Deep Learning Small Project Report

Li Xingxuan 1002189 Hao Weining 1002195

1. What we did:

a. **Model:** We used pretrained ResNet18 to perform transfer learning. To comply with the dataset, we changed the last fully-connected layer to 20 neurons. Since we are dealing with multi-class and multi-label problem, we add a sigmoid activation function to the end of the fully-connected layer.

b. **Loss function:** We used binary cross entropy loss. Since the labels are independent from each other, we need this loss function to calculate loss of each neuron of the output independently.

c. **Learning rate:** To save computational power, we set the learning rate to 0.02.

d. **Training procedures:** We have in total 3 training phases.

   i. In phase 1, we resize the data to 224 and perform a center crop. By training the pretrained ResNet18 with these data for 15 epochs, we now entered phase 2.

   ii. In phase 2, we performed data augmentation by randomly rotating the image between -10 and 10. And then we trained the model again for 15 epochs.

   iii. In phase 3, we performed another method of augmentation by setting both jitter saturation and jitter hue to 0.05, as well as randomly horizontal flipping. And finally, we trained the model again for 15 epochs.

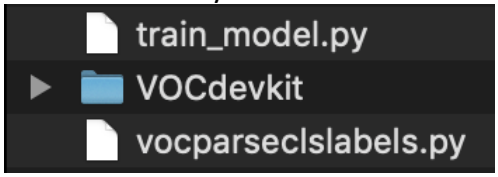
e. **Other hyperparameters:**

   i. **Threshold:** threshold of deciding whether it has such label, which will affect the precision.

2. How to reproduce:

a. **How to train the model:**

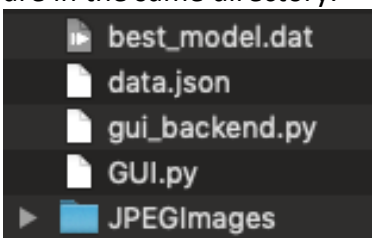
The file directory should look like this:



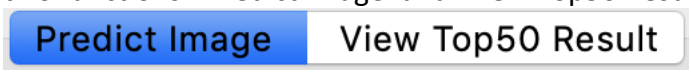
Ensure that the above three files are in the same folder. Then run the python file “train_model.py”

b. How to run the GUI:

i. After the model is trained, state of best model and data of top 50 ranked images for each label will be stored. Ensure the following files are in the same directory.



ii. Run GUI.py and the interface will show up. There two tabs indicating two functions: 'Predict Image' and 'View Top50 Result'.

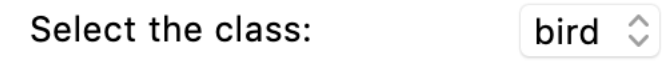


iii. The first tab 'Predict Image' allows you to select your own image and see the classification results. The result will show the objects detected in the image and will be displayed at the bottom of the image. If no objects are detected by the model, you will see the result as 'Not Identified'.

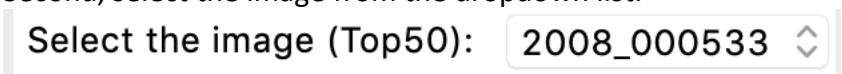
Click the 'Open File' button to select your image.

iv. The second tab 'View Top50 Result' allows you to browse the top 50 images with highest scores corresponding to each class.

First, select the class you want to look at;



Second, select the image from the dropdown list.

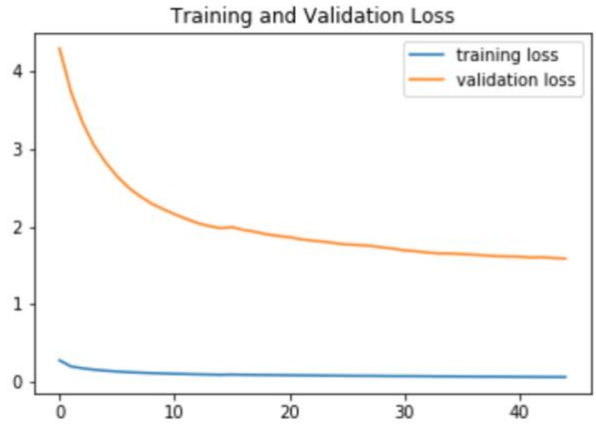


The selected image will be displayed with a confidence level at below.

v. Click the 'Close' button on the first tab 'Predict Image' to close the GUI.

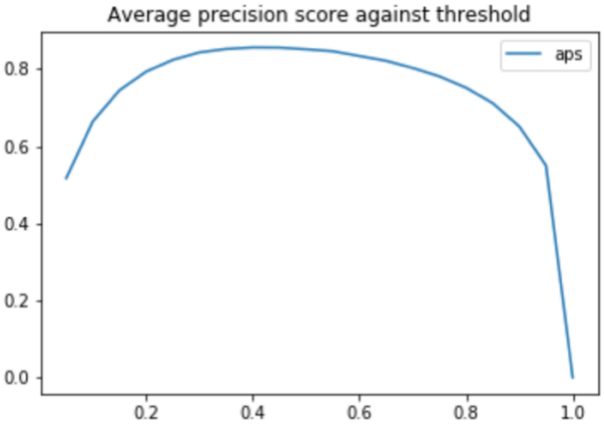
3. Result and evaluation:

a. The graph below shows the training and validation loss through 45 epochs in total:



As we can see in the graph, both training and validation loss are decreasing through the training. However, at the first 10 epochs, validation loss decreased much faster than the training loss.

b. The graph below shows a plot of prevision averaged over all 20 classes for 20 values from 0 to 1 with interval 0.05.



As we can see in the graph, the precision reaches the highest when threshold is 0.4. And the highest precision is 0.8565135759676487.