

UNDERGRADUATE PROJECT PROGESS REPORT

|  |  |
| --- | --- |
| **Project Title:** | **Semantic Segmentation on Deep Learning Applications** |
| **Surname:** | **Feng** |
| **First Name:** | **Xiang** |
| **Student Number:** | **202018010119** |
| **Supervisor Name:** | **Dr Grace Ugochi Nneji** |
| **Module Code:** | **CHC 6096** |
| **Module Name:** | **Project** |
| **Date Submitted:** | **2023/12/25** |

**Table of Contents**

[1 Introduction 3](#_Toc32219)

[1.1 Background 3](#_Toc22719)

[1.2 Aim 3](#_Toc22599)

[1.3 Objectives 3](#_Toc27007)

[1.4 Project Overview 3](#_Toc1304)

[1.4.1 Scope 3](#_Toc2174)

[1.4.2 Audience 3](#_Toc28653)

[2 Background Review 3](#_Toc18474)

[3 Technical Progress 4](#_Toc10630)

[3.1 Approach 4](#_Toc17962)

[3.2 Technology 4](#_Toc29001)

[3.3 Testing and Evaluation Plan 4](#_Toc25401)

[3.4 Design and Implementation 4](#_Toc4139)

[4 Project Management 4](#_Toc11008)

[4.1 Activities 4](#_Toc14257)

[4.2 Schedule 4](#_Toc29568)

[4.3 Project Version Management 4](#_Toc28192)

[4.4 Project Data Management 4](#_Toc14602)

[4.5 Project Deliverables 5](#_Toc6429)

[5 Professional Issues and Risk: 5](#_Toc8931)

[5.1 Risk Analysis 5](#_Toc578)

[5.2 Professional Issues 5](#_Toc15868)

[6 References 5](#_Toc19837)

[6.1 Formatting Requirements 5](#_Toc9825)

[6.2 Written Presentation 5](#_Toc22315)

# **1Introduction**

## **1.1Background**

Today's cities are challenged by traffic congestion and frequent accidents, and autonomous vehicles are considered as a potential solution to reduce accident rates and improve traffic flow [1]. However, in urban environments, we need to deal with complex backgrounds, variable weather and various traffic scenarios, which are difficult for traditional computer vision techniques to handle [2].

Therefore, deep learning-based semantic segmentation techniques become especially critical, which can associate each pixel in scenes to semantic categories such as roads, buildings, vehicles, pedestrians, etc [3]. This provides a powerful environment perception tool for autonomous driving [4], helping vehicles better understand and adapt the complex road environment.

This project is devoted to the in-depth study and application of semantic segmentation techniques to create high-precision real-time models to promote safer and more efficient autonomous driving technologies for urban transportation systems, contributing to the sustainable development of cities.

## **1.2Aim**

Semantic segmentation can accurately classify each pixel in an image to improve the accuracy of environment perception, which has important applications in areas such as autonomous driving. This project aims to improve the model performance and generalization ability through integrated learning to achieve high-precision real-time road environment detection. This helps to achieve highly reliable autonomous driving and is expected to contribute to the development of urban transportation.

## **1.3Objectives**

* When the project begins, the first task will be to study the relevant literature in-depth to understand the latest advances and breakthroughs in the area of semantic segmentation. This will help to understand the performance and efficiency of various advanced neural network models in the task.
* Next, the focus is on selecting an appropriate semantic segmentation model. Considering the performance differences of different models in various scenarios, I plan to absorb and integrate the advantageous features of several mainstream model architectures, such as the lightweight design of LinkNet and the efficient parameter configuration of ResNet, to build a model with stronger generalization capabilities and superior performance.
* During the project, model performance will be monitored, including steps such as pre-training weight assignments, training strategy adjustments, and hyper-parameter fine-tuning to ensure that the synergy of the various modules in the model is maximized.
* Finally the model performance will be evaluated using appropriate test datasets and multiple evaluation metrics. To ensure that this semantic segmentation model meets the expected results and standards and that the model has high accuracy and real-time performance.

## **1.4Project Overview**

### **1.4.1Scope**

In semantic segmentation tasks, although traditional Convolutional Neural Networks (CNN) excel in feature extraction, common CNN models such as VGG, ResNet50, and InceptionV3 often fall short in terms of generalization capabilities and performance. To overcome this challenge, we have adopted an innovative composite model strategy. This approach involves integrating these common models into the semantic segmentation framework, fully exploiting the strengths and complementarity of different networks, thereby significantly enhancing the semantic segmentation effectiveness on datasets. Moreover, the core task of semantic segmentation models is to accurately identify various features in images, such as 'roads,' 'buildings,' 'vehicles,' and 'pedestrians.' Given the complexity of the task, integrating multiple models is crucial for enhancing the accuracy and stability of predictions.

The significances of this study are as follows:

Improving Autonomous Driving Safety

Improves driving efficiency

Promote the wide application of autonomous driving technology

Solving urban transportation problems

Promote the advancement of automated driving technology

Contribute to urban sustainable development

### **1.4.2Audience**

For government and city planners, this project enhances traffic management efficiency. Real-time road condition sensing in autonomous driving can optimize traffic flow and reduce congestion [1]. Autonomous driving also mitigates accidents caused by human factors, enhancing road safety. For drivers, it ensures safe, fatigue-free driving, and selects optimal routes based on real-time conditions, saving time, reducing stress, and improving travel efficiency [5]. For the urban environment, improved traffic management and reduced congestion cut emissions, enhancing city air quality.

# **2 Background Review**

Mohammed et al. [6] employed a double-branching strategy and double-attention block to improve accuracy and focus on image information, meanwhile reducing the number of parameters using the FDSS-nbt block, which achieved 74.9% mIoU. Rudra PK et al. [7] proposed new learned downsampling module that preserves spatial details and reduces computational burden to maintain 68% mIoU. Zhou et al. [8] model based on CSP encoder-decoder architecture reduces the computational cost, enhances feature extraction, and achieves 81% mIoU. Sun et al.[9] 's ResNet18-based multi-feature fusion network achieved 75% mIoU on the CityScapes dataset, proving the efficiency of multi-feature fusion.

Yu et al. [10] proposed BiSeNet model which uses a parallel structure as well as global and local information perception to improve accuracy, achieving 68.4% mIoU. Chaurasia et al. [11]: LinkNet model aims for fast semantic segmentation and uses an efficient decoder structure to reduce the computational burden, achieving 68.3% mIoU, proving that efficient semantic segmentation is feasible. Chen et al. [12] utilized DeepLab model to extend the sensory field by deep void convolution and achieves 65.88% mIoU, showing the strong performance of deep convolutional networks in urban scenarios on the urban dataset. Badrinarayanan et al. [13] suggested SegNet model which achieves semantic segmentation by using the encoding-decoding structure and achieves 65.2% mIoU, which is key to the information transfer and reconstruction mechanism. The summary of the different models proposed by different researchers can be seen in Table 1.

|  |  |  |  |
| --- | --- | --- | --- |
| **Researchers** | **Data Set** | **Techniques** | **Performance** |
| Mohammed et al.[6] | Cityscapes,CamVid | DSANet,FDSS-nbt,SIEN | mIoU=74.9% |
| Rudra et al. [7] | Cityscapes | Fast-SCNN,Coarse | mIoU=68% |
| Zhou et al. [8] | Cityscapes,CamVid | CSP en–decode -network | mIoU=81% |
| Sun et al. [9] | Cityscapes,CamVid | ResNet18,CMCA | mIoU=75.0% |
| Yu et al. [10] | Cityscapes,CamVid, COCO-Stuff | BiSeNet | mIoU=68.4% |
| Chaurasia et al. [11] | Cityscapes,CamVid | LinkNet | mIoU=68.3% |
| Chen et al. [12] | Cityscapes | Deep-Lab CRF | mIoU=65.88% |
| Badrinarayanan et al.[ [13] | CamVid | SegNet | mIoU=65.2% |

Table 1: Performance on different Semantic Segmentation models

# **3 Technical Progress**

## **3.1 Approach**

This project is aimed at creating a composite model with excellent performance by fusing traditional CNNs with advanced semantic segmentation architectures. The implementation steps are as follows:

Firstly, the dataset is subjected to fine-grained preprocessing, including but not limited to random cropping, flipping and noise injection, in order to improve the adaptability and robustness of the model.

ResNet50 is integrated into LinkNet as a core encoder to realize the dual advantages of feature extraction and structure lightening. This step aims to maximize the accuracy and efficiency of semantic segmentation through the composite model strategy.

Conduct an in-depth evaluation of the LinkNet model, including comprehensive performance analysis and comparison with industry benchmark models, to accurately locate its advantages and room for improvement.

Based on the evaluation results, the model structure is fine-tuned to optimize its learning mechanism and parameter configuration to ensure the best performance output and generalization capability.

### **3.1** **Dataset**

The Cityscapes dataset is a large dataset focusing on urban street scenes, widely used in computer vision and autonomous driving research. It contains high-resolution images from 50 different cities and provides accurate pixel-level annotations for about 5,000 images covering 30 different categories, such as roads, pedestrians, etc. The dataset also includes a number of other images that have been annotated to provide a better understanding of the city's streetscape. In addition, the dataset includes about 20,000 roughly annotated images, enriching the training data. In this project 2975 images will be used as training dataset, 500 as validation and 500 as testing.

**3.1.1 Data Preprocessing**

During the data preprocessing phase, I explored two class configurations for the CityScapes dataset: the original 33-category categorization, and a simplified 19-category categorization that was processed by the CityScapes script. For the 33-category categorized dataset, I implemented additional data enhancements such as random flipping and random cropping to ensure that the model can effectively handle a larger number of categories and avoid performance degradation due to the large number of categories. In addition, considering the high resolution of the images, I adopted the strategy of overlay cropping the labeled images with the training images at the same time, which prevents confusion caused by label inconsistency in the cropping process, and thus guarantees the consistency between the images and their corresponding labels. These preprocessing steps provide a more robust training foundation for the model and help improve its performance in complex scenes.

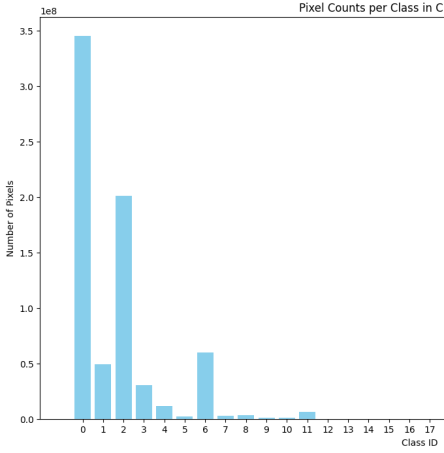
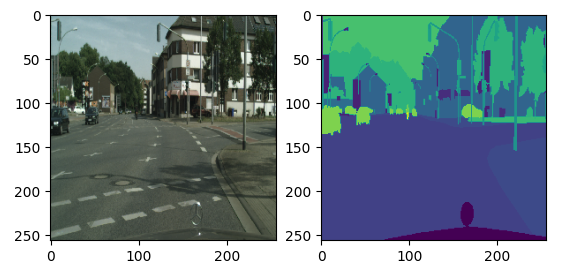


Figure 7

**3.1.2 Data Balancing**

When dealing with the semantic segmentation dataset CityScapes, a crucial step is to evaluate the balance of the different categories in the dataset, which is usually achieved by counting the number of pixels in each category. This process involves traversing the annotations of all images in the dataset to count the number of pixels in each category. Given the large volume of the CityScapes training dataset, a representative validation dataset was selected for this analysis for pixel counting. By generating graphs, we can visualize the quantitative relationships between different categories in the dataset. The analysis results reveal that in the CityScapes dataset, the number of samples for certain categories such as roads and buildings is relatively high, while the number of samples for specific vehicle or pedestrian categories is low.

In order to keep the data balanced in subsequent training, one may consider appropriately reducing the frequency of use of pixels that over-represent categories to achieve a more balanced data distribution.

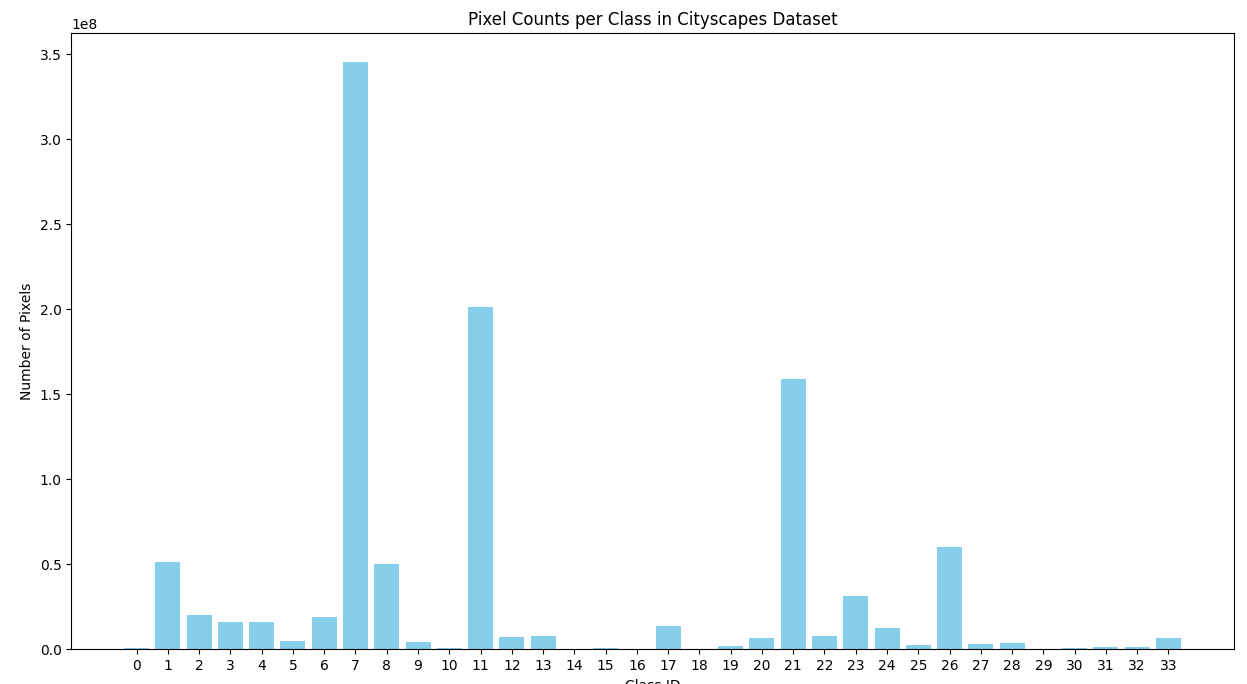
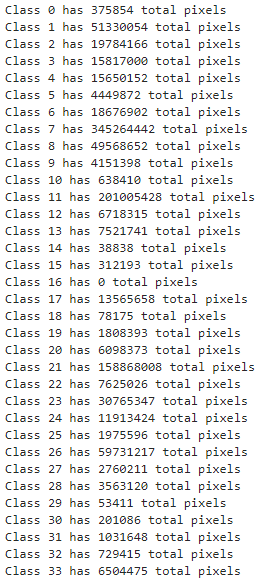


Figure 7

### **3.2** **Semantic Segmentation:**

Semantic segmentation task is to categorize each pixel in an image into predefined categories, thus providing detailed information at the pixel level [14]. Semantic segmentation helps in accurate recognition and understanding in areas such as autonomous driving and medical imaging [14], providing in-depth analysis of objects and areas in an image.

### **3.2.1 M** **(MIoU)**

MIoU [15] is mainly used to measure the degree of overlap between the segmentation results predicted by the model and the true results. In formula, 'k' represents the category number, 'Pii' represents number of overlapping pixels, and 'Pji' represents the number of misassigned pixels.' 1/(k+1)' is the average weight to ensure that each category contributes equally to the mIoU. Thus, mIoU is affected by the category number, the positive sample number, and the pixel overlap between different categories.

(equation 1)

### **3.2.2** **LinkNet**

As shown in Figure 3, LinkNet is a lightweight and efficient semantic segmentation neural network structure with an encoder-decoder architecture. It efficiently solves the gradient loss problem through special skip connections, a design that is important to maintain the integrity of the information stream. In this architecture, the encoder is responsible for capturing the core features of the image while gradually decreasing the resolution of the image, while the decoder is responsible for recovering the details of the image, fusing the feature information while increasing the resolution, and assigning the appropriate semantic category to each pixel. This design allows LinkNet to efficiently extract image features while maintaining image details, making it suitable for semantic segmentation tasks that require high accuracy and efficient processing [17].

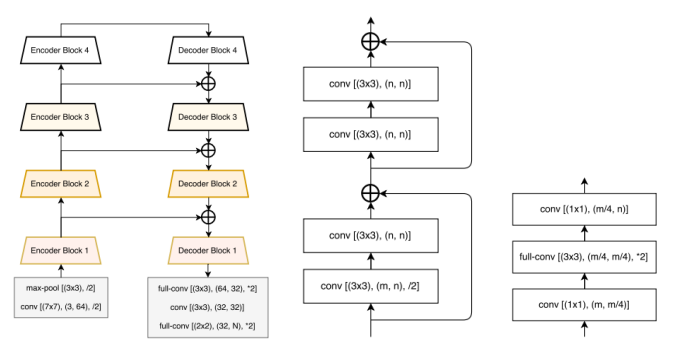


Figure 3 LinkNet structure diagram and encoder-decoder[16]

**3.2.3 ResNet50**

ResNet50 is a popular deep convolutional neural network belonging to the ResNet family with a 50-layer network structure. It solves the problem of gradient vanishing in deep networks through residual connectivity, ensuring effective information transfer and feature learning. This structure enables it to perform well in areas such as image classification, detection and segmentation, and is widely used in computer vision tasks.

### **3.2.4** **Proposed Model**

In this project, we propose to use the CityScapes dataset, as demonstrated in Figure 6. For the proposed model, I have carefully designed and adapted the encoder and intermediate layers. Specifically, for the encoder part, I used a pre-trained ResNet50 model from which I extracted the key layers to act as an encoder. The subsequent module is a specially designed tuning module that receives the output from the encoder. This module consists of a number of advanced components, including null convolution, CBAM attention mechanisms, and spatial pyramid pooling. This design allows the model to effectively recognize multi-scale features of an image while capturing expansive contextual information. In addition, to further strengthen the information extraction capability of the model, I also increase the number of encoders and decoders to ensure the model's efficiency and accuracy in processing complex visual data.

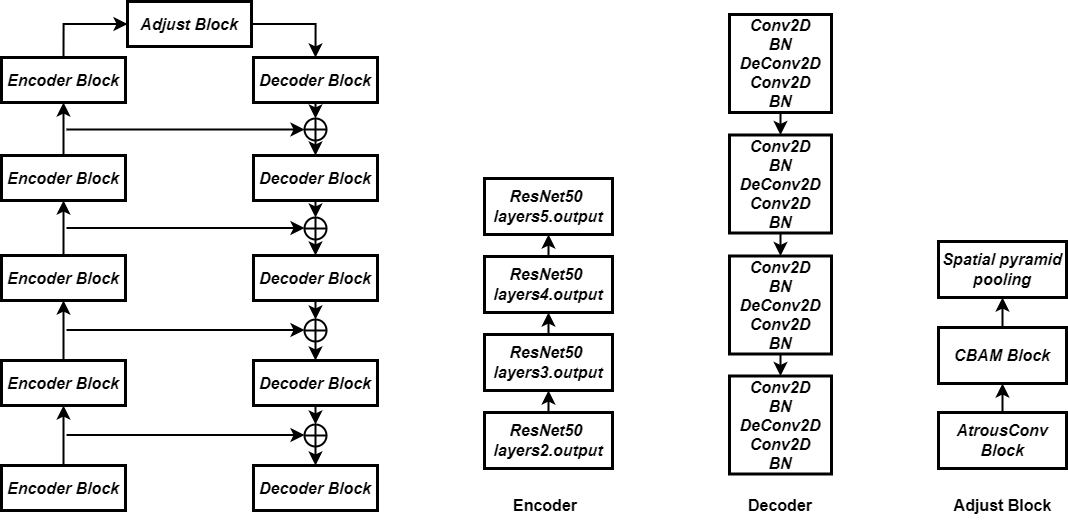


Figure 6 Proposed Model structure

**3.2.5 Spatial pyramid pooling**

Spatial Pyramid Pooling (SPP) is an innovative technique for enhancing the ability of convolutional neural networks to process input images of different sizes. By applying SPP after the convolutional layer of the network, the model is able to accept inputs of arbitrary size while maintaining consistency in output size. This property makes SPP particularly important in semantic segmentation and classification tasks, improving the flexibility and adaptability of the model.

## **3.3** **Technology**

The part information is shown in Table 2.

|  |  |  |
| --- | --- | --- |
| Software | Framework | Tensorflow 2.10.0  Cudatoolkit 11.8.0  Cudnn 8.9.2.26 |
| Language | Python 2.9 |
| Libraries | Numpy  Matplotlib  Pandas  Keras 2.10.0  Glob3 |
| Version management plan | GitHub |
| Operation System | Windows 10 |
| Hardware | Central processing unit(CPU) | AMD Ryzen 9 4900HS with Radeon Graphics 3.00GHz |
| Graphic Processing Unit(GPU) | NVIDIA GeForce RTX 2060 Max-Q |

Table 2 The technologies of the project

**3.4 Performance Evaluation Metrics**

**3.4.1 Loss**

For the model performance evaluation, I chose accuracy and mean intersection ratio (MIoU) as the key performance metrics. Accuracy measures the proportion of images correctly classified by the model, while MIoU measures the degree of overlap between the model segmentation results and the true labels. Accuracy can simply call the accuracy method provided by keras, but MIoU is not integrated, so I created a custom MIoU metric class to obtain the appropriate evaluation capabilities.

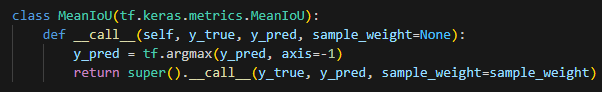


Figure 8

(1)

In addition, I also use Accuracy as an evaluation metric. In semantic segmentation, since the loss value (loss) does not directly reflect the actual performance of the model, especially on pixel-level categorization, I chose not to use the loss value as the main evaluation criterion. In contrast, Accuracy and MIoU, as more intuitive performance metrics, can more effectively evaluate and compare the effectiveness of semantic segmentation models.

**3.3.2.2 Pre-train Testing**

First, make sure that all layers and operations are configured correctly and that there are no missing or incorrect structures. Afterwards, confirm that the data loading and preprocessing steps are correct and appropriate for the input requirements of the model. Finally, evaluate the baseline performance of the model in its untrained state and check for obvious logic errors. This facilitates the early detection and correction of potential problems.

**3.3.2.3 Post-train Testing**

In this part can try to rotate, scale, and color adjust the input data to check whether the performance of the model is stable or not. And confirm whether the model can achieve a respectable recognition accuracy on the largest category.

## **3.4 Design and Implementation**

In this project, I used a composite model combining LinkNet and ResNet50. In order to enhance the feature extraction capability of the model in the encoder part, I used a pre-trained ResNet50 to build the encoder of the model. After a series of fine-tuning and appropriate adjustments to the model architecture, I obtained the following results:

This is the result of training the LinkNet semantic segmentation model with 100 batches trained, this dataset is currently 34 categories and obtained a validation accuracy of 0.8113 and an MIoU of 0.330, the lower accuracy is probable due to the dataset having too many categories:

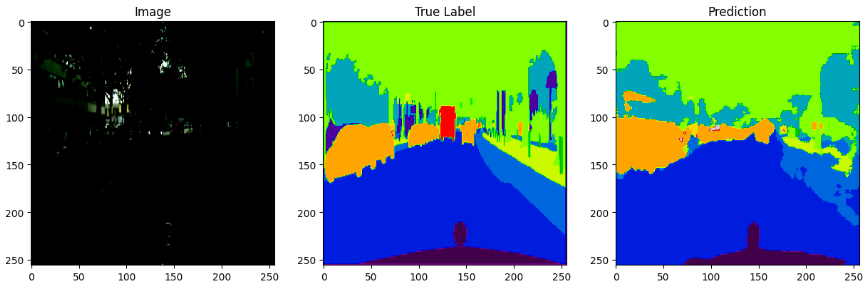




Figure 9

Below are the results of training 100 batches of a hybrid ResNet50 and LinkNet model based on the CBAM attention mechanism and spatial pyramid pooling, to ensure that the model is not overfitted, I added the operations of early stopping and learning rate decay, and ultimately obtained a MIoU of 0.520 and an accuracy of 0.8:

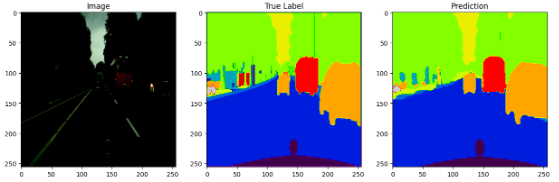




Figure 10

In order to better compare the effects of different semantic segmentation models, I used the UNet model for this semantic segmentation dataset, and obtained a MIoU of 0.429 as well as an accuracy of 0.9 after training for 60 epochs:

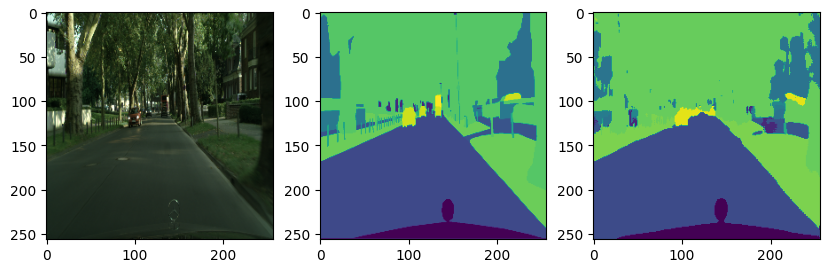


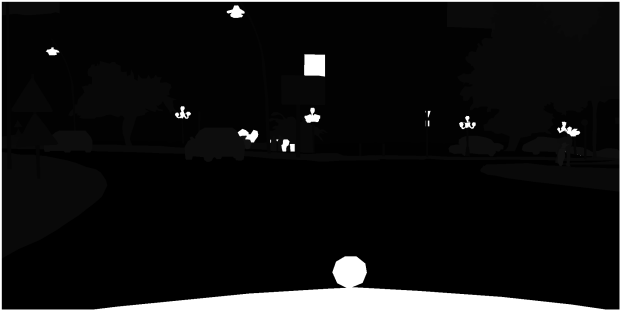


Figure 11

Therefore, in summary, the default LinkNet is not as good as UNet in terms of training results on the cityscapes dataset with 34 classifications, but after changing the encoder layer in the LinkNet model, the model acquires higher MIoU. Moreover, the number of LinkNet with integrated ResNet encoder is less in the training parameters, which means that this LinkNet is more suitable for real-time semantic segmentation projects with high accuracy.

After modifying the dataset to 19 categories, the left side is the original image with 34 categories, and the right side is the generated image with 19 categories, which may be due to improper operations during data preprocessing, resulting in the inability to detect the corresponding categories, and due to the inability to detect the corresponding categories, the loss of null was generated during training, and the specific reason is still being investigated, and the initial thought is that the color mapping in the grayscale map was not The specific reason is still under investigation. The details are shown in the figure below:





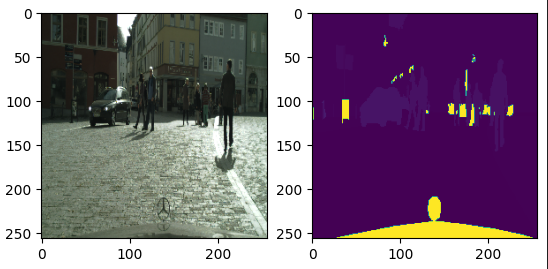


Figure 12

# **4 Project Management**

## **4.1 Activities**

|  |  |
| --- | --- |
| Tasks | State |
| A. study real-time semantic segmentation and keep abreast of the latest advances in the field. | Complete |
| B. Compare and analyze different semantic segmentation models. | Complete |
| C. Study hybrid learning in semantic segmentation, focusing on sub-model parameter assignment and hybrid strategies. | Complete |
| D.Explore multiple datasets relevant to the project and identify the required datasets. | Complete |
| E.Preprocess the selected datasets including feature transformation, data enhancement and segmentation. | Complete |
| F.Understand and construct a model based on the structure of the literature. | Complete |
| G.Based on the implementation results, select the best performing and well-structured model as the backbone network for integrated learning in this project. | Complete |
| H. Fine-tune the hyperparameters of the selected model, compare the performance with the original model, and analyze the differences. | Complete |
| I. Confirm the integration strategy, weight assignment, and integrate the adjusted model. | Complete |
| J. Record model test results from multiple training phases, fine-tune hyperparameters, and compare performance with the original model. | Complete |
| K. Analyze experimental results, measure model accuracy, and compare to the original single model. | Uncompleted |
| L. Complete an experimental report, including project background, research methodology, results, and conclusions. | Uncompleted |
| M. Design an image interface that can be interacted with by the user to facilitate more intuitive observation of the results | Uncompleted |
| N. Create a PowerPoint presentation based on the report to summarize the course. | Uncompleted |

**Table 3 Activities Check Table**

**4.2 Schedule**

The project schedule designed with Gantt can be seen in Figure 7

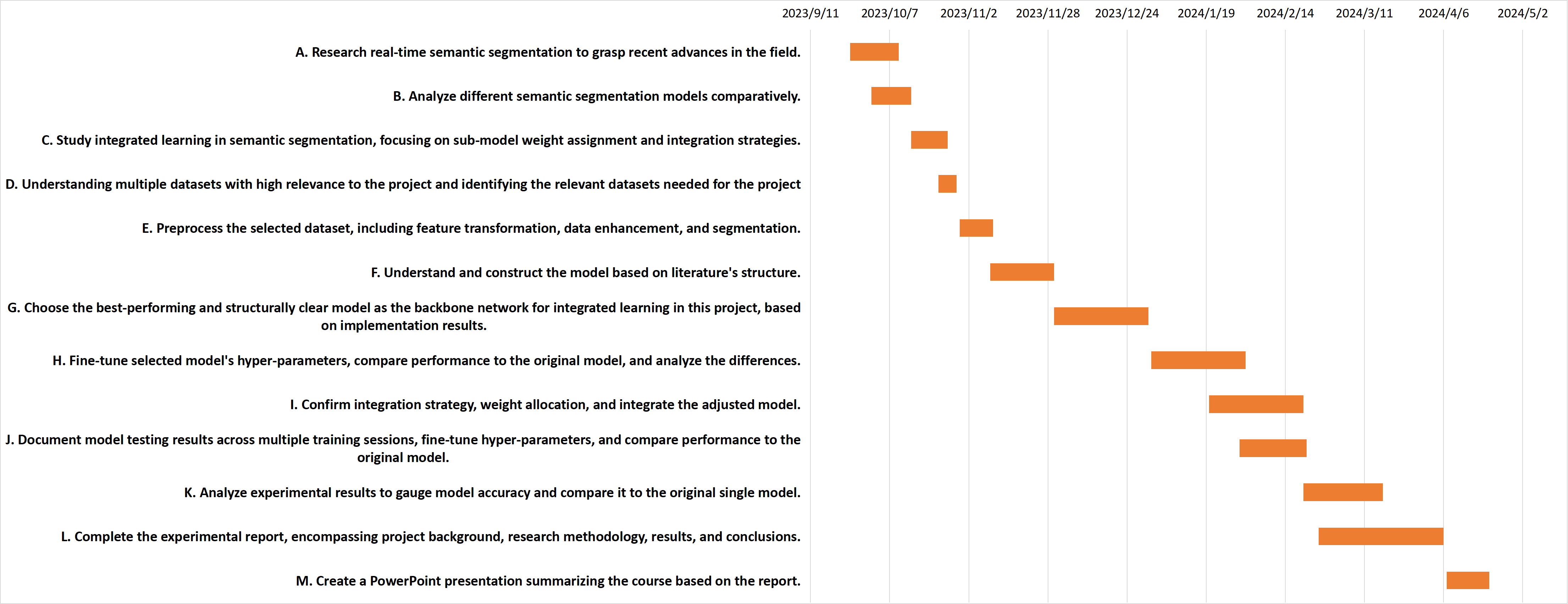


Figure 13 The project schedule.

## **4.3Project Version Management**

This project will use GitHub as the primary version control tool for regular commits and updates of project data. Also, a local and remote backup strategy is used, including code, documentation, and every version to ensure data redundancy and availability and to easily track project development.

|  |  |
| --- | --- |
| Version 1 | Implement different semantic segmentation models and record the results and predict the consumption time. |
| Version 2 | Implementing the LinkNet semantic segmentation model. |
| Version 3 | Some more advanced concepts such as null convolution, spatial pyramid pooling, and depth-separable convolution are added to the model, and the effects are recorded. |
| Version 4 | Implemented a class change for the model, converting 34 classes to 19 classes. |

Table 4 Version Management

## **4.4Project Data Management**

This project will use github for data management, this includes figures, code, logs.Github will upload weekly project results as well as updating weekly project progress and also, ensuring backup of data. It will create local folders and record or manage the relevant data for easy local recall. Use Zotero to manage references for literature that is being used.

## **4.5 Project Deliverables**

1. The project proposal
2. Weekly report
3. Progress Report
4. Final Project Report
5. Project codes
6. Project presentation’s ppt
7. Github links for project management

**5 Professional Issues and Risk.**

**5.1 Risk Analysis**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Potential Risk | Potential Causes | Severity | Likelihood | Risk | Mitigation |
| Dataset imbalance | Skewed data distribution,  Insufficient samples | 3 | 3 | 9 | Balance dataset,  Increase sample size |
| Model overfitting | Excessive training, Limited dataset variety | 4 | 3 | 12 | Regularize models, Use cross-validation techniques |
| Technology implementation complexity | Deep learning complexity, Modular design absence | 3 | 3 | 9 | Implement modular design, Conduct thorough testing |
| Real-time processing challenges | Hardware limitations, Software inefficiencies | 4 | 3 | 12 | Optimize for real-time, Upgrade hardware and software |
| Complex urban environments for self-driving vehicles | Variable weather, Traffic scenarios | 4 | 3 | 12 | Develop adaptive algorithms, Enhance traffic data analysis |
| Miss the deadline | Inadequate project timeline, Unexpected delays | 4 | 3 | 12 | Develop a detailed project plan, Include buffer times in schedule, Regular progress tracking |

Table 5 Risk Analysis

**5.2 Professional Issues**

**5.2.1 Identification and Discussion of Legal, Social, Ethical and Environmental Issues**

In the context of the project, we identify and discuss relevant legal, social, ethical and environmental issues. For example, the development of self-driving technology may give rise to legal challenges in terms of privacy and security, as well as potential impacts on the job market.

**5.2.2 Adherence to professional codes of conduct**

I refer to the BCS (British Computer Society) and ACM (Association for Computing Machinery) codes of professional conduct to ensure that the project complies with industry standards. This includes ensuring fair use of data, maintaining user privacy and ensuring fairness and accessibility of technical solutions.

**5.2.3 Environmental Responsibility**

Given that the goal of the project is to contribute to the sustainable development of the urban transportation system, we place special emphasis on environmental responsibility. By improving the accuracy and efficiency of automated driving technologies, we expect to reduce traffic congestion and emissions, and contribute to the quality of the urban environment.

# **6. 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. Hao, S., Zhou, Y. and Guo, Y. (2020) ‘A Brief Survey on Semantic Segmentation with Deep Learning’, *Neurocomputing*, 406, pp. 302–321. Available at: https://doi.org/10.1016/j.neucom.2019.11.118.
4. Li, J. *et al.* (2021) ‘Lane-DeepLab: Lane semantic segmentation in automatic driving scenarios for high-definition maps’, *Neurocomputing*, 465, pp. 15–25. Available at: https://doi.org/10.1016/j.neucom.2021.08.105.
5. Martínez-Díaz, M. and Soriguera, F. (2018) ‘Autonomous vehicles: theoretical and practical challenges’, XIII Conference on Transport Engineering, CIT2018, 33, pp. 275–282. Available at: https://doi.org/10.1016/j.trpro.2018.10.103.
6. Elhassan, M.A.M. et al. (2021) ‘DSANet: Dilated spatial attention for real-time semantic segmentation in urban street scenes’, Expert Systems with Applications, 183, p. 115090. Available at: https://doi.org/10.1016/j.eswa.2021.115090.
7. Poudel, R.P.K., Liwicki, S. and Cipolla, R. (2019) ‘Fast-SCNN: Fast Semantic Segmentation Network’. arXiv. Available at: http://arxiv.org/abs/1902.04502 (Accessed: 23 October 2023).
8. Zhou, L. et al. (2023) ‘Real-time semantic segmentation in traffic scene using Cross Stage Partial-based encoder–decoder network’, Engineering Applications of Artificial Intelligence, 126, p. 106901. Available at: https://doi.org/10.1016/j.engappai.2023.106901.
9. Sun, J. and Li, Y. (2021) ‘Multi-feature fusion network for road scene semantic segmentation’, Computers & Electrical Engineering, 92, p. 107155. Available at: https://doi.org/10.1016/j.compeleceng.2021.107155.
10. Yu, C. et al. (2018) ‘BiSeNet: Bilateral Segmentation Network for Real-Time Semantic Segmentation’, in V. Ferrari et al. (eds) Computer Vision – ECCV 2018. Cham: Springer International Publishing (Lecture Notes in Computer Science), pp. 334–349. Available at: https://doi.org/10.1007/978-3-030-01261-8\_20.
11. Chaurasia, A. and Culurciello, E. (2017) ‘LinkNet: Exploiting encoder representations for efficient semantic segmentation’, in 2017 IEEE Visual Communications and Image Processing (VCIP). 2017 IEEE Visual Communications and Image Processing (VCIP), St. Petersburg, FL: IEEE, pp. 1–4. Available at: https://doi.org/10.1109/VCIP.2017.8305148.
12. Chen, L.-C. et al. (2018) ‘DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs’, IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(4), pp. 834–848. Available at: https://doi.org/10.1109/TPAMI.2017.2699184.
13. Badrinarayanan, V., Kendall, A. and Cipolla, R. (2017) ‘SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation’, IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(12), pp. 2481–2495. Available at: https://doi.org/10.1109/TPAMI.2016.2644615.
14. Lateef, F. and Ruichek, Y. (2019) ‘Survey on semantic segmentation using deep learning techniques’, Neurocomputing, 338, pp. 321–348. Available at: https://doi.org/10.1016/j.neucom.2019.02.003.
15. Garcia-Garcia, A. et al. (2017) ‘A Review on Deep Learning Techniques Applied to Semantic Segmentation’. arXiv. Available at: http://arxiv.org/abs/1704.06857 (Accessed: 24 October 2023).
16. Chaurasia, A. and Culurciello, E. (2017) ‘LinkNet: Exploiting Encoder Representations for Efficient Semantic Segmentation’, in 2017 IEEE Visual Communications and Image Processing (VCIP), pp. 1–4. Available at: https://doi.org/10.1109/VCIP.2017.8305148.
17. K. Wu and F. Cai (2022) ‘Dual Attention D-LinkNet for Road Segmentation in Remote Sensing Images’, 2022 IEEE 14th International Conference on Advanced Infocomm Technology (ICAIT), pp. 304–307. Available at: https://doi.org/10.1109/ICAIT56197.2022.9862683.
18. Ronneberger, O., Fischer, P. and Brox, T. (2015) ‘U-Net: Convolutional Networks for Biomedical Image Segmentation’. arXiv. Available at: http://arxiv.org/abs/1505.04597 (Accessed: 25 October 2023).
19. 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.