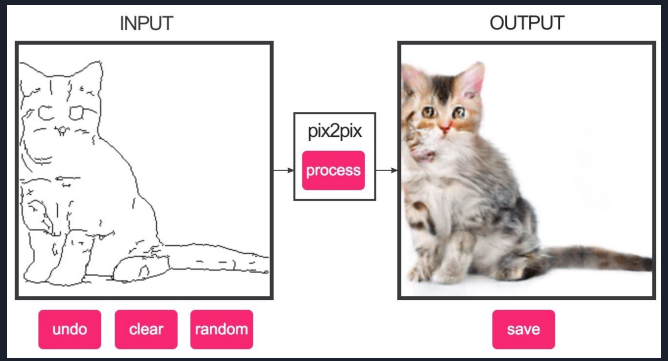
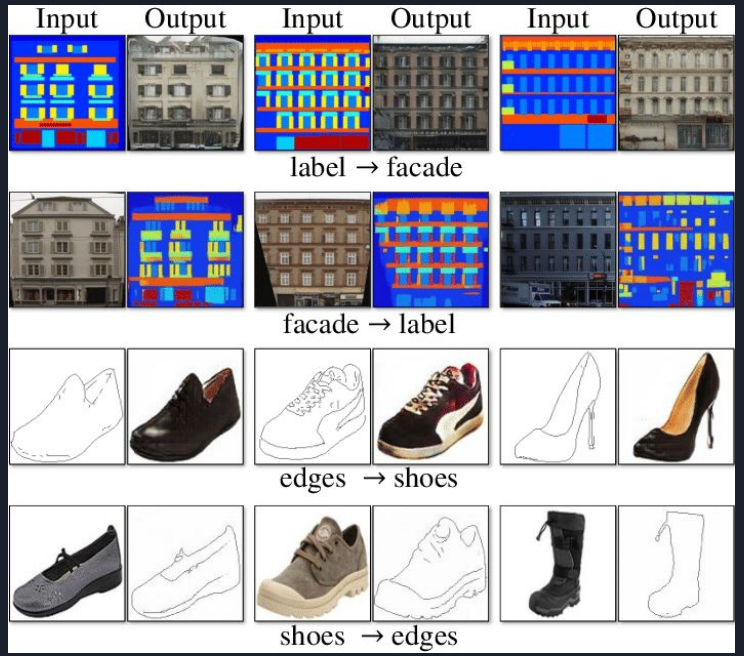
GAN Cars pix2pix

Shatrugna Rao Korukanti
U43517028

Pix2pix GAN

- Pix2Pix is a Generative Adversarial Network, or GAN, model designed for image-to-image translation.

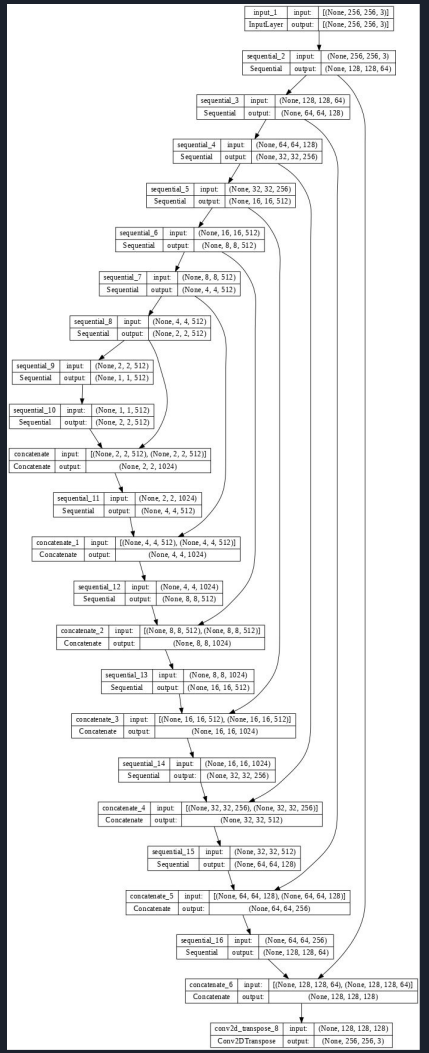




Pix2pix architecture

U-Net Generator Model
The generator model takes an image as input, and unlike a traditional GAN model, it does not take a point from the latent space as input.

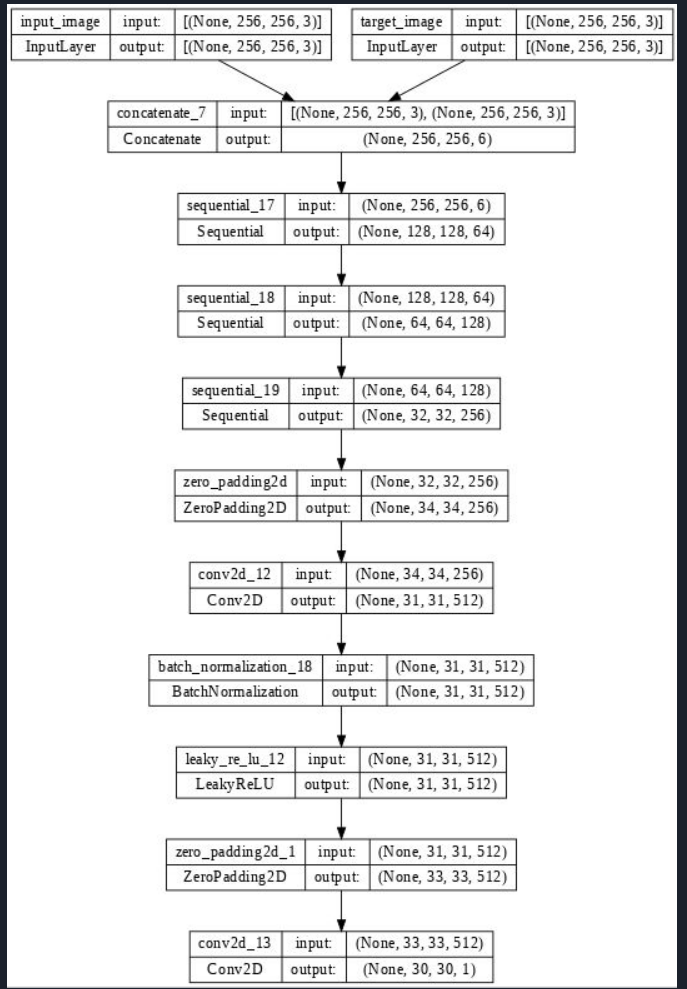
Instead, the source of randomness comes from the use of dropout layers that are used both during training and when a prediction is made.
The pix2pix paper mentions the L1 loss, which is a MAE (mean absolute error) between the generated image and the target image.



Pix2pix architecture

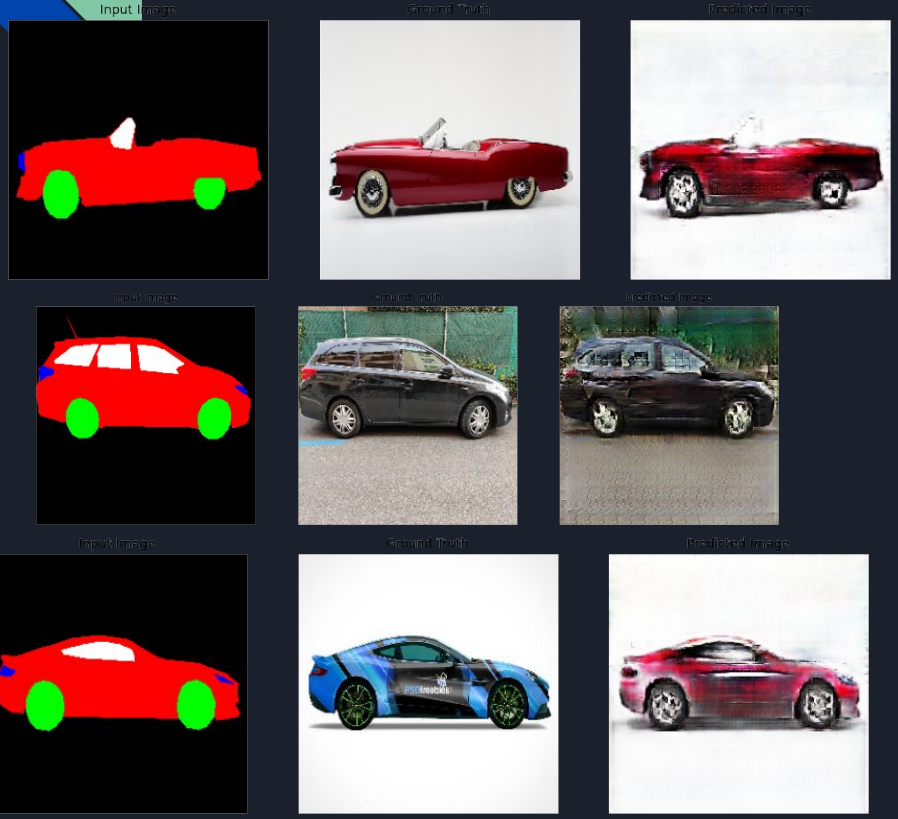
PatchGAN Discriminator Model
The discriminator model takes an image from the source domain and an image from the target domain and predicts the likelihood of whether the image from the target domain is a real or generated version of the source image.

- Input: Image from source domain, and Image from the target domain.
- Output: Probability that the image from the target domain is a real translation of the source image.



Results

The model generated images with considerable quality despite training it only for 1 hour in google colab.



Future Scope

- Should be able to train the model with variable image sizes. The model is constrained to use 256x256 pixel image.
- Beside Cars, should be able to train images with multiple more segmentation like road, tree, buildings, and more.