Deep Learning - Image Reconstruction Challenge
Spring 2023
Group number - 14
Mid-Eval Report

Group Members:

1. Vivek Jaiswal (20CS10077)
2. Amit Kumar (20CS30003)
3. Anuska Srivastava (20CS10009)
4. Anand Parikh (20CS10007)

Description of Dataset:

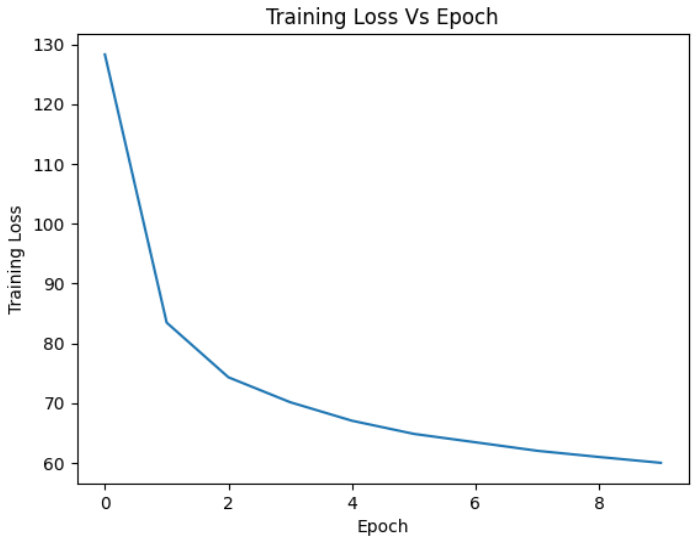
- The dataset contains 256*256 pixel images of 4 animals: cats, dogs, elephants and tigers.
- In the folder: Testing_Data, there are many masked images of the animals in .jpeg files. Each image has 2 masked squares of dimension 75*75 pixels. There is also one masked_info.csv file, in which each entry is a 4-value tuple: the top left row and column coordinates of both the masked squares.
- The objective is to fill in these blacked-out masks of each image to perform accurate inpainting of the pictures.
- In the folder: Training_Data, there are 4 subfolders (one for each animal). Each subfolder further contains 2 subfolders:
    - Masked_Train has contents similar to the Testing_Data folder (masked .jpeg images and a masked_info.csv file). They are the inputs to the model for training.
    - Unmasked_Train has un-damaged (complete) images. They are the expected outputs of the model for training the corresponding masked images.

Data processing: The images were converted from a 256 * 256 RGB image to 1 * 3 * 256 * 256 tensor to feed into the model

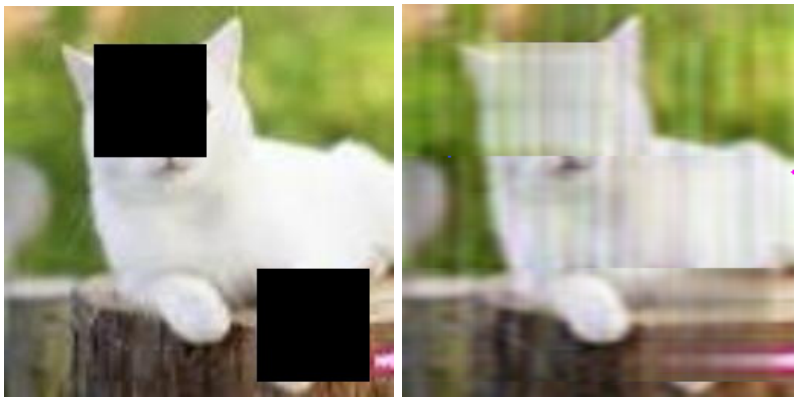
Architecture used: A Bi-LSTM model with the following parameters was trained from scratch-

```
# Model Params
model = BiLSTM(input_size=256, hidden_size=128, num_layers=2)
# Define the loss function
loss_fn = ReconstructionLoss()
# Define the optimizer
optimizer = torch.optim.Adam(model.parameters(), lr=0.001)
```

1. Training Loss is taken as the Sum of the MSE of all the images:



Example reconstruction of destroyed cat image:



Next step: To find an appropriately pre-trained model and fine tune it with our dataset.