U-Net Image Segmentation using Keras: A Technical Report

Dr. Saurabh Jain

March 2, 2024

1 Introduction

The aim of this report is to demonstrate the application of the U-Net architecture for image segmentation using Keras. Image segmentation is a crucial task in computer vision, involving the partitioning of an image into multiple segments to simplify its representation and make it more meaningful for analysis. U-Net, with its encoder-decoder architecture, has shown promising results in various segmentation tasks.

2 Methods and Materials

In this study, we utilize the U-Net architecture implemented in Keras, a high-level neural networks API, with TensorFlow as the backend. The dataset used for training and validation is The Oxford-IIIT Pet Dataset `https://www.robots.ox.ac.uk/ vgg/data/pets/`, which consists of images of cats and dogs. Each image is annotated with pixel-level segmentation masks, delineating the regions corresponding to the pets in the images.

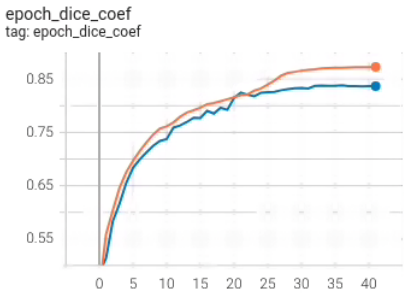
To prepare the dataset, we randomly divided it into a training set and a validation set. The validation set comprises 1000 images, ensuring a sufficient amount of data for evaluating the model's performance. We employed Python as the programming language, utilizing the PyCharm IDE for development.

The model was trained utilizing the Adam optimizer, incorporating a combination of focal loss and dice loss as the chosen loss functions.

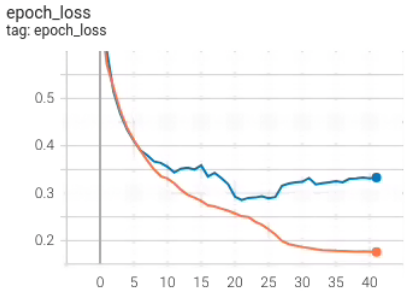
3 Results

Figure 1 shows training performance. For the evaluation of the model, we present the segmentation results on the first four images from the validation dataset. These images as shown in Figures 2-5 are chosen to provide a representative sample of the model's performance across different instances.

The segmentation results demonstrate the effectiveness of the U-Net architecture in segmenting the pets in the images.



(a) Dice coefficient as accuracy.



(b) Loss curves.

Figure 1: Performance curves. Orange colour is training curve and blue color is for validation.

4 Discussion and Conclusion

The U-Net architecture, with its encoder-decoder design, has proven to be highly effective in image segmentation tasks. By leveraging the hierarchical features learned through the encoder pathway and the precise localization provided by the decoder pathway, U-Net can produce accurate segmentation masks.

In this study, we successfully applied the U-Net architecture to the task of pet segmentation using the Oxford-IIIT Pet Dataset. During training we observe that the model starts over fitting around epoch 23. Therefore, we evaluate the model performance at epoch 22. The segmentation results on the validation set indicate the model's ability to delineate the pets from the background with high accuracy.

In conclusion, U-Net implemented in Keras offers a robust solution for image segmentation tasks, and its effectiveness has been demonstrated in the context of pet segmentation. Further experimentation and fine-tuning may enhance the model's performance on this and other segmentation tasks.

5 Code

The code for this project is available on https://github.com/saurabhjain-1986/UNet. It includes implementations of the U-Net architecture in Keras, data preprocessing steps, training procedures, and evaluation scripts.

Figure 2: Image 1 segmentation result.

Figure 3: Image 2 segmentation result.

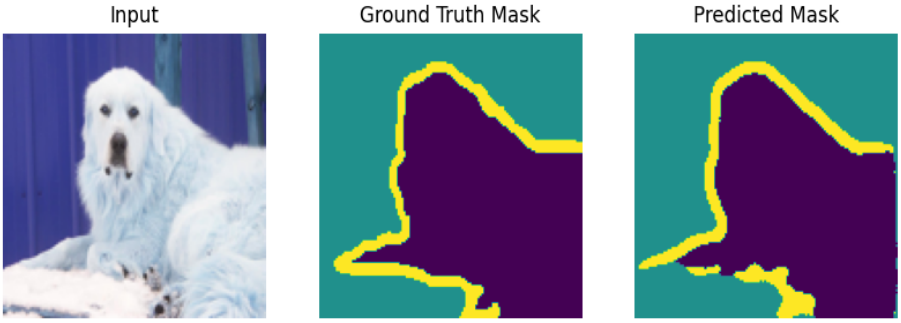


Figure 4: Image 3 segmentation result.

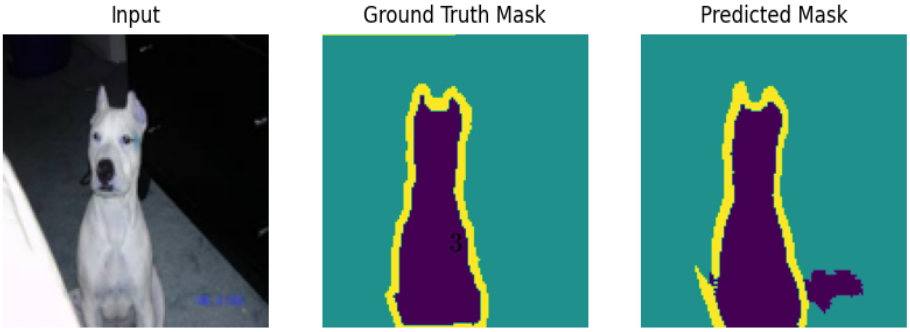


Figure 5: Image 4 segmentation result.