Building a deep learning model for image segmentation using Pytorch library

Lakshmikar Reddy Polamreddy Yeshiva University
lakshmikarpolamreddy@gmail.com

Abstract

An attempt has been made to build a deep learning neural network model in Pytorch library for the sake of image segmentation. To develop our model, we have used an architecture similar to Unet model. This model has been trained on 800 images and respective masks of images of lungs.. After training our model for 200 epochs, we have used it to predict the results of 200 unseen images. We have used the metric IoU (Intersection over Union) to check the performance of our model.

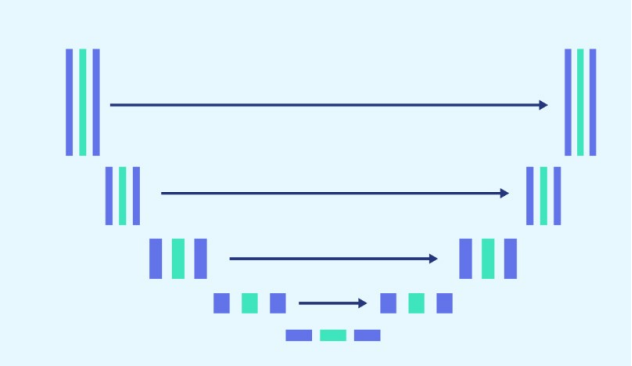
1. Introduction

Image segmentation is a method in which a digital image is broken down into various subgroups called Image segments which helps in reducing the complexity of the image to make further processing or analysis of the image simpler. Segmentation in easy words is assigning labels to pixels. All picture elements or pixels belonging to the same category have a common label assigned to them. For example: Let's take a problem where the picture has to be provided as input for object detection. Rather than processing the whole image, the detector can be inputted with a region selected by a segmentation algorithm. This will prevent the detector from processing the whole image thereby reducing inference time.

In this project, this task is done by building deep a learning neural network model using libraries such as Keras, Tensorflow and Pytorch. For this project, we have used Pytorch because of its better debugging capabilities. In addition to this, Pytorch is faster than Keras especially when dealing with large datasets.

Image segmentation has many applications in Medical field in addition to other fields. In this project, we have used this for medical images of lungs

We have developed the model from scratch which looks similar to the Unet models in terms of encoder and decoder. In this case, we preferred to use GoogleNet. Its architecture is shown below



2. Related Work

We have referred to the project CARLA Image Semantic Segmentation with DeepLabV3+ [3] related to self driving cars, which was found in Kaggle platform. This project objectives were to create a Model to predict semantic segmentations of CARLA images and predict masks using the model and compare with ground-truth masks. They have achieve a mean IoU score of 0.65

A tutorial on [1] have been very helpful to understand how to prepare datasets and how to pass them to a dataloader in pytorch library. It has also thrown insights on Unet model architecture and metrics such as IoU and Dice coefficient.

3. Methods

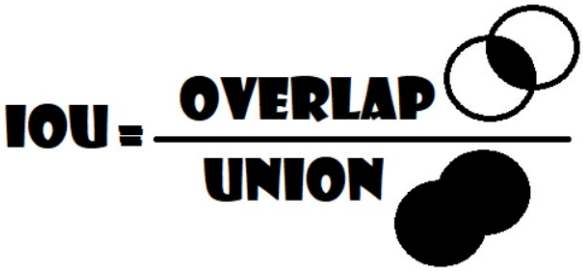
We have considered cross entropy loss as this project is related to image classes and below is the equation

$$\text{Cross Entropy Loss} = \sum_{c=1}^{M}(y_{o,c}log(p_{o,c}), \qquad (1)$$

where M is number of classes, y is binary indicator (0 or 1) if class label c is the correct classification for observation o and p is the predicted probability observation o is of class c

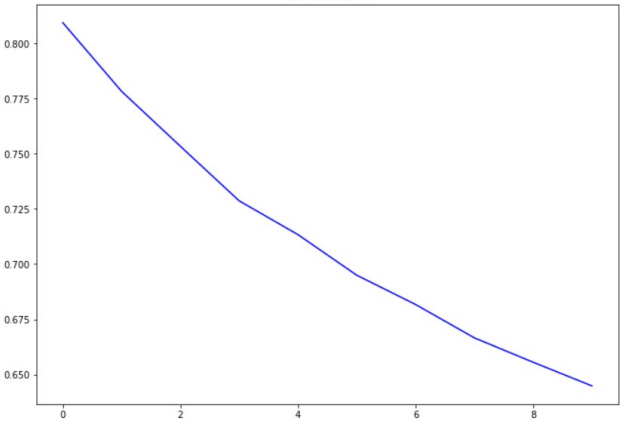
intersection over union (IoU) is also known as the Jaccard index, is the metric used to check the performance of the model. Using IoU implies that we have two images to compare: one is our prediction and the other one is the ground truth, if the value obtained approaches number 1

that means the prediction is similar and close to our ground truth.



4. Results

Below graph shows loss value for validation images versus number of epochs for my model that was run for 200 epochs. This has been plotted with number of epochs on X-axis and loss values on Y-axis.

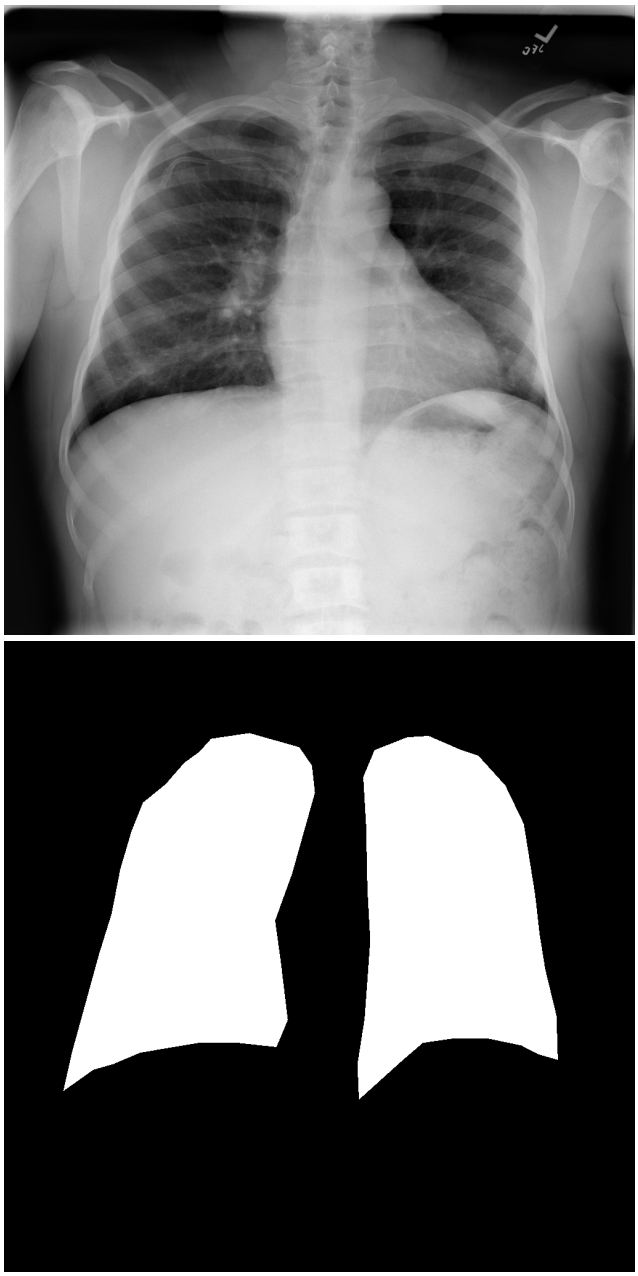


IoU score on the validation set and test set is abnormally low as there might be some problem in the training. We will work on this in future to resolve this issue to attain a good IoU score.

4.1. Datasets

We have considered 800 medical scan images of lungs for training our model. 100 images out of these have been used for validation purpose. A batch size of 16 images have been selected for running each epoch. Along with the images, we have been provided with respective masks. We have sourced this data from github link [2]

We would like to show a below sample image and respective mask to get an idea about the training data.



The trained model was applied on 200 unseen images to predict the masks.

Data set	Train data set	Validation data set	Test data set
Number of images	700	100	200

5. Discussion

As Unet model is superior for semantic segmentation tasks, we have developed a model architecture similar to it.

Regarding loss, I have used Cross Entropy loss function for to deal with the classes of images. I would like to mention that we have only one object i.e. lungs in these images.

Regarding hyper parameters, we have used Color jitter, horizontal flip, learning rate, batch normalization to achieve better training.

Regarding number of epochs, we have run for 200 epochs and stopped as we have achieved convergence in the loss values.

Regarding the evaluation metric of the model, we have used IoU score to measure the performance of the model.

6. Conclusion

This project has been started to predict the masks of unseen medical scan images of lungs with high accuracy to attain a score of 0.82. But, We could not achieve this and the reason might be the architecture and improper training of the model. In future, we will work on this to improve IoU score.

References

[1] https://blog.paperspace.com/image-segmentation-using-segmentation$_{m}odels_{p}ytorch/$. 1

[2] https://github.com/youshanzhang/lung$_{s}egmentation$. 2

[3] https://www.kaggle.com/code/oluwatobiojekanmi/carla-image-semantic-segmentation-with-deeplabv3. 1