NightRide: Semantic Segmentation of Motorcycle-Perspective Imagery During Nighttime Driving

Yena Choi

September 21, 2023

Abstract

This study addresses semantic segmentation challenges in images captured from a motorcycle driver's viewpoint during nighttime. The nighttime driving scenario presents its unique set of challenges and risks, making accurate image interpretation paramount. We introduce a model that categorizes each pixel into one of six predefined categories, namely undrivable areas, road, lane markings, the motorcycle itself (my bike), other riders, and movable objects. Initial results showcase the model's proficiency in classifying these categories, signifying potential advancements in nighttime motorcycle navigation safety tools. This segmentation model is poised to significantly bolster safety during nighttime motorcycle journeys by offering lucid visual understandings of road scenarios.

1 Introduction

This project was inspired by a Kaggle competition titled "Motorcycle Night Ride (Semantic Segmentation)". The dataset comprises 200 frames, each encompassing six classes, tailored for on-road object detection. The six classes are: undrivable, road, lane mark, my bike, rider, and movable. Our objective is their distinct identification. Numerous semantic segmentation projects of this nature exist, ranging from competition challenges to real-world applications. A prominent example is medical image analysis, which entails classifying different tissue types, organs, or tumors.

2 Method

2.1 Data Augmentation

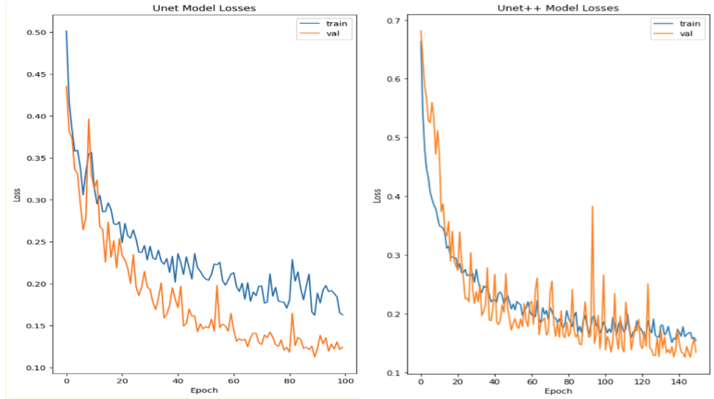
For data augmentation, we employed techniques like Horizontal Flip, Random Sized Crop, and Resize. Images were flipped horizontally with a 50% probability and randomly cropped at the same probability. All input images were resized to 224x224 dimensions. For the test dataset, the sole transformation was resizing to 224x224.

2.2 Train and Test Data Generators

After setting up the augmentation procedures, we constructed train and test data generators. The parameters used were: batch size=5, learning rate=1e-4, input shape=(224,224,3), optimizer='Adam', and loss function='binary crossentropy'.

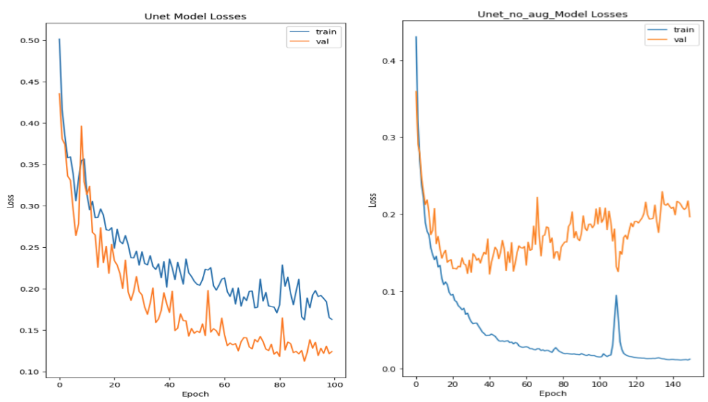
2.3 Model Selection

Given that our focus is image segmentation, we chose U-Net as our baseline model. For enhanced performance, U-Net++ was considered as an alternative due to its additional skip connections. We produced four models for evaluation: U-Net (100 epochs with augmentation), U-Net (150 epochs with augmentation), U-Net++ (150 epochs with augmentation), and U-Net (150 epochs without augmentation). The models' predictions were then assessed and contrasted using graphical illustrations. For a performance metric, we adopted Intersection over Union (IoU) to gauge the closeness of the predictions to the original labels.



(a) Loss graph of U-Net (b) Loss graph of U-Net++

Figure 1: Comparing loss graphs of two models



(a) Loss graph using augmentation (b) Loss graph without augmentation

Figure 2: Comparing loss graphs of two models

3 Results

Based on graphical interpretations and IoU scores, the U-Net++ model (150 epochs with augmentation) exhibited superior performance without any overfitting issues.

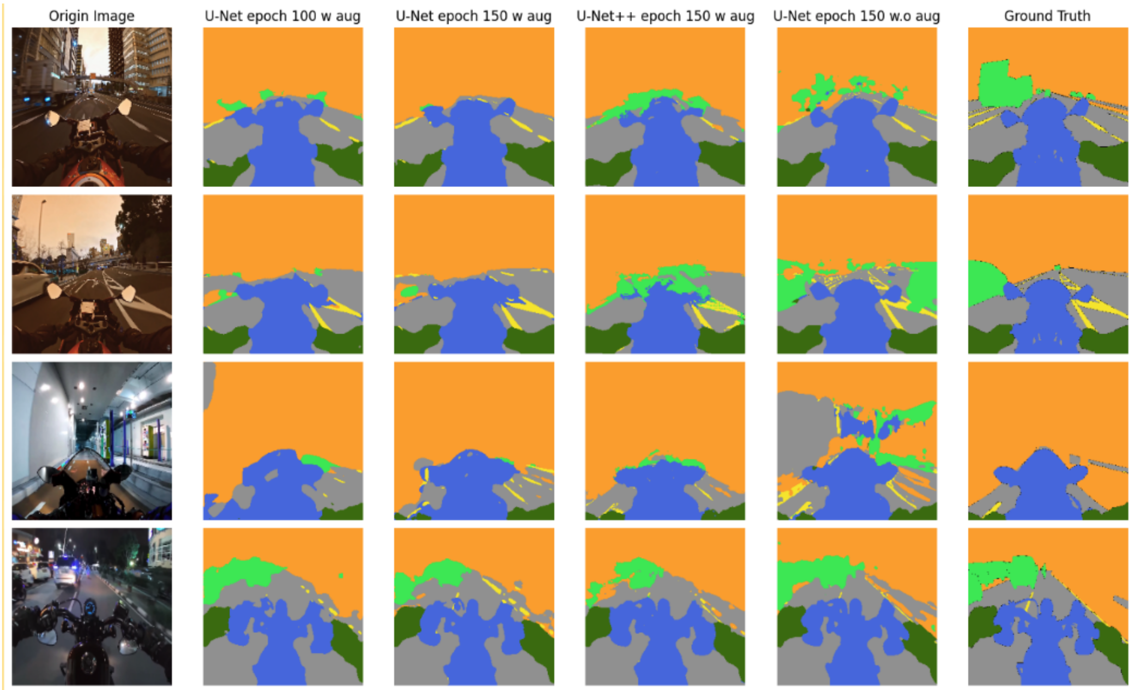


Figure 3: semantic segmentation

```
Index 0 of Batch
U-Net epoch 100 w aug IoUs
IoU of Rider : 0.8907    IoU of My bike : 0.8654    IoU of Moveable : 0.4382    IoU of Lane Mark : 0.0633    IoU of Road : 0.7994    IoU of Undrivable : 0.9617
Mean IoU : 0.6698    Pixel Accuracy : 0.8665

U-Net epoch 150 w aug IoUs
IoU of Rider : 0.9425    IoU of My bike : 0.9175    IoU of Moveable : 0.1107    IoU of Lane Mark : 0.2183    IoU of Road : 0.8877    IoU of Undrivable : 0.9472
Mean IoU : 0.6706    Pixel Accuracy : 0.8891

U-Net++ epoch 150 w aug IoUs
IoU of Rider : 0.8866    IoU of My bike : 0.8802    IoU of Moveable : 0.3265    IoU of Lane Mark : 0.0000    IoU of Road : 0.6797    IoU of Undrivable : 0.9538
Mean IoU : 0.6211    Pixel Accuracy : 0.8248

U-Net epoch 150 w.o aug IoUs
IoU of Rider : 0.9619    IoU of My bike : 0.9599    IoU of Moveable : 0.5083    IoU of Lane Mark : 0.2807    IoU of Road : 0.9057    IoU of Undrivable : 0.9560
Mean IoU : 0.7621    Pixel Accuracy : 0.9155
```

Figure 4: Intersection of Unions

4 Conclusion

Our findings indicate that among the models tested, U-Net++ particularly stands out in performance. While our models achieved respectable accuracy and IoU scores, there's potential for further enhancement using alternative architectures. Specifically, leveraging pre-trained models for training and experimentation might yield better outcomes. In future endeavors, we aspire to experiment with advanced segmentation models like DeepLabV3+.

5 References

Acme AI Ltd., Sadhli Roomy, Md. Mominul Islam, Abu Bakar Siddik Nayem, Ashik Mostofa Tonmoy, amp; Sikder Md. Saiful Islam. (2022). ¡i¿Motorcycle Night Ride (Semantic Segmentation)¡/i¿ [Data set]. Kaggle. https://doi.org/10.34740/KAGGLE/DSV/4208825