

Computer Vision Assignment 4

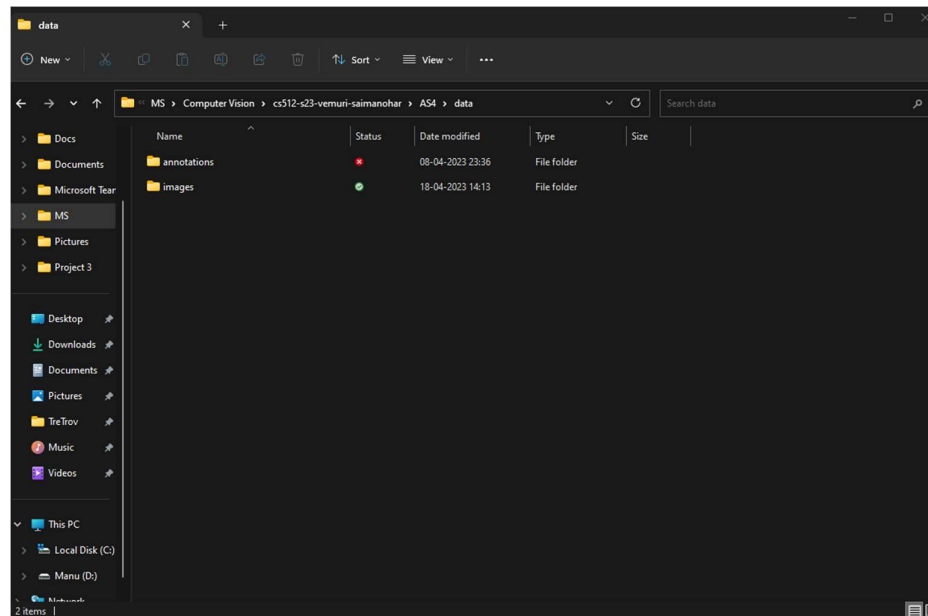
Sai Manohar Vemuri

A20514848

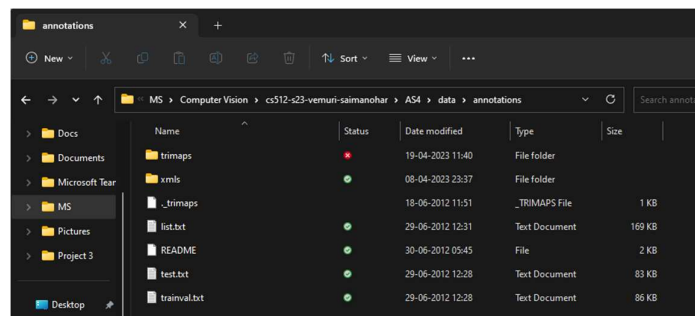
1. A) I have downloaded the oxford pet dataset from the link provided.

B) I have converted the cat/dog breed labels to cat/dog labels using the file “list.txt” which is found in the dataset. And we have converted it into dataframe and modified cat as “0” and dog as “1”

“../data/annotations/list.txt” for this to work we need to have this file in the same directory. For the above and below codes to work we must download the original dataset folder and it should be in this file structure



And annotations folder must and should contain all these files and folders.



C) I wrote a function to copy all the images with the species name into a folder named "categorized". Where all the images with cat and dog are stored with the file name "cat_" + {i} in the loop. Same for the dog.

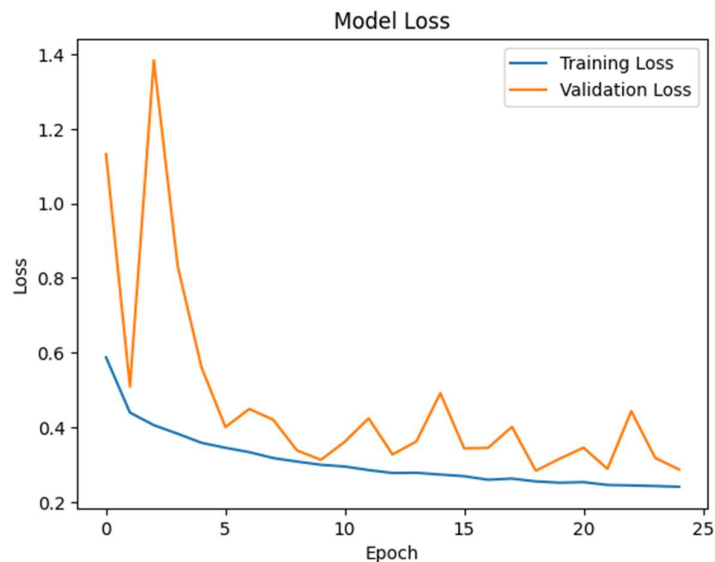
This will create a folder in the images folder which we have downloaded. So do remember to download the original dataset.

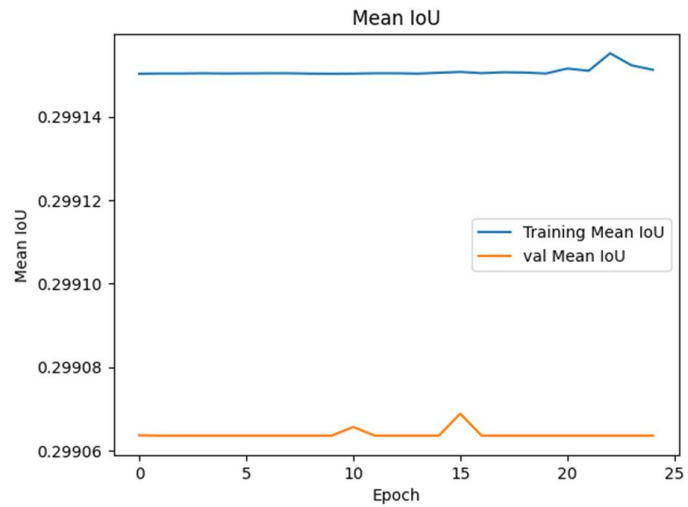
Next i have retrieved all the images and stored in the two different arrays called `img_array` and `mask_array`. And then i have preprocessed the images and masks using the `preprocess_data` function. next I combined both masks and images and shuffled. Then next I have plotted the image and mask side by side.

For this to work, we need the masks stored in the `'../data/annotations/trimaps'` folder and images should be in `'../data/images'` folder.

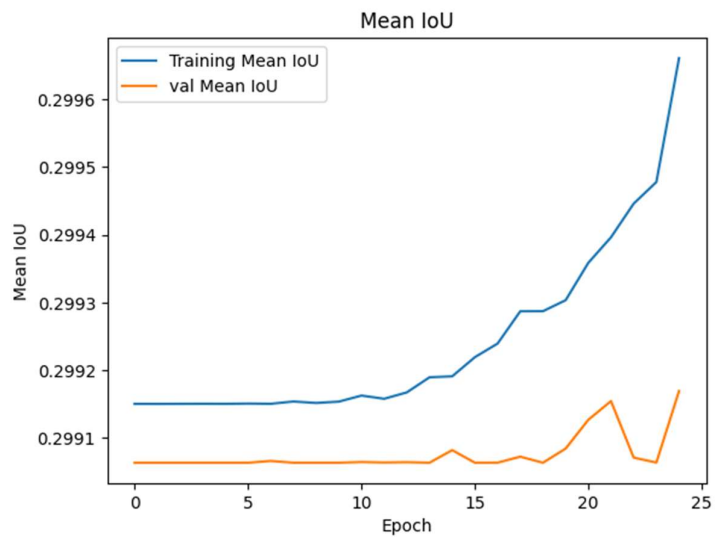
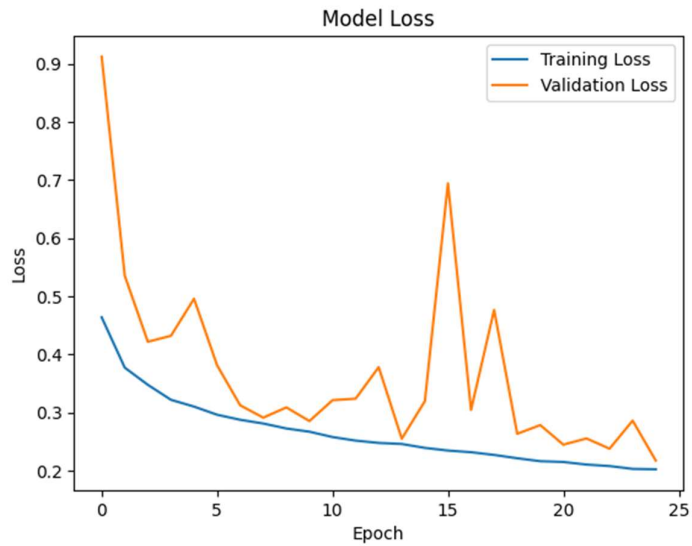
D) Next, I have split the dataset into train, test and validation datasets using the 60:20:20 ratio.

E) In this question I have built a simple unet model without any skip connections between encoder and decoder. I have used adam optimizer with learning rate of 0.01. we ran the model for 25 epochs with batch size of 32. And we got the mean iou of 0.2992 and it's not improving.

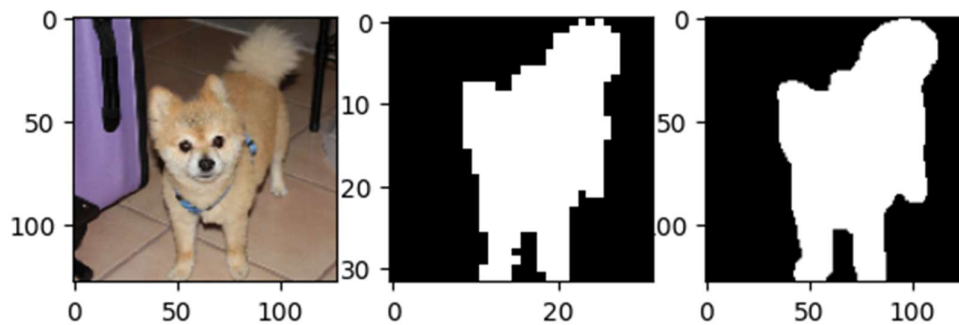
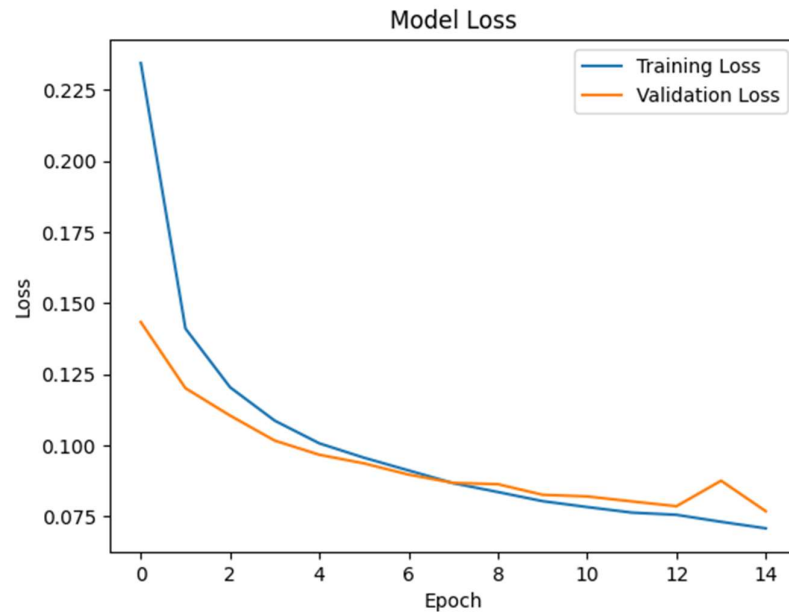




F) In this we included the skip connections for the previous model. And below I have provided the graphs for loss and meaniou. We can see that it fluctuating more compared to the previous model



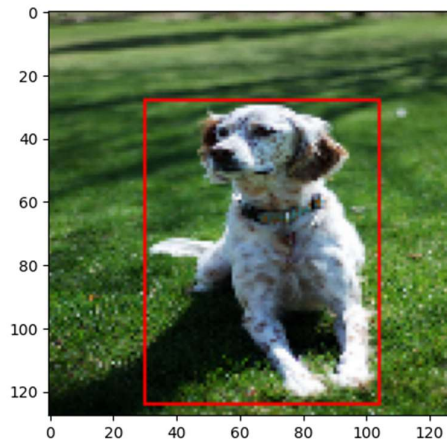
G) In this question we have used the pretrained model and for that we have converted the data into tensorflow datasets. We ran this for 15 epochs and we can see from the below graph that loss is decreasing and it is performing better than what we defined in the previous question.



The above image shows the output predicted from the model.

2. A) In this question we have defined a bounding box using the `find_countours()` function we defined. Here we are passing the masks data which we previously computed and images which was stored in the images dataset. And the corresponding bounding boxes are stored in the `bounding_box_vals` numpy array and this array has been dumped as a pickled format in the local storage.

For this to work we need data folder. And the resulting pickle file will be stored in `'../data'` folder



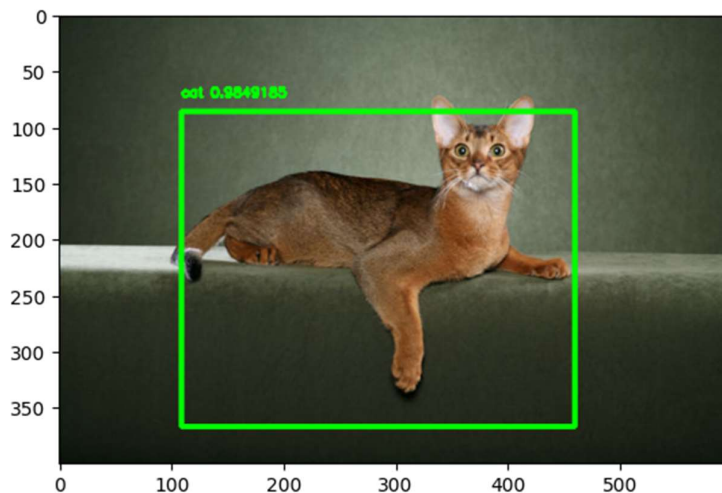
In the same question we have used XML parser and created a XML file containing bounding boxes for each image. these are stored in the “/data/xml” directory.

B) In this question we have downloaded the pre-trained yolo model weights using the curl script and we have used the yolo functions from the git repository mentioned in the class lecture pdf.

C) I read the weights using WeightReader() function and saved the model weights using model.save() function. saved format is in .h5 format.

You need to give the path to the variable where the weights file is stored.

D) We have loaded the same model which we saved to the local storage and printed the summary of the model. Next, we have predicted the output by passing the image as input to the model.



If this has to work, we may need to install the keras and tensorflow libraries which I am using because I have changed few things in the repository that I have downloaded to work in my local system. I have included the environment library versions in environments.yml or requirements.txt.

Note:

Please download the entire keras-yolo3 folder which I have uploaded as I have made some changes to the python files to work on my system. And I have renamed the weights file which I have downloaded as yolov3.weights and then I have saved it as "yolo_model.h5".
E) In this question we have calculated the mAP for the images and but cannot complete it. But I have computed the MeanIOU for the first row of both gt_label.txt and pred_label.txt. And it was 0.23.

Note: I have uploaded the entire assignment folder including the data, as I have made few changes to the .py files downloaded from the git repo for the 2nd question. So, please download entire assignment folder.