

# A model for object detection and semantic segmentation applicable to the CamVid dataset

\*Based on YoloV3 and Deeplabv3 architectures - COMP0248

Jiaqi Yao

**Abstract**—This paper presents a computer vision model based on YoloV3 and Deeplabv3 architectures for object detection and semantic segmentation tasks on the CamVid dataset. The model is optimized for the 5 classes present in the dataset, with special attention to addressing class imbalance issues. Experimental results demonstrate that our approach achieves [performance metrics] in terms of mean Average Precision (mAP) and Intersection over Union (IoU). Additionally, we explore the impact of various class imbalance handling techniques on model performance and discuss potential applications in robotic systems.

**Index Terms**—computer vision, object detection, semantic segmentation, YoloV3, Deeplabv3, CamVid, class imbalance

## I. Introduction

This section introduces the research background and motivation. It describes the importance of object detection and semantic segmentation in computer vision and their applications in robotics. The characteristics of the CamVid dataset and its value in urban scene understanding are explained. The main objectives and contributions of the research are clearly stated.

## II. Dataset

### A. CamVid Dataset Overview

This subsection describes the basic information about the CamVid dataset, including its source, scale, and purpose. It focuses on the 5 classes used for this research and their distribution within the dataset.

### B. Data Preprocessing

This subsection details the preprocessing steps applied to the original data, such as image resizing, normalization, data augmentation, etc.

### C. Data Visualization and Analysis

This subsection presents visualizations of the dataset, analyzes the imbalance in class distribution, and discusses the potential impact on model training.

## III. Methodology

### A. Model Architecture

This subsection provides a detailed description of the model architecture based on YoloV3 and Deeplabv3, including network structure diagrams, layer functions, and connection methods.

### B. Design Choices and Innovations

This subsection explains the key decisions in model design and specific optimizations made for the CamVid dataset.

### C. Training Strategy

This subsection introduces the training process, loss function selection, optimizer settings, and other details.

## IV. Experimental Setup

### A. Experimental Environment

This subsection describes the hardware and software environment used for the experiments.

### B. Hyperparameter Settings

This subsection lists in detail the hyperparameters used in the training process, such as learning rate, batch size, number of epochs, etc., and explains the reasons for these choices.

### C. Evaluation Metrics

This subsection explains the metrics used to evaluate model performance, such as mAP, IoU, etc., and the calculation methods for these metrics.

## V. Results

### A. Overall Performance

This subsection presents the overall performance of the model on the test set, including average precision and IoU scores.

### B. Class Analysis

This subsection analyzes the performance differences across classes, identifying which classes perform well and which perform poorly, and discussing possible reasons.

### C. Qualitative Results

This subsection presents some intuitive visual results, comparing model predictions with ground truth annotations.

## VI. Class Imbalance

### A. Class Imbalance Handling Techniques

This subsection introduces various techniques tested to address the class imbalance issues in the dataset, such as resampling, weighted loss functions, etc.

## B. Impact on Performance

This subsection analyzes the impact of various class imbalance handling techniques on overall model performance and performance across different classes.

## C. Technique Comparison

This subsection compares the advantages and disadvantages of different techniques and proposes the most suitable solution for this task.

# VII. Discussion

## A. Result Analysis

This subsection provides an in-depth discussion of the experimental results, analyzing the strengths and weaknesses of the model.

## B. Limitations

This subsection honestly points out the limitations of the research, such as dataset constraints, computational resource limitations, etc.

## C. Potential Applications

This subsection explores the potential uses of the model in practical application scenarios, especially in the field of robotics.

# VIII. Conclusion

This section summarizes the main findings and contributions of the research. It reflects on the lessons learned throughout the research process and proposes possible future research directions and improvements.

## Acknowledgment

Thanks to [relevant personnel/institutions] for their support of this research.

## References

- [1] [YoloV3 related citation]
- [2] [Deeplabv3 related citation]
- [3] [CamVid dataset related citation]
- [4] [Class imbalance problem related citation]
- [5] [Semantic segmentation evaluation method related citation]