Image-to-Image Translation with Conditional Adversarial Networks

Mansour DIA

27 août 2021

Introduction

General Architecture

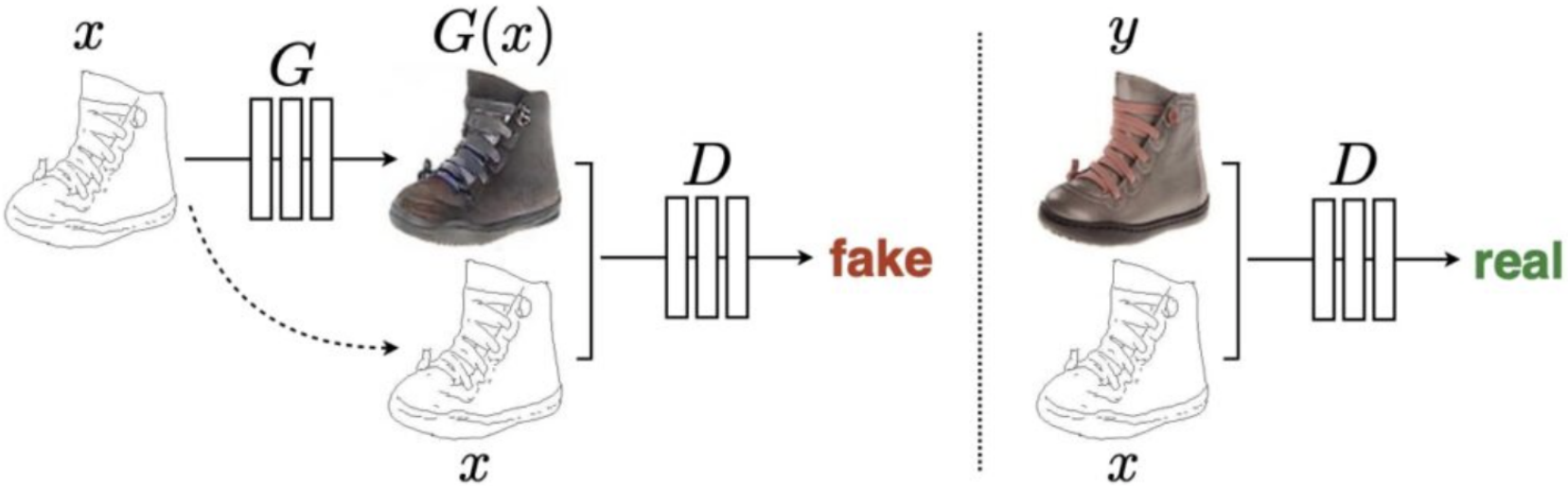

FIGURE – Pix2Pix Architecture.

PatchGAN Discriminator

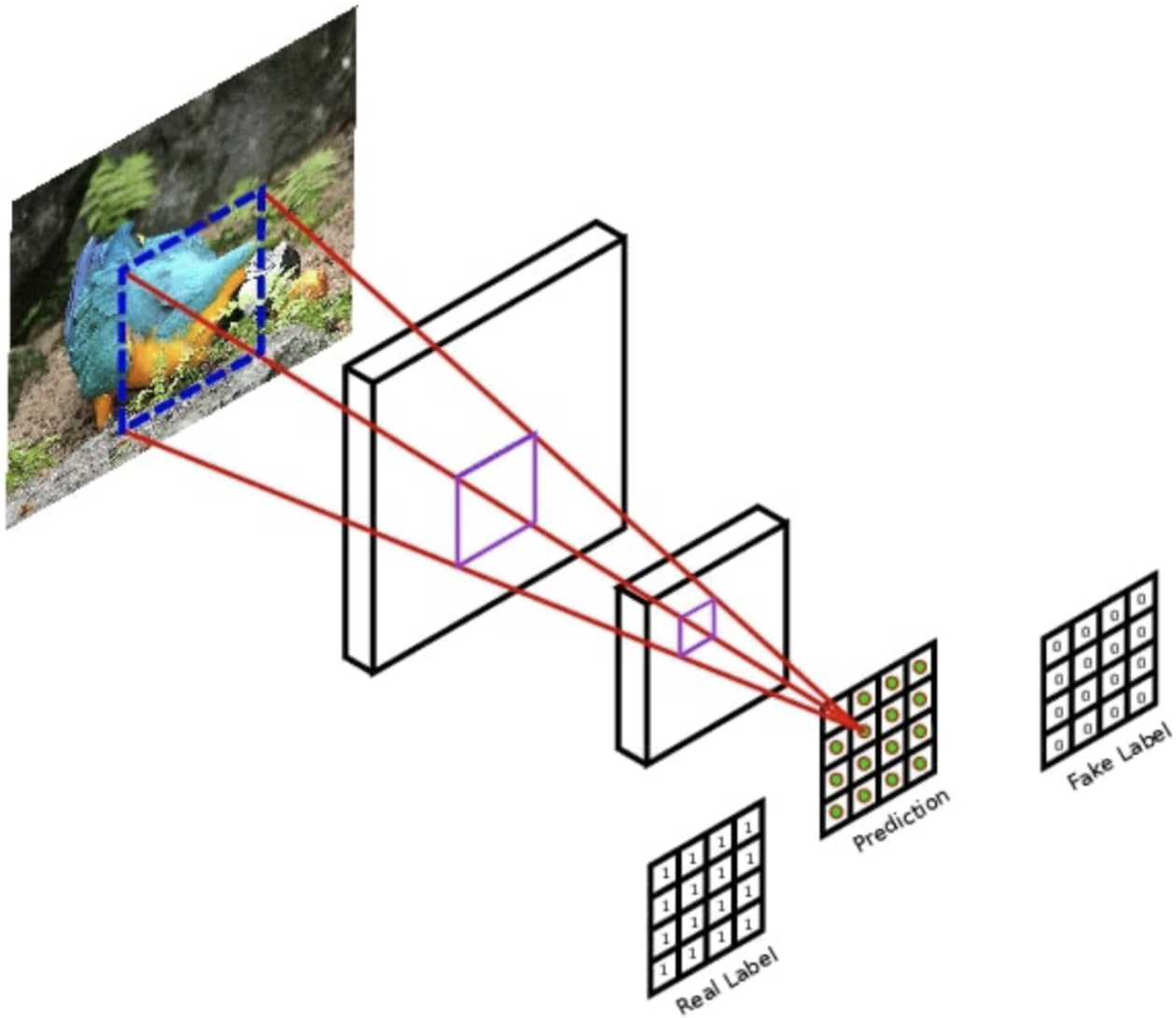

FIGURE – PatchGAN architecture.

Unet Generator

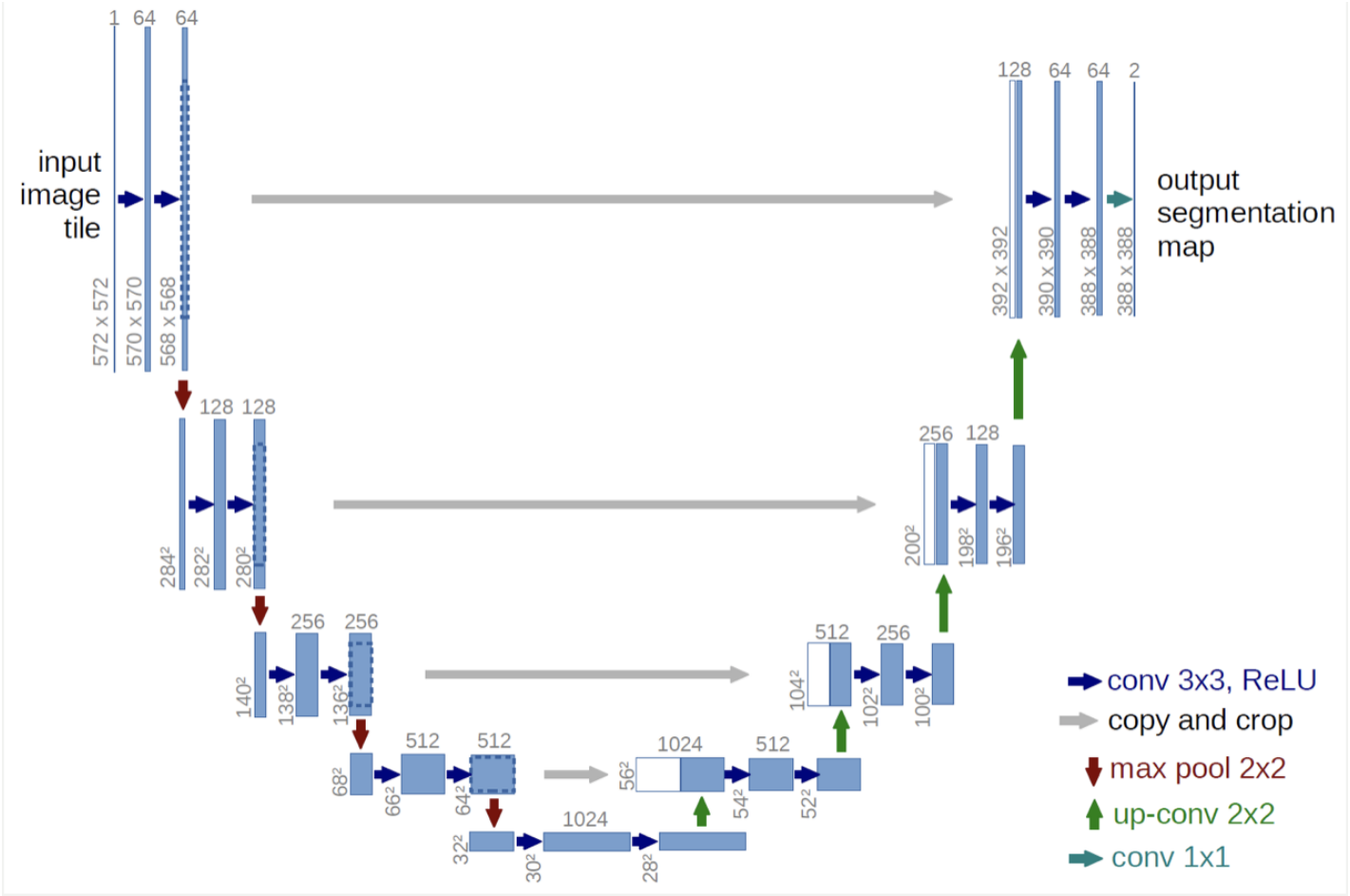

FIGURE – Unet architecture.

Pix2Pix loss function

$$\mathcal{L}_{cGAN}(G, D) = \mathbb{E}_{x,y}[\log D(x, y)] + \mathbb{E}_{x,z}[\log(1 - D(x, G(x, z))]$$

$$\mathcal{L}_{L1}(G) = \mathbb{E}_{x,y,z}[\|y - G(x, z)\|_1].$$

$$G^* = \arg \min_G \max_D \mathcal{L}_{cGAN}(G, D) + \lambda \mathcal{L}_{L1}(G).$$

patGAN hyperparameters

Unet hyperparameters

- block (Conv2D/ConvTranspose2d + BatchNorm + Activation) ,

Unet hyperparameters

- block (Conv2D/ConvTranspose2d + BatchNorm + Activation) ,
- stride =2 ,

Unet hyperparameters

- block (Conv2D/ConvTranspose2d + BatchNorm + Activation) ,
- stride =2 ,
- input_channels=3 and output_channels=3

Unet hyperparameters

- block (Conv2D/ConvTranspose2d + BatchNorm + Activation) ,
- stride =2 ,
- input_channels=3 and output_channels=3
- feature_map_dims = 64, 128, 256, 512

.

Training loop

Final hyperparameters setting

- $n\_epochs = 250$,
- $noise\_vect\_dim = 64$,
- $batch\_size = 128$
- $lr = 0.00001$

Results

FIGURE – loss of the generator.

Results

FIGURE – loss of the discriminator.

Results

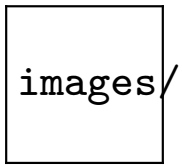

FIGURE – epoch 1.

Results

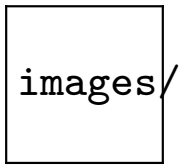

FIGURE – epoch 20.

Results

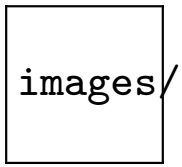

FIGURE – epoch 250.

Future work

- Convolutional layers insted of MLP ,

End

Thank you for your attention !