CMPM202 PROJECT TWO

Training Neural Network for Creative Tasks

Presented by Alex Mayben
and Jeffrey Weekley

PROJECT 2 IN 2 PARTS

- Generate Text: Walt Whitman
- Generate Media: CycleGAN for Puffins and Mallards

GENERATE TEXT

- Using the GPT-2 model on Google Colab
- Learning rate: 1000 epochs, temperature 0.7
- Walt Whitman was chosen because:
  - Large corpus of public-domain works
  - Easily gleaned from gutenberg.org
  - In Leaves of Grass alone, there are 129,066 words

ORIGINAL SAMPLE

Facing West from California's Shores

Facing west from California's shores,
Inquiring, tireless, seeking what is yet unfound,
I, a child, very old, over waves, towards the house of maternity,
the land of migrations, look afar,
Look off the shores of my Western sea, the circle almost circled;
For starting westward from Hindustan, from the vales of Kashmere,
From Asia, from the north, from the God, the sage, and the hero,
From the south, from the flowery peninsulas and the spice islands,
Long having wander'd since, round the earth having wander'd,
Now I face home again, very pleas'd and joyous,
(But where is what I started for so long ago?
And why is it yet unfound?)

From Whitman's Leaves of Grass

GENERATED SAMPLE

What is wrong

What is wrong with me? it is you, the earth,
It is I, it is every one, it is you yourself—
The minute you do not answer me the
mysterious glimmering illuminates,
The moment you do not see the gleam I give you,
The moment you are not yourself as I am, then I am.

NOT From Whitman's Leaves of Grass

GENERATED SAMPLE

Whitman is

*Whitman is the poet of sin, and
the poet of death;
What poet is there who does not write his country's poet?
Who but leaves his state's poet behind him that writes in his
state?*

*The name of the poet is hideous to me, it is rank to me,
Offhand to you it comes some day or at night,
Olle warm you tapestry of my life, I know not why,
But with you it comes every day, you fascinate me.*

*Why should I not be fascinated by the name of the poet?
The name is terrible to me, (it is in the Latent Fecund
on my soul,)
I do not want the effusion of the name that drips from me,
I want the effusion of the name that comes from me daily and
forever.*

GENERATED SAMPLE

How is it

How is it I extract strength from the beef I eat?
Is it the same beef my wife eats?
The same grass I eat? the same vegetables?
Is it the same hog-shoot I eat? the same mutton I eat?
The hay-cake I eat? the bread I eat?
The flax-husks I eat? the seeds I eat?
The beans I eat? the oil I eat?
The nitrogenous compounds I eat? the alkaline hydrolysis of
the roots of trees?

GENERATED SAMPLE

Whitman is

Whitman is the poet of sin, and
the poet of death;
What poet is there who does not write his country's poet?
Who but leaves his state's poet behind him that writes in his
state?

The name of the poet is hideous to me, it is rank to me,
Offhand to you it comes some day or at night,
Olle warm you tapestry of my life, I know not why,
But with you it comes every day, you fascinate me.

Why should I not be fascinated by the name of the poet?
The name is terrible to me, (it is in the Latent Fecund
on my soul,)
I do not want the effusion of the name that drips from me,
I want the effusion of the name that comes from me daily and
forever.

MYSTERY SAMPLE

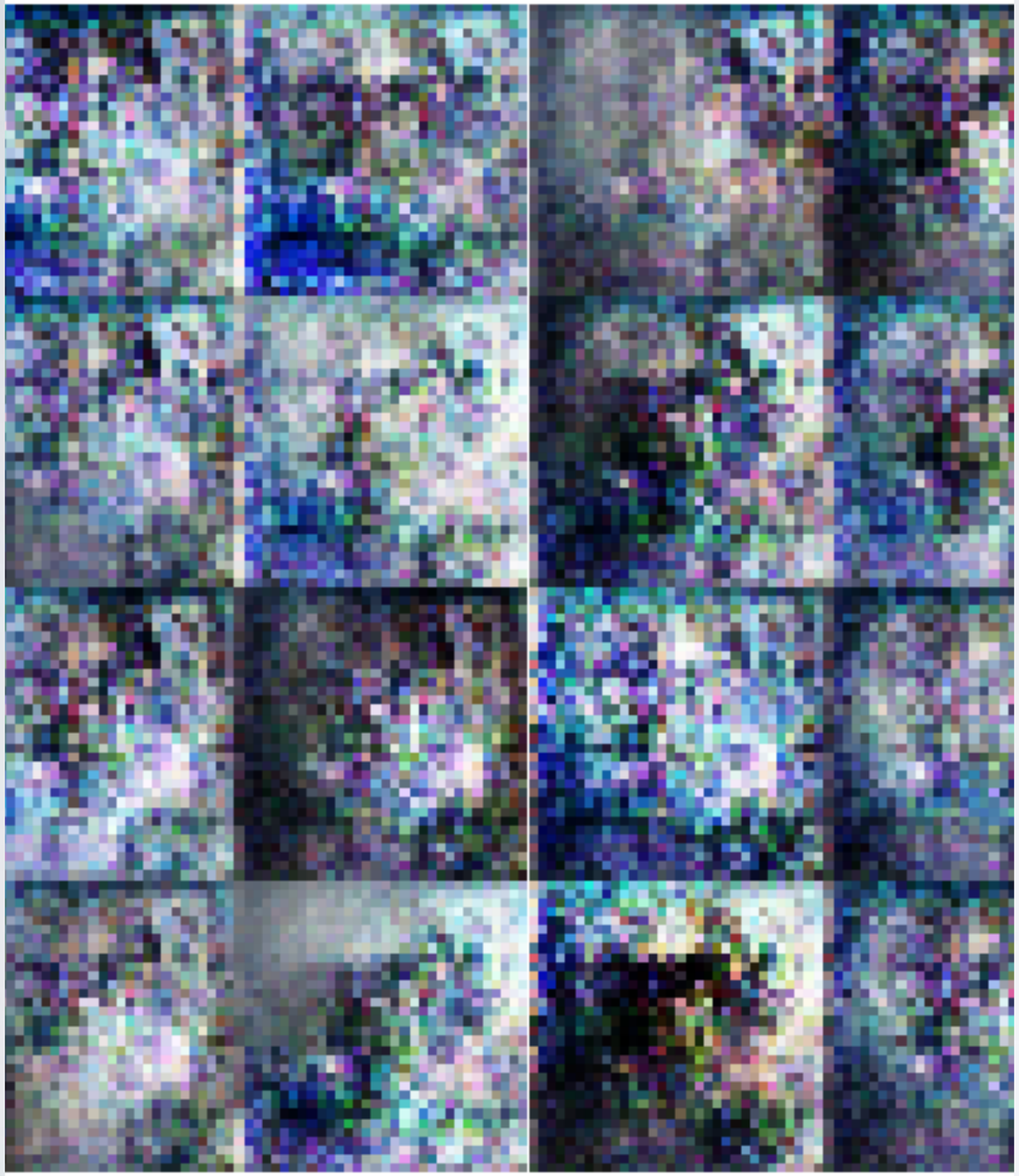
Turn the page

The song is over, the day is done,
I announce the close of the book—the true ending,
The great themes, the triumphs of the human race,
The many great characters—the varied, vast, separate races,
The unique, stern, remorseless, fire-breathing character,
This serene, burning soul of mine.

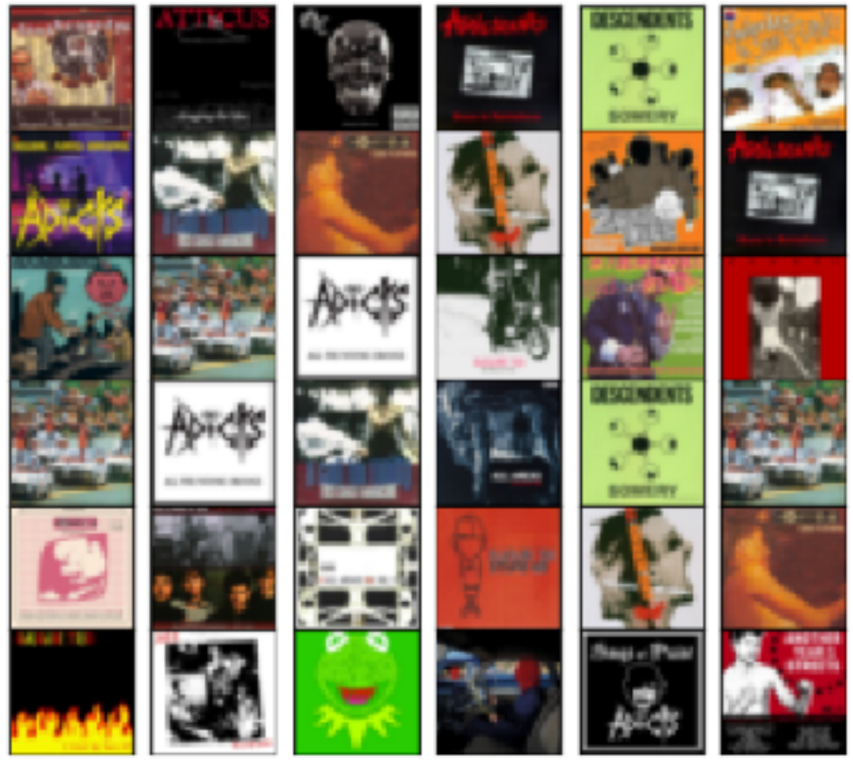
GENERATING IMAGES OF BIRDS

Pix2Pix, a CycleGAN for Birds

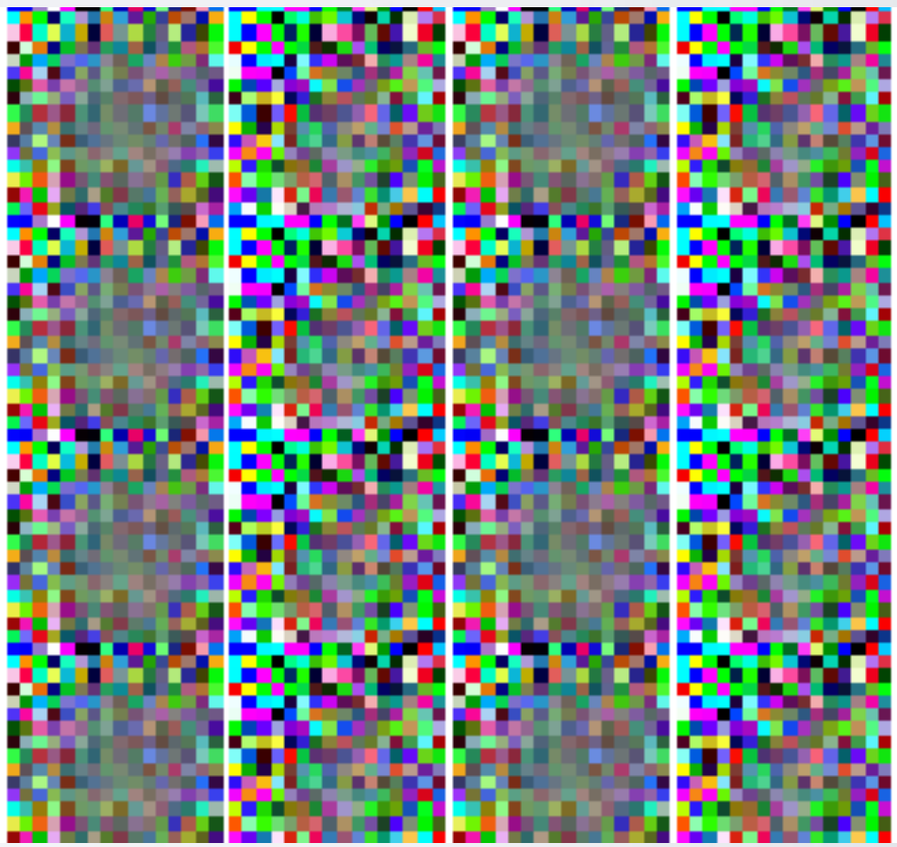
GAN Training against 400 Punk Album Covers

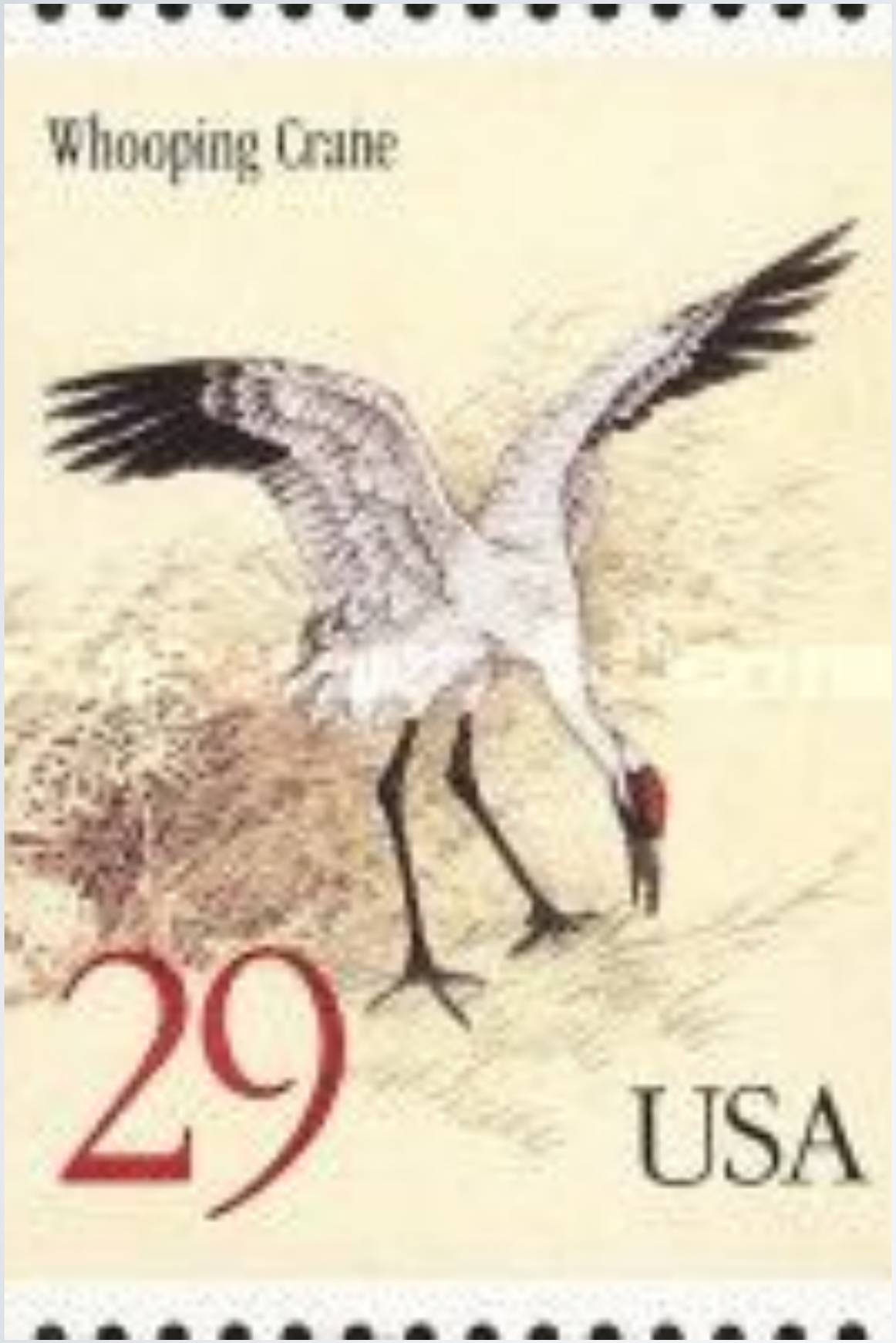


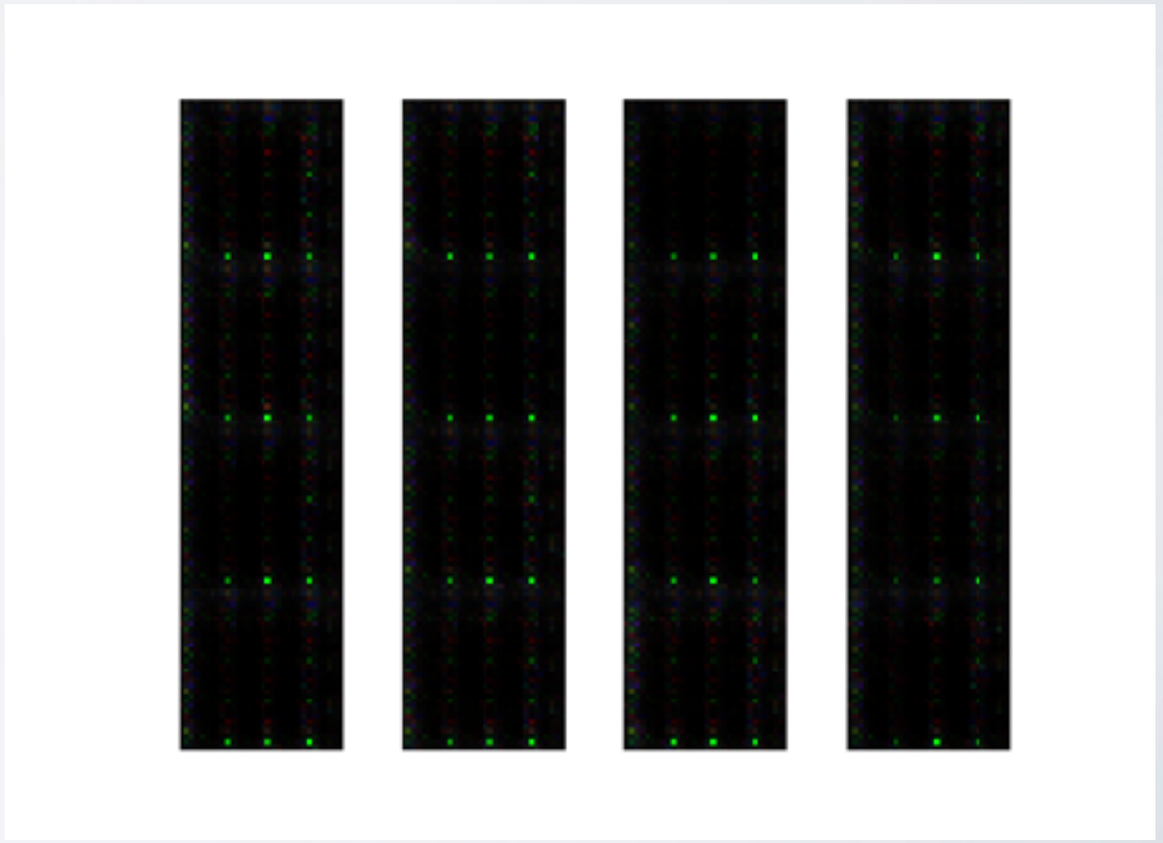
Input



Kernel




Whooping Crane
29
USA



GENERATE IMAGES

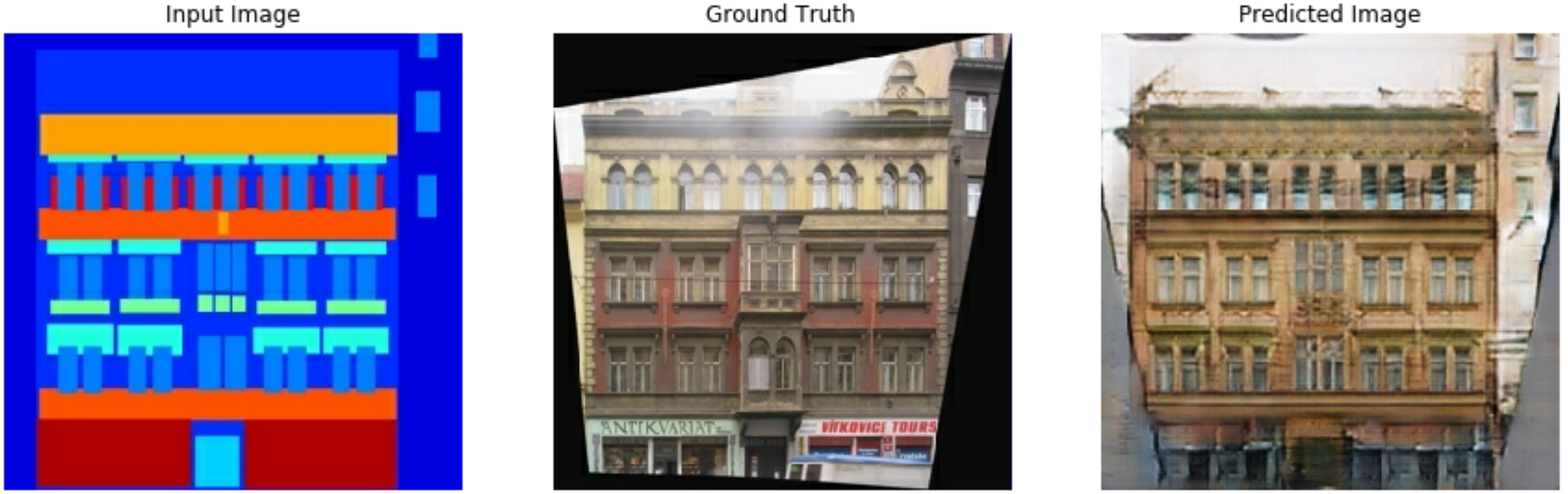
- Using the Pix2Pix on Google Colab
- CUB-200-2011 Dataset from CalTech

- Large corpus of tagged bird images
- 200 species, each with 60 images

PREPROCESSING

- Scaled to 512x512, then cropped each image to be exactly 256x256 (Presuming the area of interest was centered and this would preserve features)

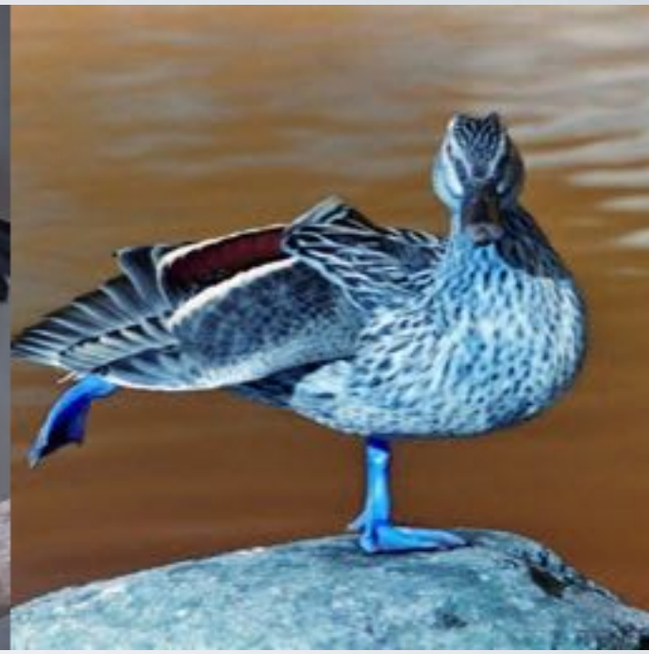
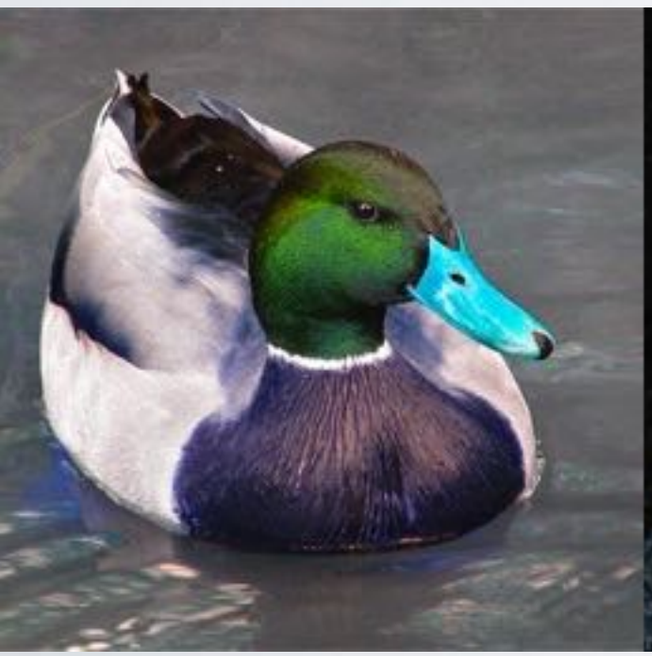
- Images need to be side-by-side in two folders (A and B) with same names and flip-flopped relative to each other.

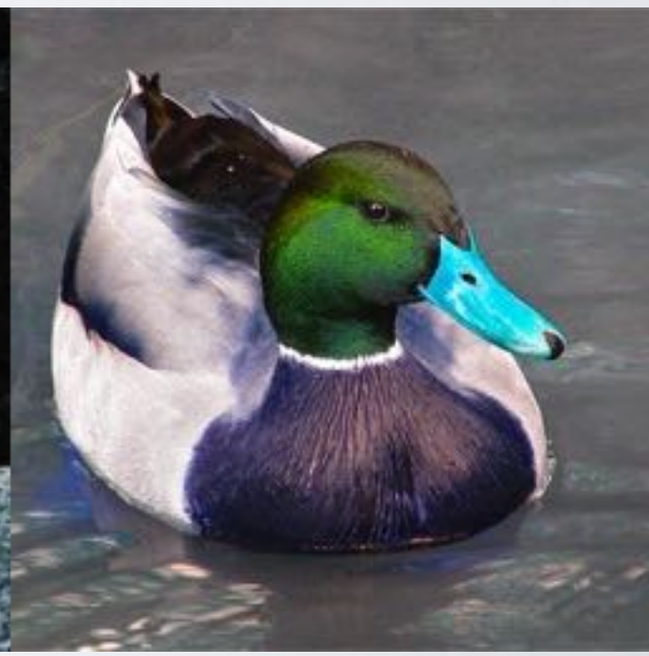


From https://www.tensorflow.org/tutorials/generative/pix2pix

Validation

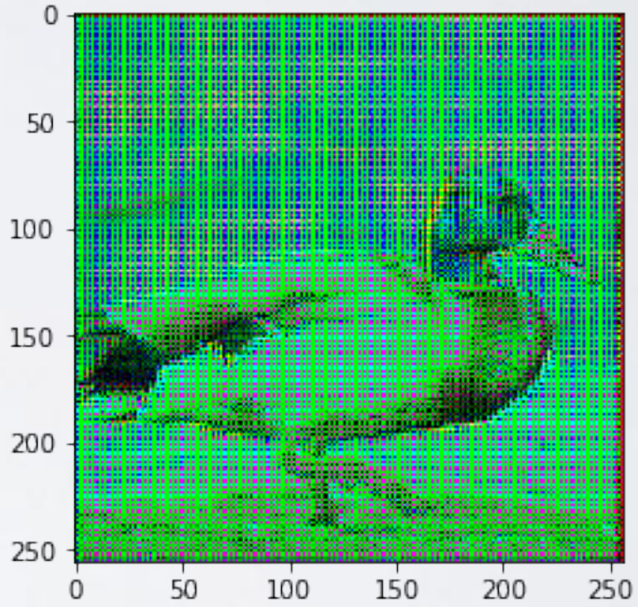
Test



BUILD THE GENERATOR

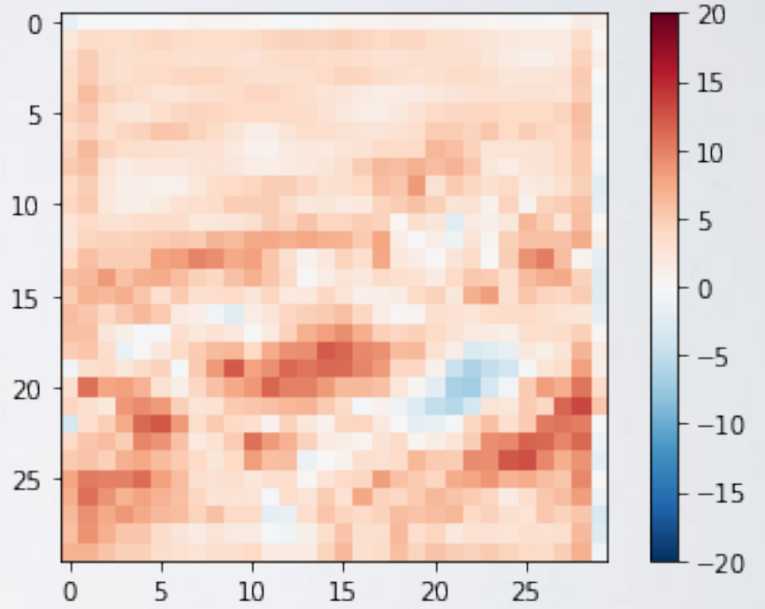
- The architecture of generator is a modified U-Net.
- Each block in the encoder is (Conv -> Batchnorm -> Leaky ReLU)
- Each block in the decoder is (Transposed Conv -> Batchnorm -> Dropout(applied to the first 3 blocks) -> ReLU)
- There are skip connections between the encoder and decoder (as in U-Net).



Generator Loss Sample

BUILD THE DISCRIMINATOR

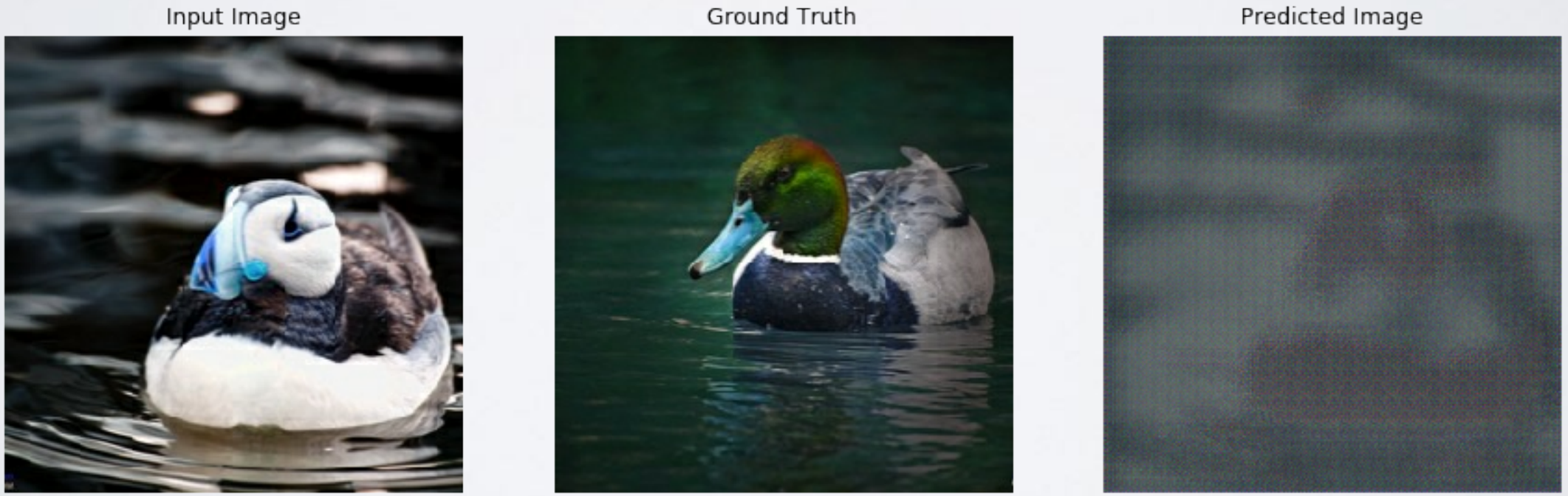
- The Discriminator is a PatchGAN.
- Each block in the discriminator is (Conv -> BatchNorm -> Leaky ReLU)
- The shape of the output after the last layer is (batch_size, 30, 30, 1)
- Each 30x30 patch of the output classifies a 70x70 portion of the input image (such an architecture is called a PatchGAN).
- Discriminator receives 2 inputs.
- Input image and the target image, which it should classify as real.
- Input image and the generated image (output of generator), which it should classify as fake.



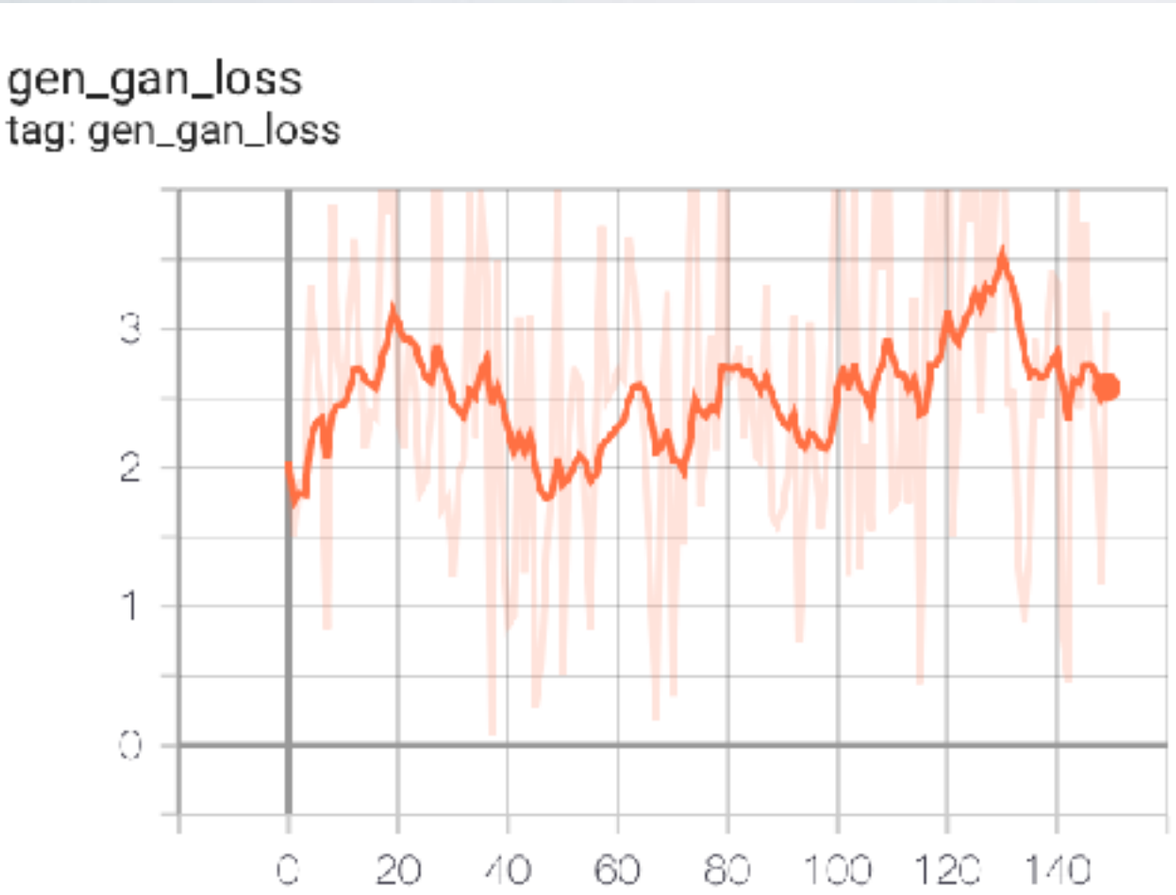
Discriminator Loss Function

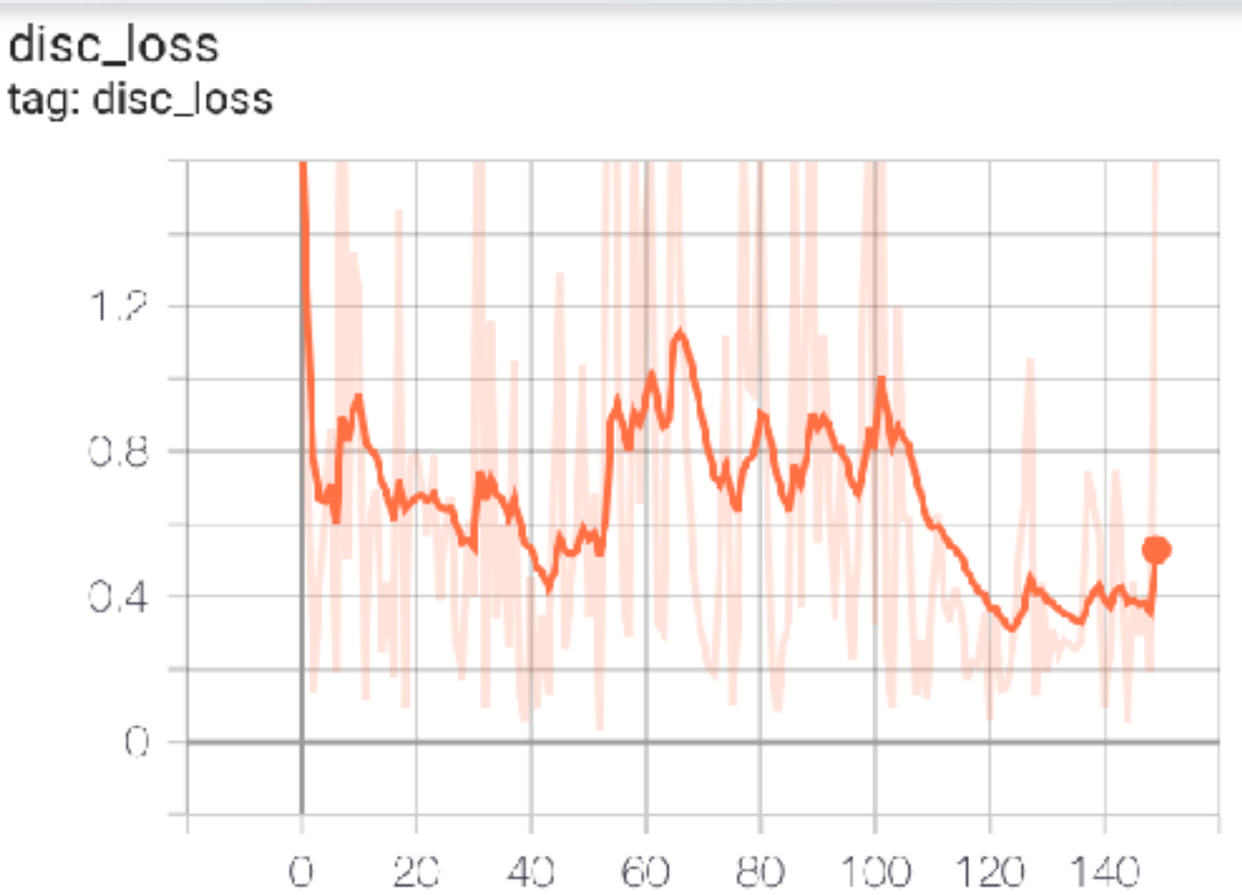
TRAIN THE MODEL

- 150 epochs
- Lambda 100 (don't know why this was chosen)

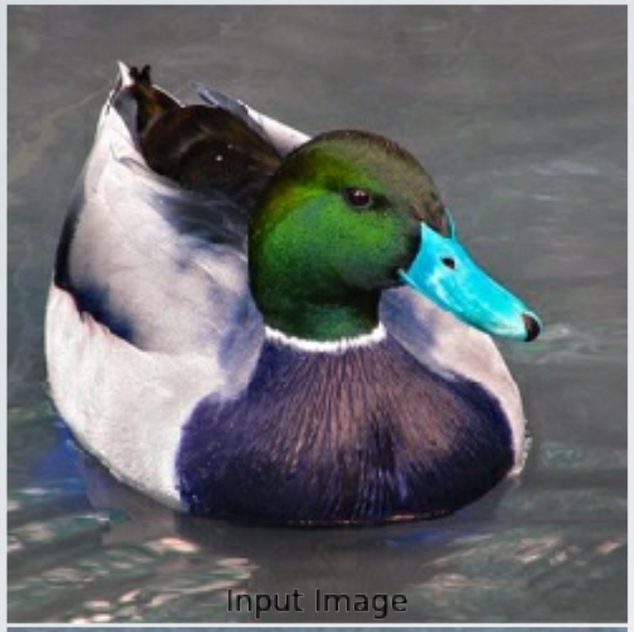


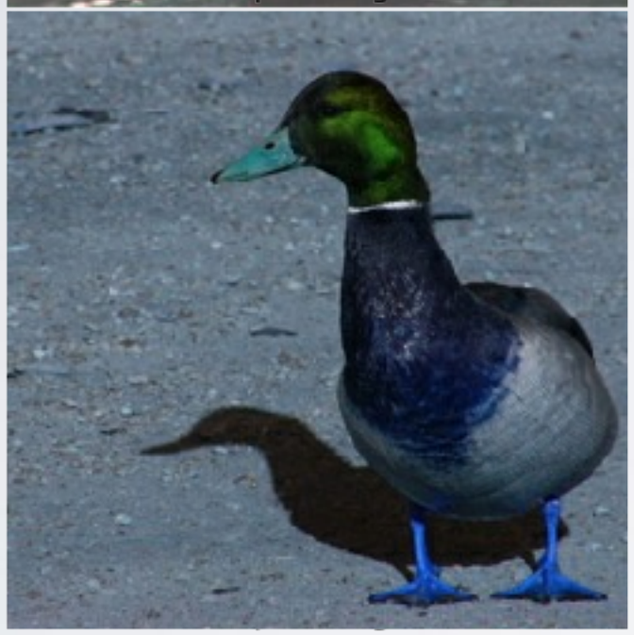
Tensorboard Sample Plots (only 5 epochs)

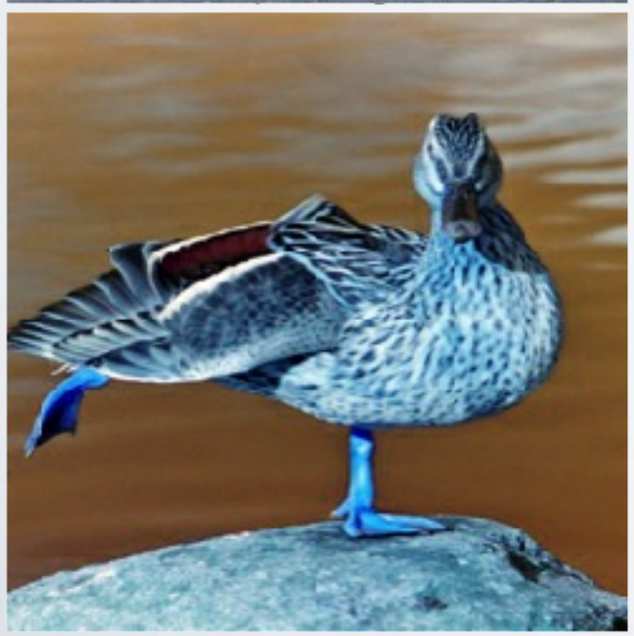




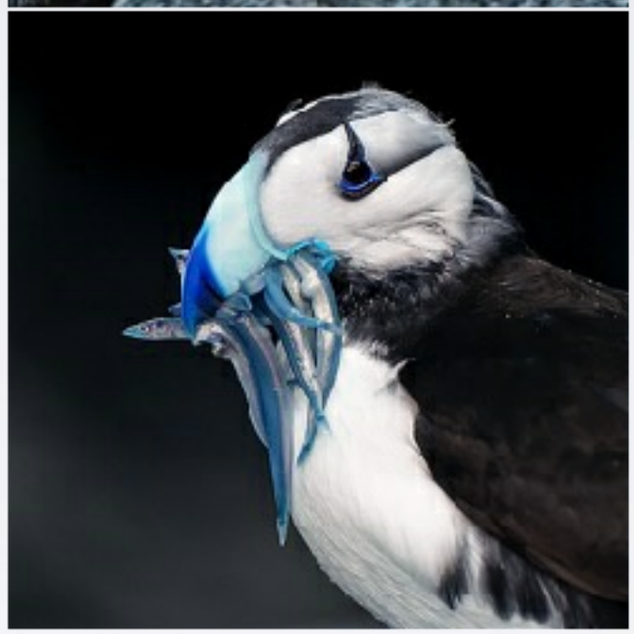
Input Image



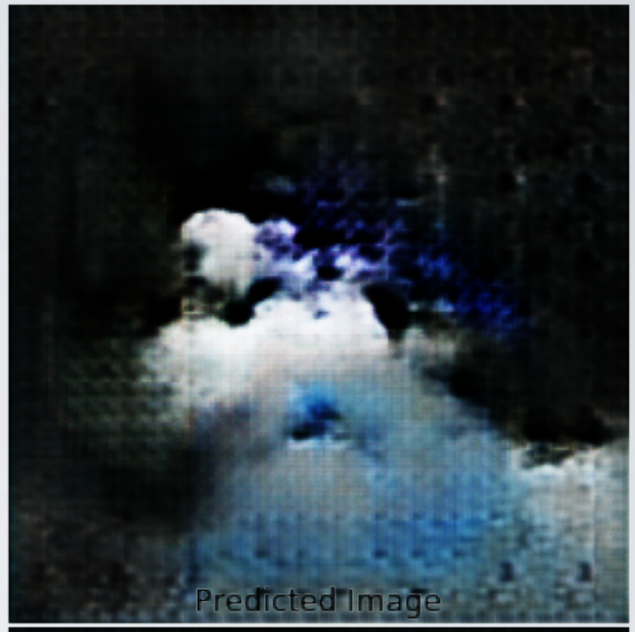


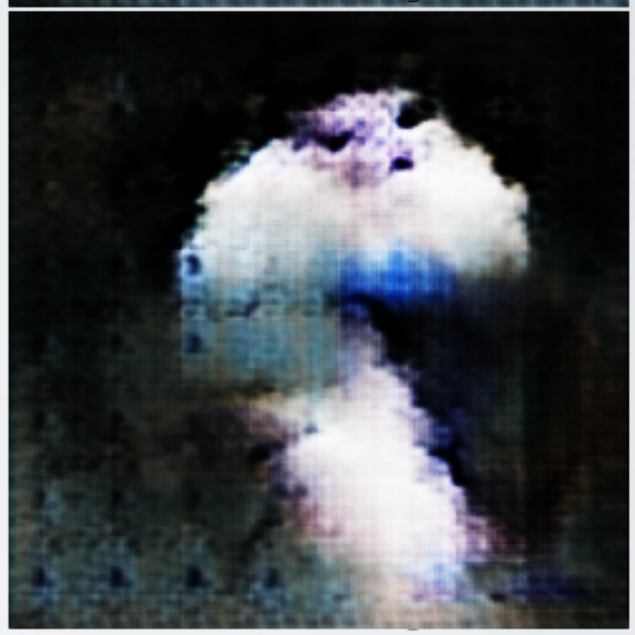


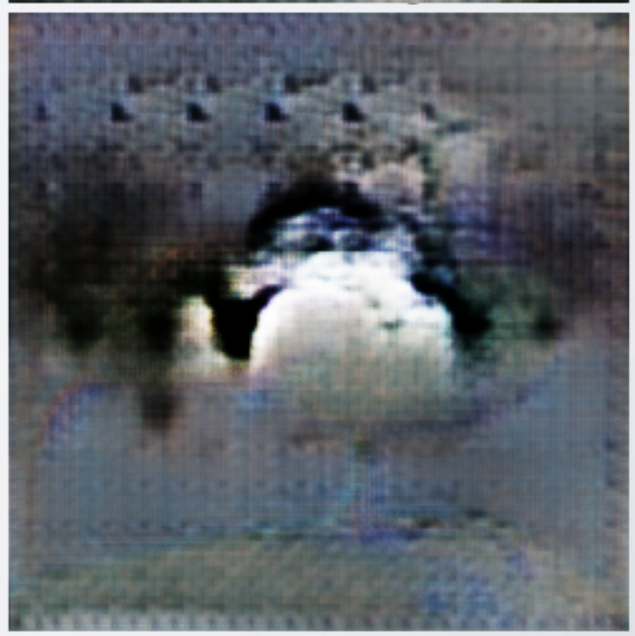
Ground Truth



Predicted Image



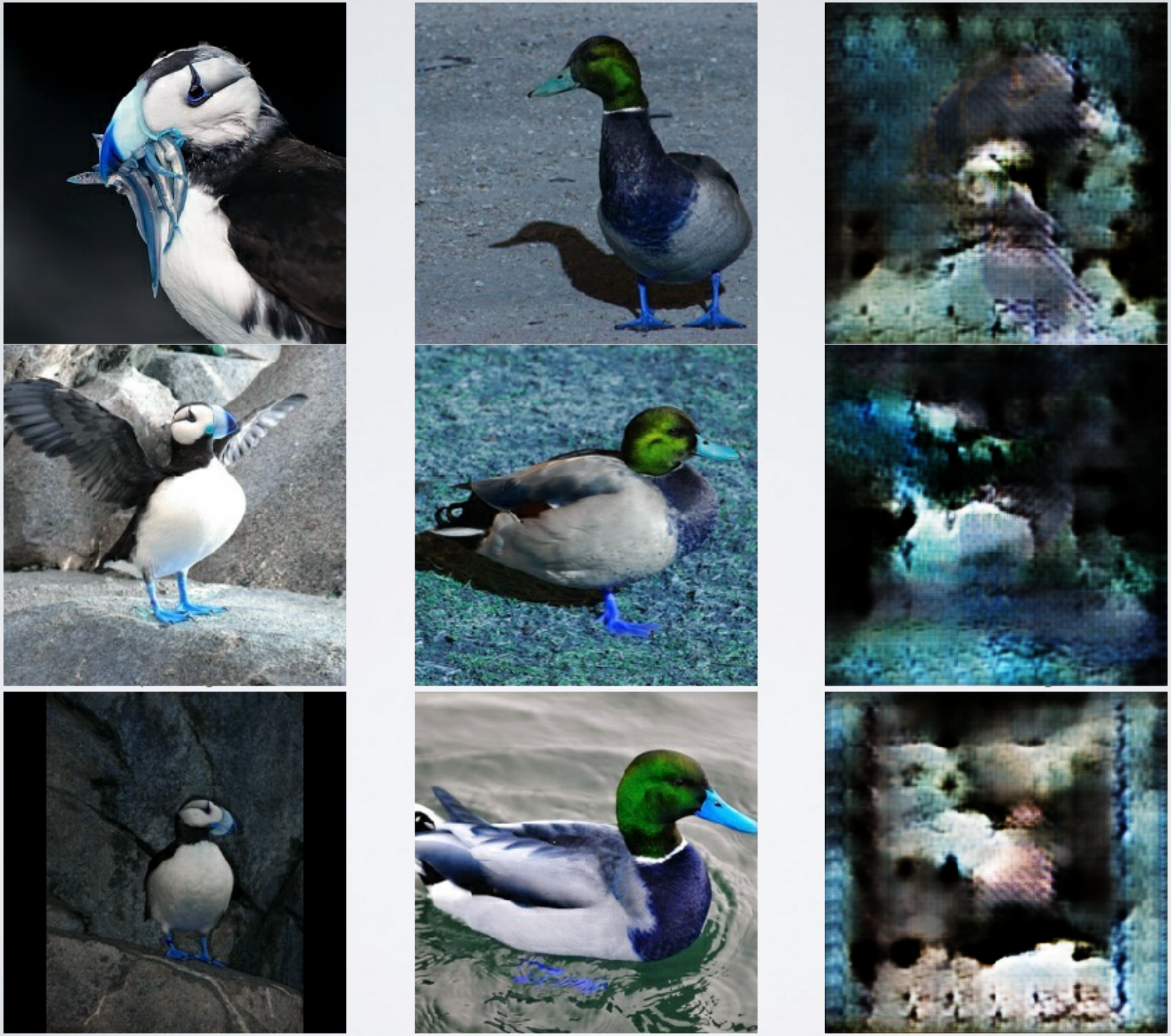




Input Image

Ground Truth

Predicted Image



POSSIBLE WAYS TO IMPROVE THE RESULTS

- Add More Data (Puffins!)
- Clamp color channels to easily discriminated color samples (posterize the images)
- Align the geometry between the samples

Whitman GPT-2 Code and Data available here

Pix2Puffins Code available here

THANKS…

- To Manu for walking us through the failures of DCGAN

- Henry for reminding me that I needed to mount my drive

- Grad Student strikers for giving us the long weekend