Noah

CHENGDU UNIVERSITY OF TECHNOLOGY
OXFORD BROOKES UNIVERSITY

Application of Semantic Segmentation

Computer Science, Oxford Brookes University
Supervised by Dr Grace Ugochi Nneji

Abstract

This research designs a high-performance semantic segmentation framework to address the challenges of self-driving in complex urban environments. The framework combines a deeply separable convolution-based ASPP module, a Transformer module and an edge extraction module to optimise multi-scale information capture, global dependency processing and feature processing capabilities. Experimental results show that the model performs well on the Cityscapes dataset, achieving a MIOU of 0.77 and FPS of 63, demonstrating the accuracy and segmentation speed of the model.

Introduction

With the increasing urban traffic problems, self-driving cars have become one of the solutions[1]. However, in complex urban environments, self-driving faces many challenges, such as complex scene structures, variable weather conditions, and diverse traffic scenarios[2]. To deal with these challenges, this research designs and implements a high-performance semantic segmentation framework and achieves efficient and accurate segmentation results in the Cityscapes dataset. In addition, the project builds a user-friendly GUI to facilitate the popularisation and application of semantic segmentation techniques.

Semantic segmentation model

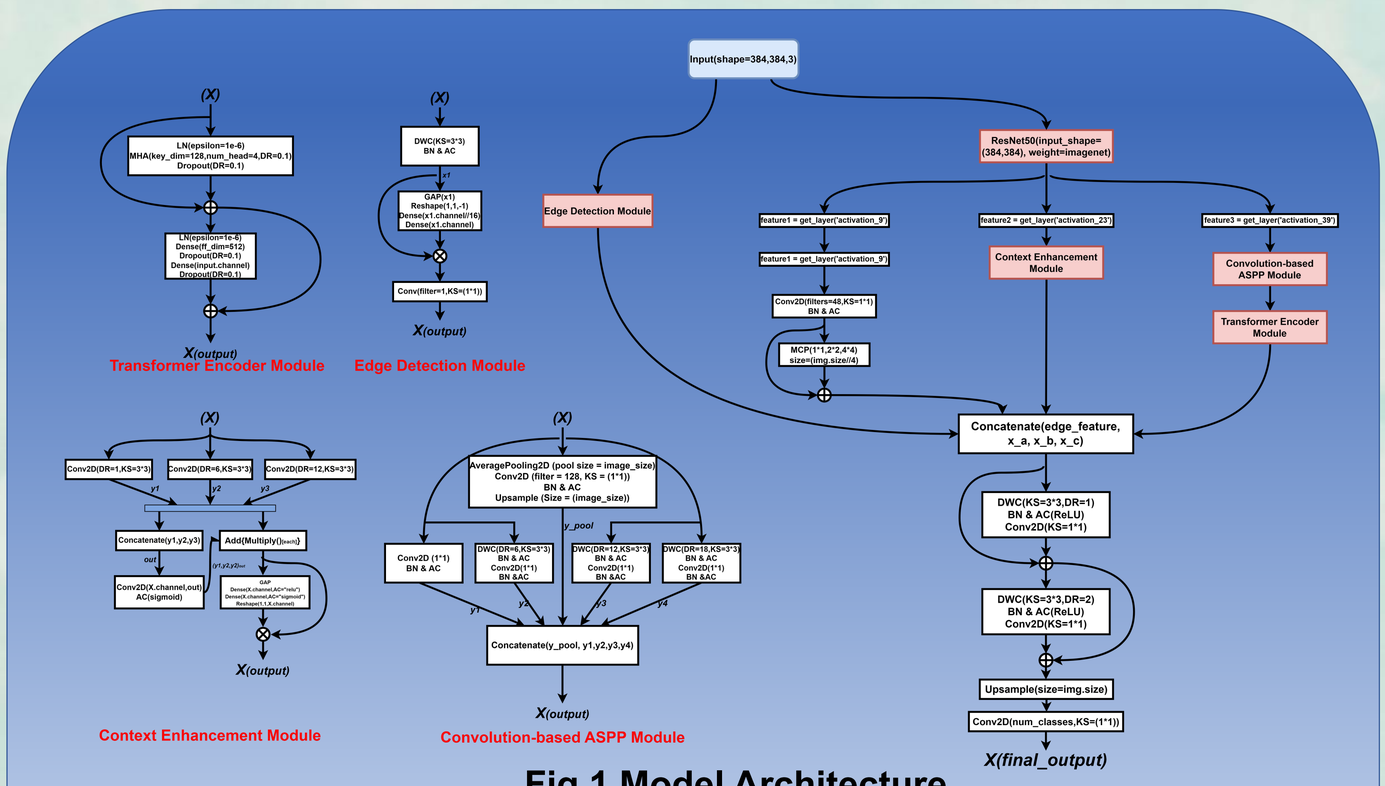


Fig.1 Model Architecture

- Transformer Encoder module: enhances the ability of the model to capture global dependencies.
- Edge Detection module: improve the accuracy of edge detection.
- Context Enhancement module: enhances the ability to process contextual information.
- Deep Separable Convolutional ASPP module: for capturing multi-scale context information.

Result & Evaluation

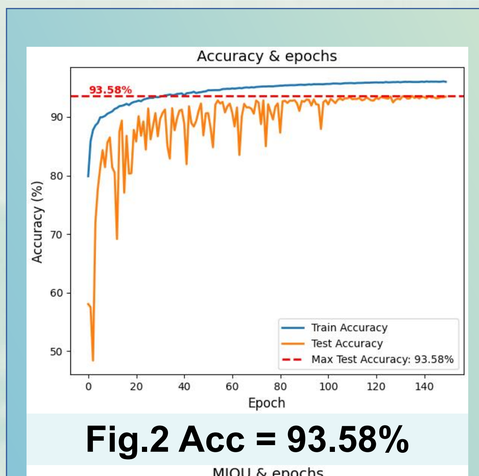


Fig.2 Acc = 93.58%

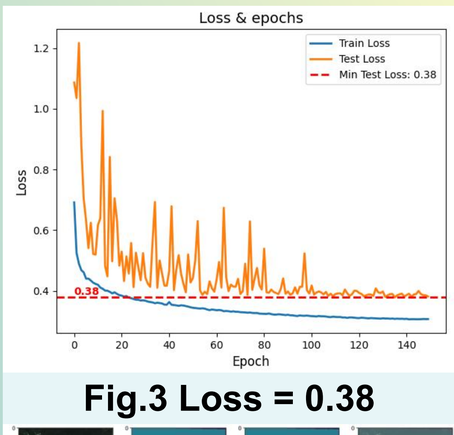


Fig.3 Loss = 0.38

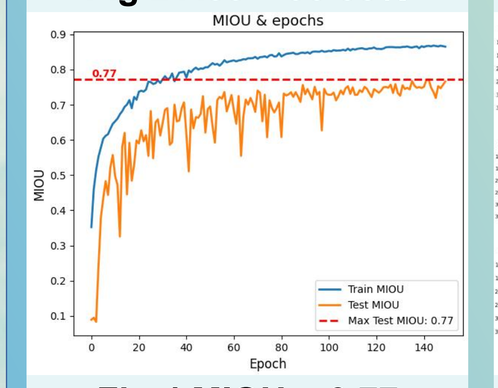


Fig.4 MIOU = 0.77

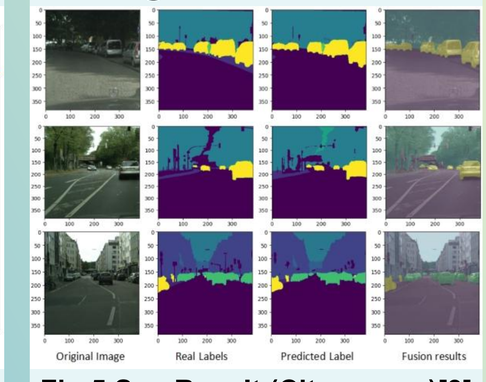


Fig.5 Seg Result (Cityscapes)[3]

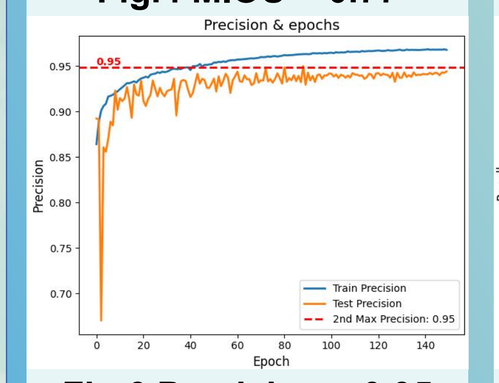


Fig.6 Precision = 0.95

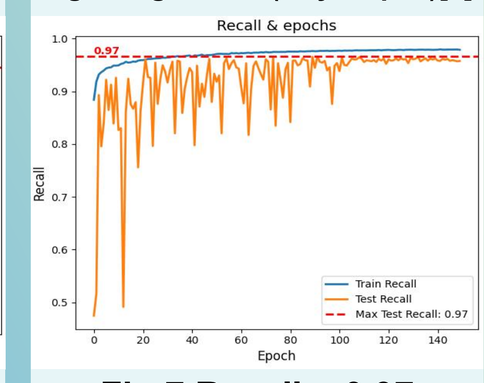


Fig.7 Recall = 0.97

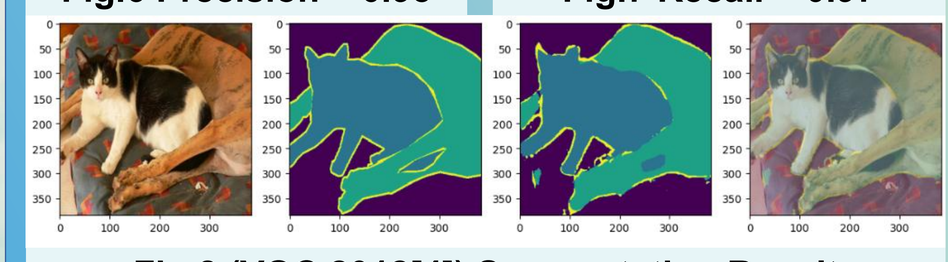


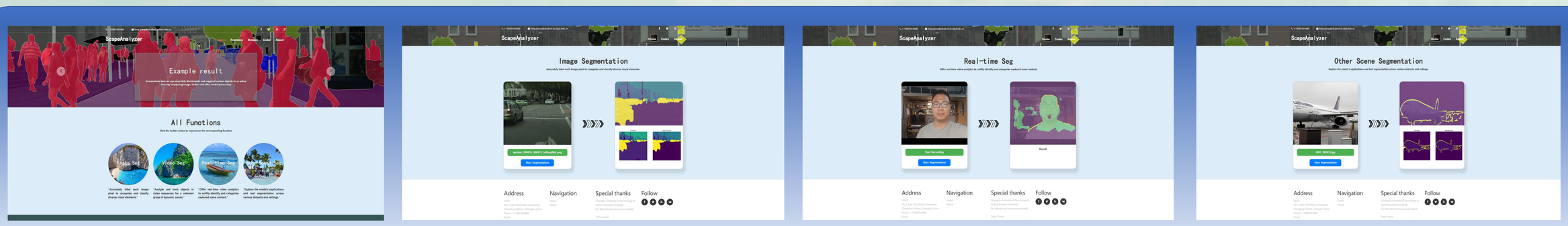
Fig.8 (VOC 2012[4]) Segmentation Result

PA	CPA	Dice Coefficient	Kappa	FPS
0.93	0.85	0.85	0.89	63.28

Tab.1 Other Metrics

In this project, I used a combined loss function of cross-entropy and Dice and trained for 150 epochs based on the SGD optimizer. By fine-tuning and setting other hyperparameters, the above results were finally achieved.

Project GUI



Future Work

- Investigate small target object detection techniques
- Explore lightweight model architectures
- Enhance the model's ability analyze dynamic scenes

Conclusion

This research proposes an efficient and accurate semantic segmentation framework and develops a user-friendly GUI interface to facilitate the popularisation of semantic segmentation techniques. This project shows the great potential of deep learning for semantic segmentation for self-driving.

References

[1] Duarte, F. & Ratti, C., 2018. The Impact of Autonomous Vehicles on Cities: A Review. Journal of Urban Technology, 25(4), pp.3-18. Available at: https://doi.org/10.1080/10630732.2018.1493883

[2] Janai, J. et al. (2020) 'Computer Vision for Autonomous Vehicles: Problems, Datasets and State of the Art', Foundations and Trends® in Computer Graphics and Vision, 12(1–3), pp. 1–308. Available at: https://doi.org/10.1561/0600000079.

[3] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele, "The Cityscapes Dataset for Semantic Urban Scene Understanding," in Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[4] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The PASCAL Visual Object Classes (VOC) Challenge," International Journal of Computer Vision, vol. 88, no. 2, pp. 303–338, 2010