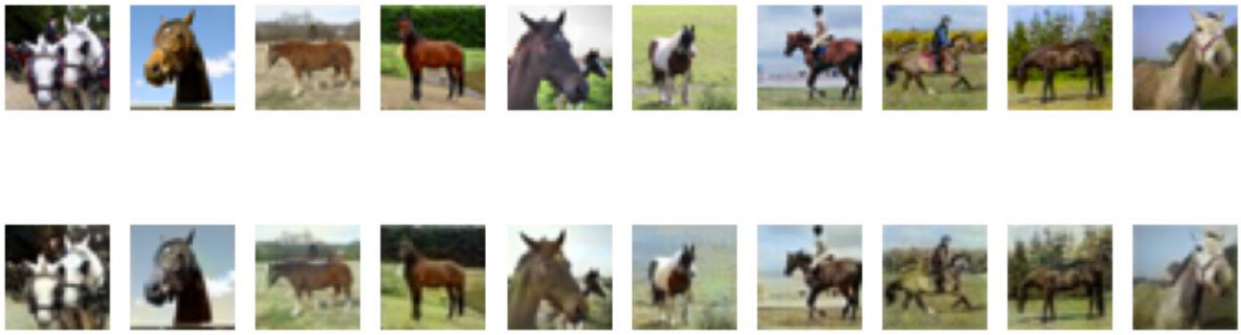
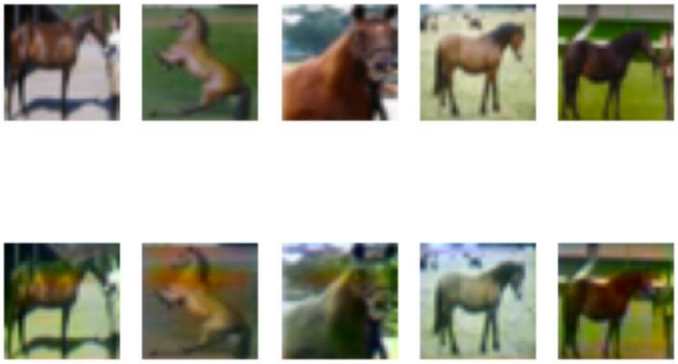
This project was based on the 6000 3*32*32 horse images in the CIFAR-10 dataset, in which 5000 were for training and 1000 were for testing. The inputs were converted to 1*32*32 grayscale and the labels were the original RGB 3*32*32 images.

The autoencoder structure had 6 layers, in which 2 were upscaling convolutional, followed by a constant size convolutional layer, then 2 downscaling convolutional layers, and finally a constant size convolutional layer. Except the output layer, all layers were relu activated. Skip connections were employed to keep the resolution of the output image. By applying the model to a random subset of the separated 1000 grayscale testing images, the coloring results are shown below, with the first row being labels and the second row being model predictions:



The GAN structure in RGB color space had the generator almost identical to the described autoencoder but with leaky relu activations instead of relu. Its discriminator had two convolutional layers and two linear layers with a sigmoid activation in the output layer to give binary result. The GAN structure in LAB color space had the generator almost identical to the RGB model but only 2 channels instead of 3 at the output, being the A (red-green) and B (blue-yellow) channels. This color space locked the lightness of the image in the L channel and represents the distances between colors more reasonably than RGB. The discriminator was unchanged. By applying the model to a random subset of the separated 1000 grayscale testing images, the coloring results are shown below, with the first row being labels and the second row being model predictions:



The GAN results gave higher saturations with some abbreviations, while the autoencoder results looked more realistic.