Building a Segmentation Model

Anonymous submission

Paper ID

Abstract

This paper proposes a segmentation model to separate road areas from driving data. Using the KITTI dataset, we trained and compared the performance of U-Net and U-Net++ models. Data augmentation techniques were applied to enhance the generalization performance of the models, and various evaluation metrics were used to analyze their performance. Experimental results showed that the U-Net++ model achieved higher accuracy and IoU scores, confirming its effectiveness in road area segmentation.

1. Introduction

Accurately segmenting road areas is a crucial task in autonomous driving and driver assistance systems. In this study, we propose a method to segment road areas using a segmentation model. We train U-Net and U-Net++ models on the KITTI dataset and compare their performances to identify the most suitable model for road area segmentation.

2. Background, Related Works

Segmentation models are techniques used to separate specific areas in images and are employed in various fields such as autonomous driving and medical imaging. U-Net is a model primarily used in the medical imaging field and can be effectively applied to road area segmentation. U-Net++ is an extended version of U-Net, which incorporates DenseNet architecture to enhance performance.

3. Method

In this study, we trained a road area segmentation model using the KITTI dataset. Data augmentation techniques were employed to increase the diversity of the training data, and both U-Net and U-Net++ models were trained. The structure of each model is as follows:

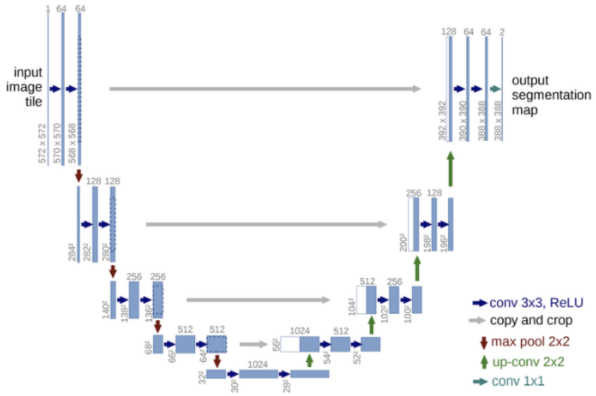
3.1. Dataset Preparation

• KITTI Dataset • Data Augmentation: Albumentation

3.2. Model Structure

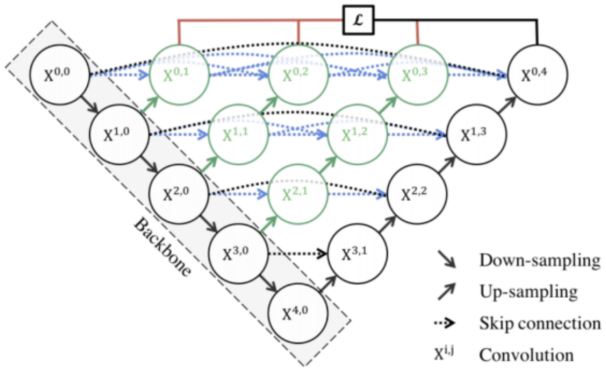
3.2.1 U-Net

Figure 1. U-Net Input: (16, 224, 224, 3) Output: (16, 224, 224)



3.2.2 U-Net++

Figure 2. U-Net++ Input: (16, 224, 224, 3) Output: (16, 224, 224)



3.3. Training

Train/Test Data Generation. Model Training and Evaluation Metrics Calculation (Accuracy, IoU).

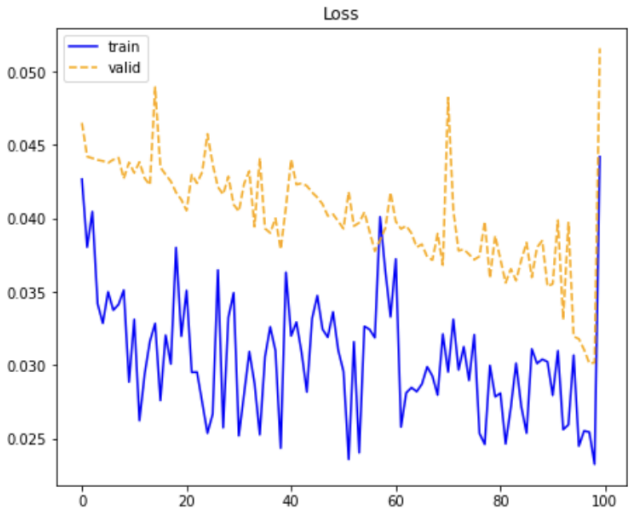


Figure 3. U-Net 100 Epoch

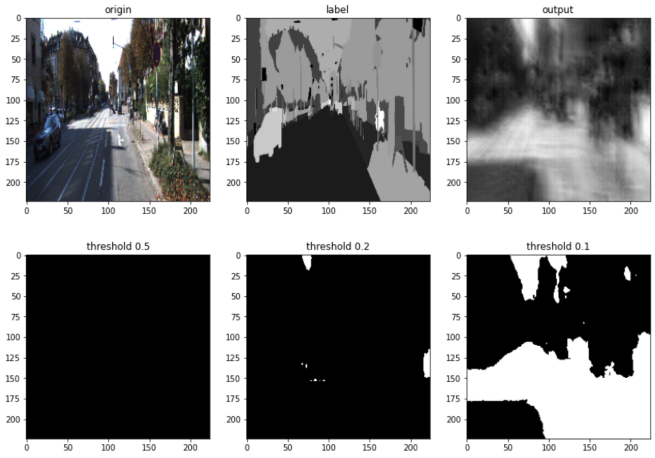


Figure 4. 100 Epoch Predict Images

4. Result

The training results showed that the U-Net model had high initial loss and low accuracy and IoU scores. In contrast, the U-Net++ model, which applied the DenseNet structure, achieved higher accuracy and IoU scores. Ultimately, the U-Net++ model trained for 1000 epochs exhibited the best performance.

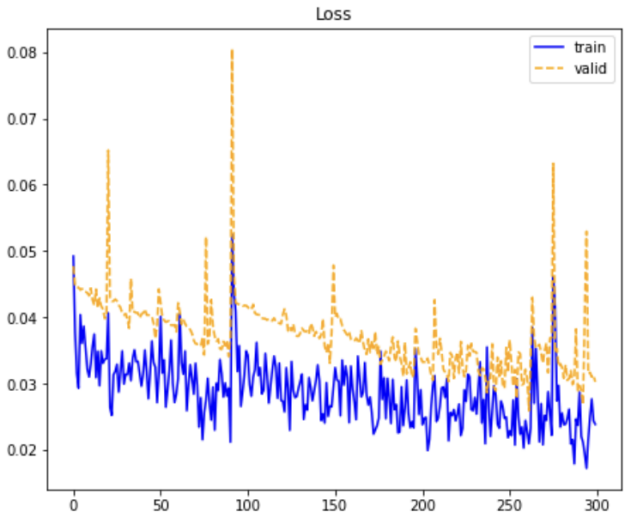


Figure 5. U-Net 300 Epoch

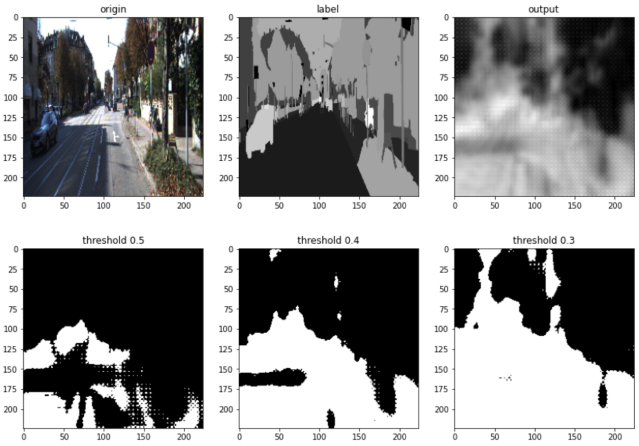


Figure 6. 300 Epoch Predict Images

5. Conclusion

In this study, we proposed a segmentation model to separate road areas using the KITTI dataset. A comparison of U-Net and U-Net++ models showed that the U-Net++ model exhibited better performance. Future research can further improve performance by diversifying data augmentation techniques and making additional improvements to the model.

6. Acknowledgment

This research was supported by the Modulabs Research Center. We thank the KITTI research team for providing the dataset.

References

[1] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The KITTI dataset," *International Journal of Robotics Research*, vol. 32, no. 11, pp. 1231–1237, 2013.

[2] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2015, pp. 234–241.

[3] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "Unet++: A nested u-net architecture for medical image segmentation," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, 2018, pp. 3–11.