Final Project - Luyang Wang 65788904

First task- Least Significant Bits

Algorithm:

The basic idea is to replace the last 4 bits in original image's pixel with the first 4 bits of secret image's corresponding pixel.

Example:

Choose figure 1 as an original image, figure 2 as a secret image, and figure 3 as the output image.

Figure 1: Original image

Figure 2: Secret image

Figure 3: Output image

It is obvious that there are noticeable differences comparing figure 1 and figure 3, but the figures are the mostly the same.

Second Task- Neural Network

Networks:

In this assignment, three neural networks have been applied: a preparation network is applied to the secret image to reduce its information, a hiding network is applied to merge the original image and the prepared secret image, and a reveal network is applied to recover the secret image.

5 layers of convolutional neural networks with 3*3, 4*4, 5*5 patches are used to approximate the mappings.

Image Dataset:

The images of size 256*256 pixels and 3 channels are divided into training and test datasets in 8:2 ratio. The image is first normalized before sending into the model.

The model is trained with the full training set for 3 epochs. Two images are randomly selected to form a batch of the original image and secret image.

Results:

There is no obvious difference between the output image of neural network with the original image. The model loss is defined as the sum of the difference between the original and output image, and the difference between the recovered secret and the real secret image.

Training loss:

Figure 4 shows the training loss of different batches. The training loss decreases rapidly within the first 20 batches and vibrates afterwards. The mean training loss is 0.0874.

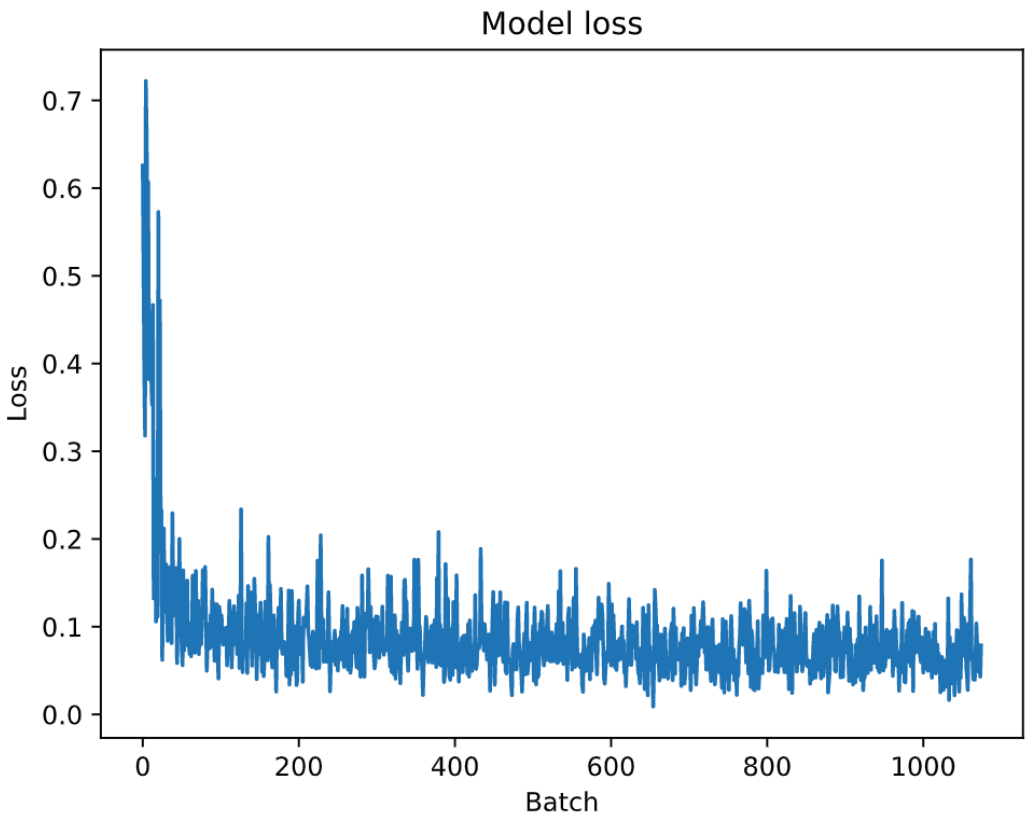

Figure 4: Training loss of different batches

Test loss:

Figure 5 shows the test loss of different batches. The mean test loss is 0.07.

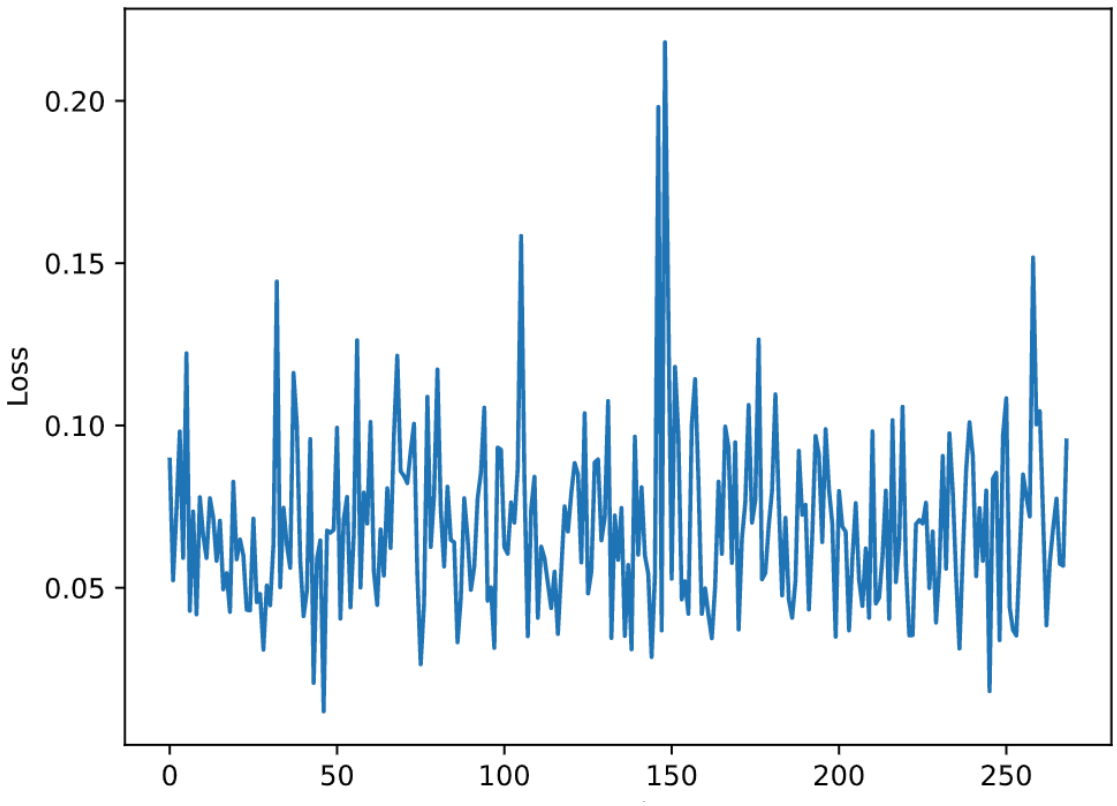

Figure 5: Test loss of different batches

Kodak test loss:

The trained model is also used to test the Kodak image dataset. Figure 6 shows the test loss. The mean test loss is 0.06, which indicates the neural network works with the Kodak images.

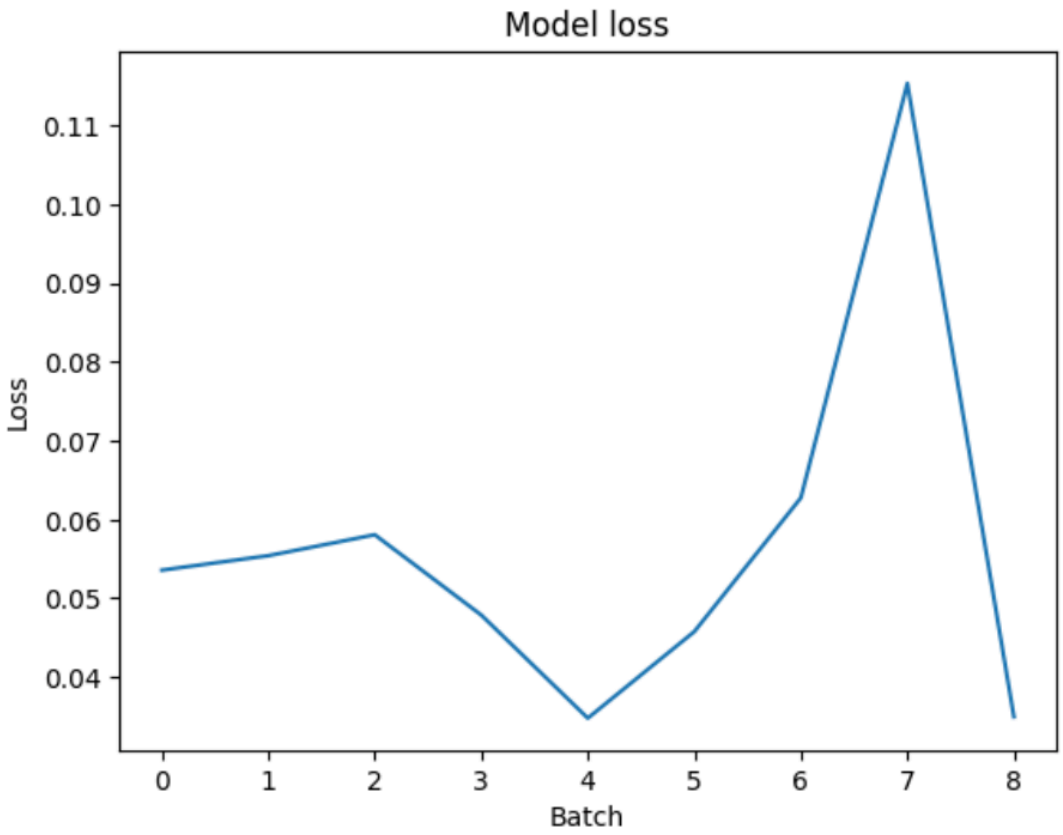

Figure 6: Kodak test loss of different batches

Conclusion:

The neural network works well wit new image dataset, which may indicate it is a more reliable tool over LSB.

Reference:

https://www.kaggle.com/gaz3ll3/optimization-ii-project-3

https://github.com/kelvins/steganography/blob/main/steganography.py

https://github.com/fpingham/DeepSteg/blob/master/DeepSteganography.ipynb

http://r0k.us/graphics/kodak/