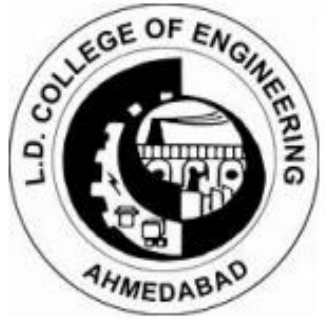
GUJARAT TECHNOLOGICAL UNIVERSITY
L. D. COLLEGE OF ENGINEERING, AHMEDABAD



"Semantic Segmentation for Autonomous Vehicles"

Dissertation Phase - 1

M.E. Sem-3

Computer Engineering Department

Prepared by:
Mohit Parmar
220280702013

Internal Guide:
Prof. Maitrik K. Shah
Assistant Professor

Outline

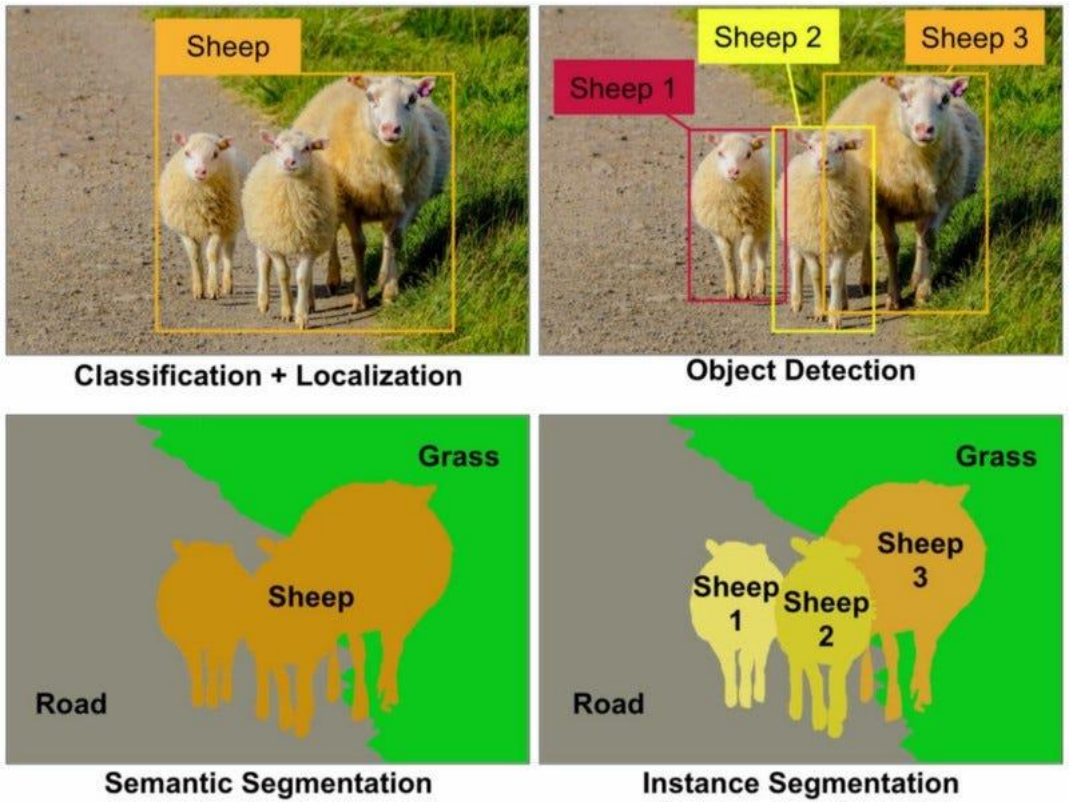
- Introduction
- Motivation
- Literature Review
- Research Gap
- Research Objective
- Proposed Methodology
- Dataset Details
- Conclusion
- References

Introduction

- Significant progress has been made in Autonomous driving. Since it offers high potential to decrease traffic congestion, improve road safety, and reduce carbon emissions.
- To make vehicle autonomous, the vehicle needs to perceive all relevant traffic participants and objects accurately, robustly, and in real-time. For that, autonomous cars are usually equipped with multi-modal sensors (e.g. Cameras, LiDARs, Radars).
- Now these cameras take real time pictures of the surrounding and perform semantic segmentation models to understand the scene.

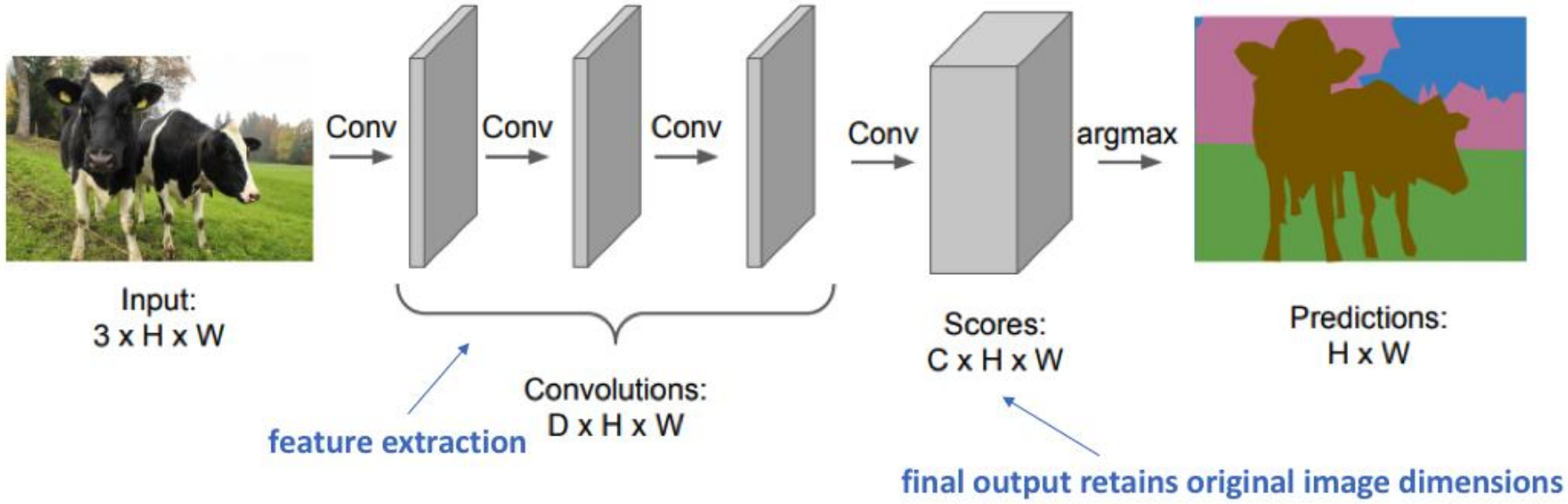
Semantic Segmentation

Semantic segmentation is the process of assigning each pixel of the received image into one of the predefined classes. These classes represent the segment labels of the image, e.g., roads, cars, signs, traffic lights, or pedestrians



Difference between Object Detection Semantic segmentation and instance segmentation[13]

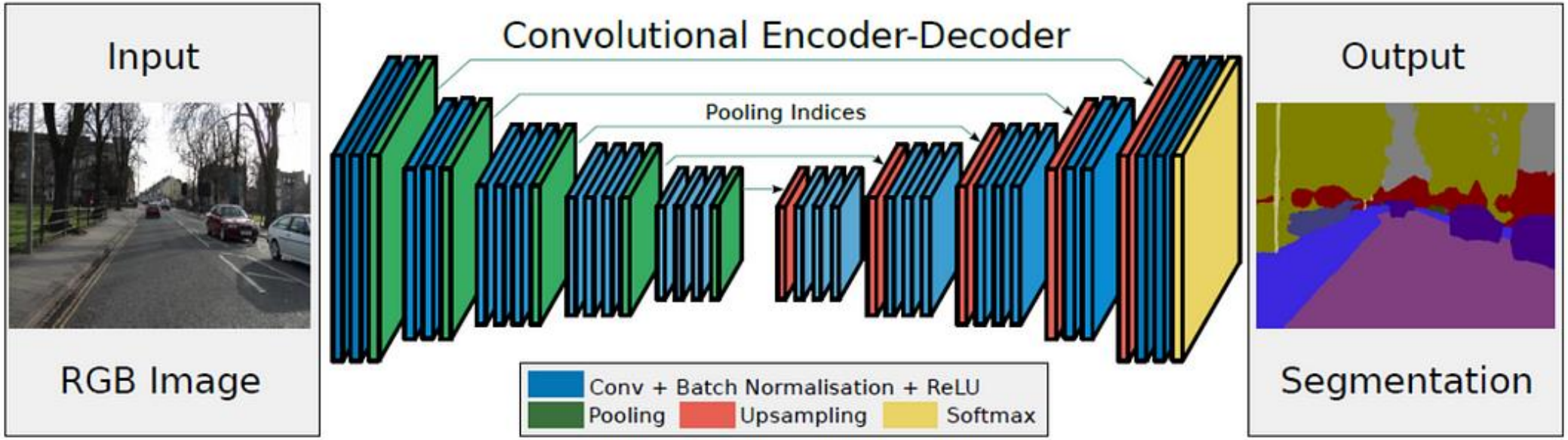
Semantic Segmentation

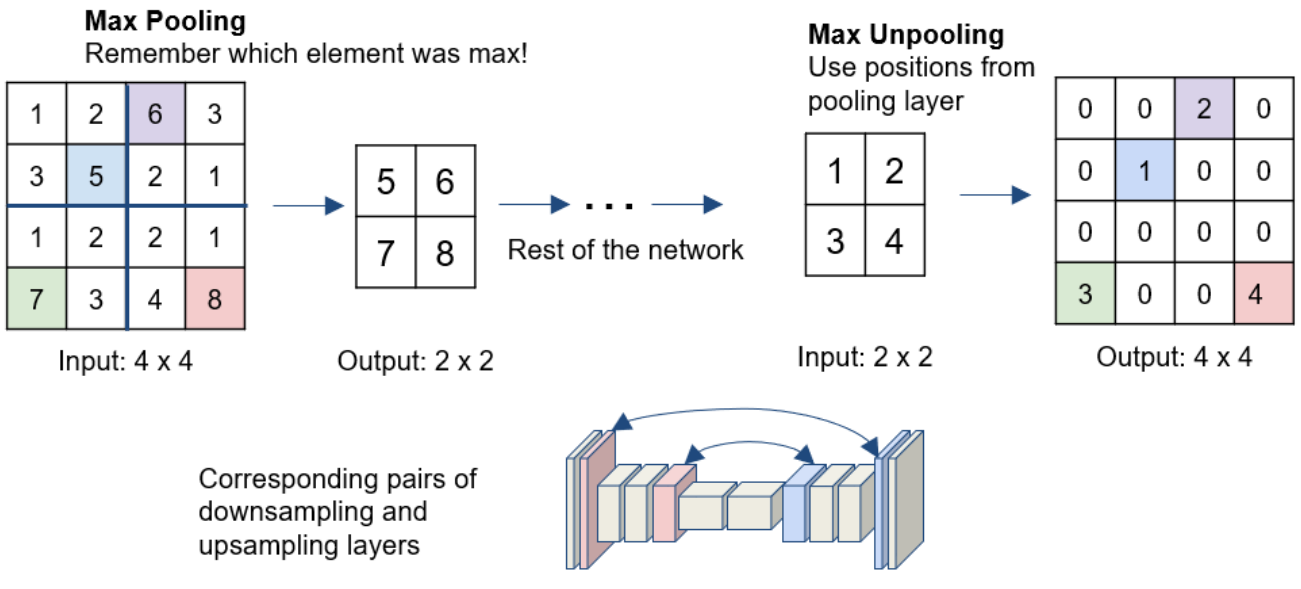


Convolution Layers for segmentation[14]

Motivation

- **Enhancing Safety**: Enhanced Semantic segmentation can help to avoid obstacles, pedestrians and potential hazards
- **Efficiency and Traffic Flow**: Vehicles can make decisions faster and hence navigate smoothly
- **Economic and Environmental Benefits**: Improved navigation help in fuel saving and reduction in emissions will reduce air pollution.
- **Urbanization and Traffic Congestion**: Rapid increase in population requires efficient management of traffic in the city.





SegNet Architecture[2]

Performance Metric

- Different We will be examining IoU to determine the accuracy of the model.
- The IoU for a single object or class is calculated as the ratio of the area of overlap between the predicted and ground truth masks to the area of union:

$$IoU = \frac{Area\ of\ Overlap}{Area\ of\ Union}$$

- The mean IoU is then computed by averaging the IoU values across all classes or objects. The formula for mean IoU is summation of each class's IoU over the number of class

Literature Review

Sr. No.	Paper Title	Publication	Methods / Results	Research Gap
1.	SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation	**Journal**: IEEE Transactions on Pattern Analysis and Machine Intelligence **Author**: Vijay Badrinarayanan, Alex Kendall, Roberto Cipolla **Year**: 2015	• 13 convolutional layers from VGG16 used. • The decoder network uses pooling indices computed in the max-pooling step of the corresponding encoder for up-sampling. • Accuracy**:** 50.02% mean IoU	• Accuracy can be improved • SegNet model can be further optimized to address overfitting

Literature Review

Sr. No.	Paper Title	Publication	Approach Used	Research Gap
2.	Real-time object detection and semantic segmentation for autonomous driving	**Journal:** Proceedings of the SPIE **Author**: Baojun Li Shun Liu Weichao Xu Wei Qiu **Year**: 2018	• DeepLab-v2 model is used (Dialated Convolutions) • Accuracy**:** 80.82% mean IoU	• There only worked on two parameters which are road and car. • Additional loss functions or training strategies can be implied to enhance object representation

Literature Review

Sr. No.	Paper Title	Publication	Approach Used	Research Gap
3.	Deep Road Scene Understanding	**Journal**: IEEE Signal Processing Letters **Author**: Wujie Zhou, Sijia Lv, Qiuping Jiang, Lu Yu **Year**: 2019	• Proposed deep encoder-decoder architecture for road scene understanding based on VGG16 • At the end of architecture concatenation layer is added • Accuracy**:** 56.05% mean IoU	• Exploration of other deep architectures can improve accuracy • Prediction of depth information is not being done

Literature Review

Sr. No.	Paper Title	Publication	Approach Used	Research Gap
4.	Real Time Image and Video Semantic Segmentation For Self-Driving Cars	**Journal**: Journal of Survey in Fisheries Sciences **Author**: Dr Moorthy A, Sivashanmugam B, Sriram R, Swathi M **Year**: 2023	• Deep learning algorithms, such as YOLOv4 and COCO dataset • Offline simulations and on-road trials are conducted to evaluate the effectiveness of the proposed systems • Accuracy: 85.5%	• Multiple algorithms can be tested on the same dataset for better results

Literature Review

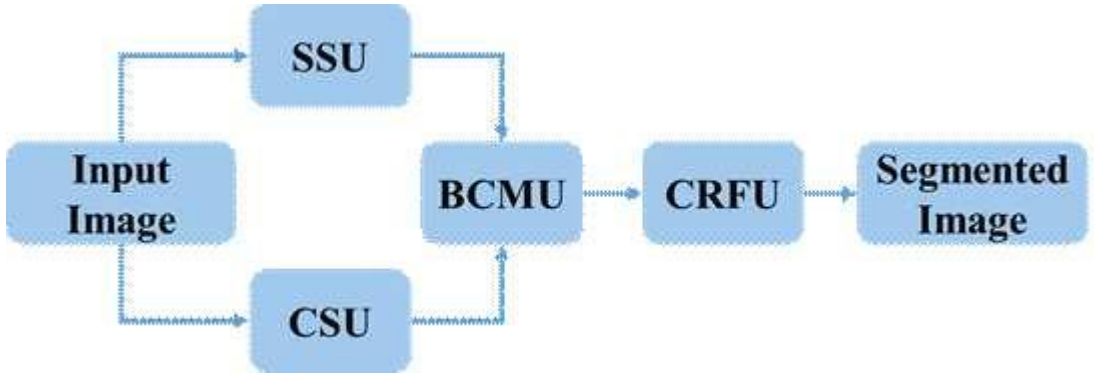
Sr. No.	Paper Title	Publication	Approach Used	Research Gap
5.	Intelligent Semantic Segmentation for Self-Driving Vehicles Using Deep Learning	**Journal**: Computational Intelligence and Neuroscience **Author**: Qusay Sellat, Sukant, Kishoro Bisoy, Rojalina Priyadarshini, Ankit Vidyarthi , Sandeep Kautish, Rabindra K. Barik **Year**: 2022	• Deep learning architectures: convolutional neural networks • Autoencoders and compared with baseline models • Accuracy**:** 58.275% mean IoU	• Number of parameters are much more than other models • mIoU can be improved by considering some other approaches

Literature Review

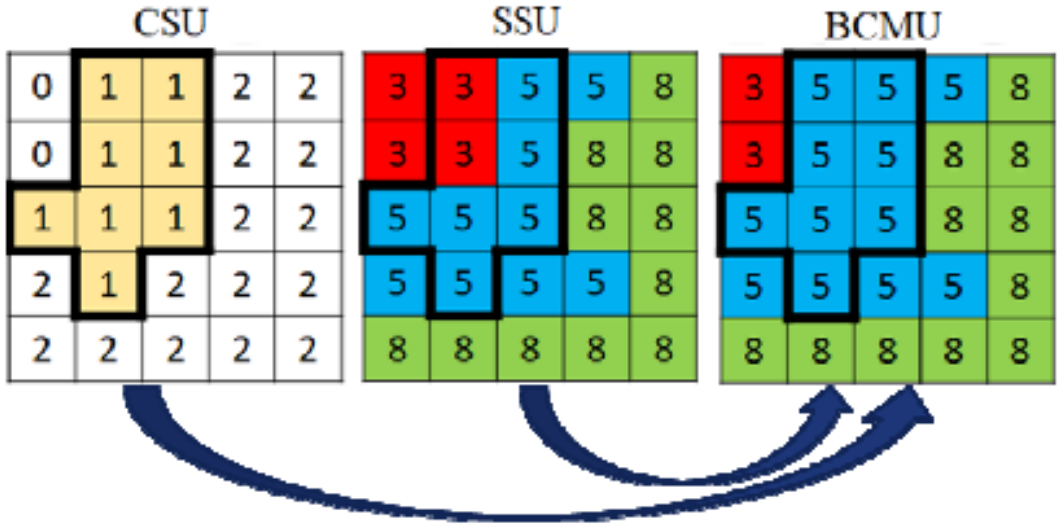
Sr. No.	Paper Title	Publication	Approach Used	Research Gap
6.	Image Segmentation for Self-Driving Car	**Journal**: International Conference on Intelligent Technologies (CONIT) **Author**: Sanchit Gautam, Tarosh Mathuria, Shweta Meena **Year**: 2022	• U-Net model for semantic segmentation in self-driving cars. • Color clustering using K-means was used to find essential colors in the image and represent classes. • Accuracy**:** 71.27%	• Comparison with different architectures like VGG and ResNet can improve the accuracy

Literature Review

Sr. No.	Paper Title	Publication	Approach Used	Research Gap
7.	Semantic Segmentation of Autonomous Driving Images by the Combination of Deep Learning and Classical Segmentation	**Journal**: International Computer Conference, Computer Society of Iran **Author**: M. Hosein Hamian A. Beikmohammadi A. Ahmadi B. Nasersharif **Year**: 2021	• Semantic segmentation unit (SSU) i.e. DeepLab v3+ • Classical segmentation unit (CSU) i.e. K Means • Boundary correction and merge unit(BCMU) • Conditional random field unit (CRFU) • Accuracy**:** 81.73% for Xception, 76.31% for MobileNetV2.	• Complexity is much more since it uses segmentation methods as well as post processing steps. • Other deep model should be explored as a backbone to improve the accuracy



Architecture of the proposed model consisting of four units[7]



Semantic segmentation unit (SSU), Classical segmentation unit (CSU),[7] Boundary correction and merge unit(BCMU)

- A conditional random field is a discriminative statistical modelling method that is used when the class labels for different inputs are not independent.

Literature Review

Sr. No.	Paper Title	Publication	Approach Used	Research Gap
8.	A Comparative Study of Real-Time Semantic Segmentation for Autonomous Driving	**Journal**: IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops **Author**: M. Siam, M. Gamal, M. Abdel-Razek, S. Yogamani, M. Jagersand **Year**: 2018	• SkipNet and U-Net meta-architecture comparison • Dilation Frontend meta-architecture • Accuracy: 70.1%(Skip Net)	• Computational efficiency and scalability of efficient semantic segmentation algorithms can be improved • Limited exploration of mathematical formalization

Encoder	Decoder	GFLOPs	mIoU	Road	Sidewalk	Building	Sign	Sky	Person	Car
SkipNet	MobileNet	13.8	61.3	**95.9**	73.6	**86.9**	57.6	91.2	66.4	**89.0**
SkipNet	ShuffleNet	**4.63**	55.5	94.8	68.6	83.9	50.5	88.6	60.8	86.5
UNet	ResNet18	43.9	57.9	95.8	73.2	85.8	57.5	91.0	66.0	88.6
UNet	MobileNet	55.9	61.0	95.2	71.3	86.8	**60.9**	**92.8**	**68.1**	88.8
UNet	ShuffleNet	17.9	57.0	95.1	69.5	83.7	54.3	89.0	61.7	87.8

Research Gap

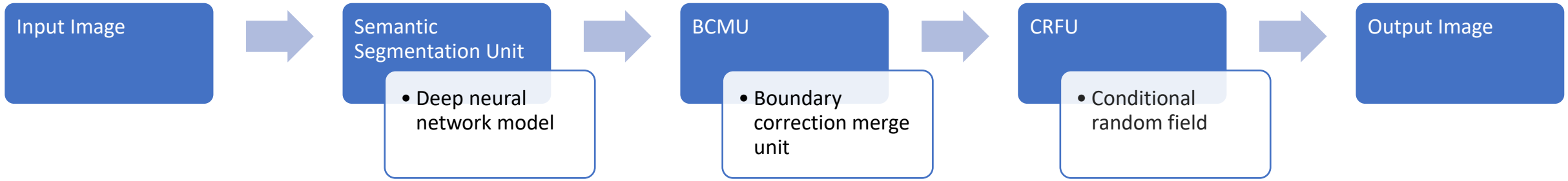
- Models have low mean IoU (Intersection over Union) which is around 70% currently.
- In VGG16, Prediction of depth information of an image is not considered.
- More parameters can be considered for the evaluation of DeepLab v2 model.
- Scalability issue with the comparatively smaller dataset(CamVid).

Research Objective

- Improve the accuracy of semantic segmentation by examining various deep neural network and their procedures.

Methodology

- Exploring other deep neural network as a backbone of DeepLab V3+ and post processing units (i.e. BCMU and CRFU) in the model and Comparing its accuracy with other backbones
- Use the Cityscapes dataset for training and testing.



Dataset

The proposed dataset is the CityScapes dataset for our model. It can be found on their website:

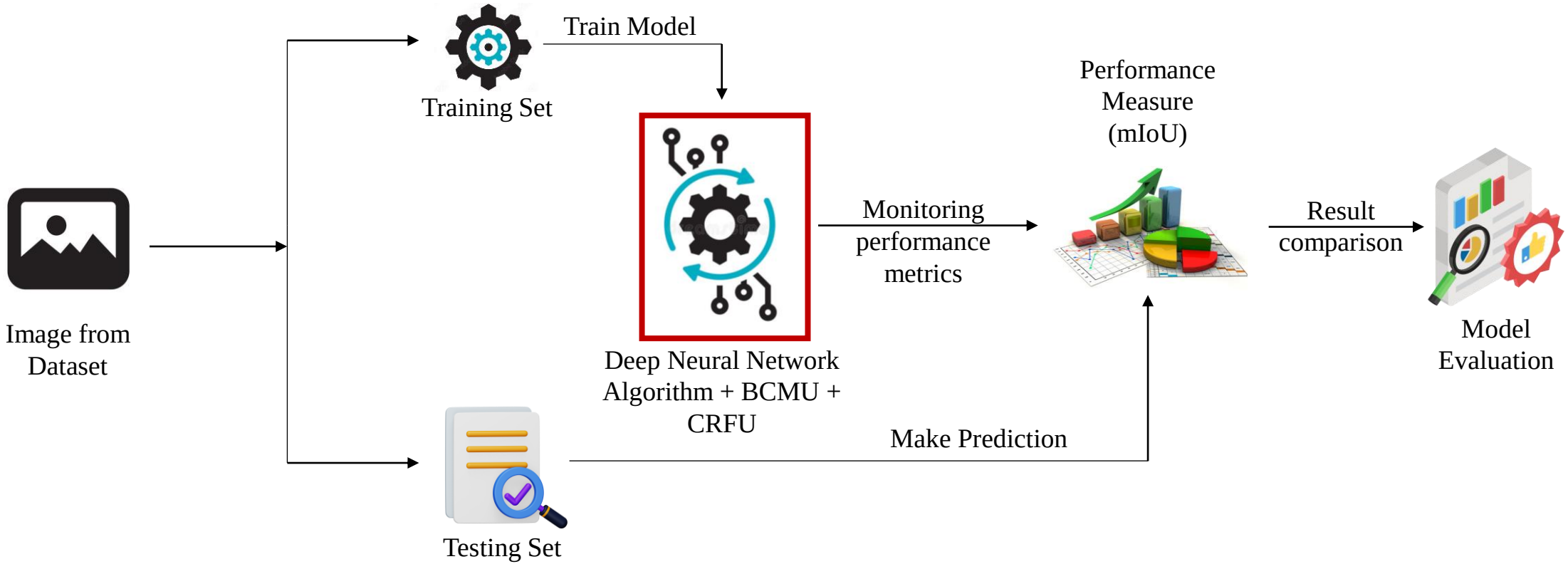
This dataset contains dashboard images and their segmented halves of a driving car in Germany.

Dataset consists of 5000 annotated images 1024px X 2048px with fine annotations. 2975, 500, and 1525 images are considered for training, validation, and testing respectively

https://www.cityscapes-dataset.com/.

Example of annotated image in Cityscapes[11]

Work Flow


Train Model
Training Set
Performance Measure (mIoU)
Image from Dataset
Deep Neural Network Algorithm + BCMU + CRFU
Monitoring performance metrics
Result comparison
Model Evaluation
Make Prediction
Testing Set

Conclusion

Semantic segmentation is helpful in understanding the surrounding scene for vehicles but it has to be modified in a way that gives good accuracy.

On the basis of Literature survey, some post processing steps and a relevant large dataset might do a better job for classifying classes. Evaluation of the segmentation models and their encoder decoder network architecture is done by the help of mIoU. Accurate detection of the classes can help in traffic congestion as well as Environmental benefits.

References

1. Vijay Badrinarayanan, Alex Kendall, Roberto Cipolla ”SegNet Architecture for semantic segmentation”, arXiv,2015
2. Baojun Li, Shun Liu, Weichao Xu, Wei Qiu “Real-time object detection and semantic segmentation for autonomous driving”, SPIE, 2018
3. Wujie Zhou,Sijia Lv, Qiuping Jiang, Lu Yu “Deep Road Scene Understanding”, IEEE Signal Processing Letters, 2019
4. Dr Moorthy A, Sivashanmugam B, Sriram R, Swathi M “Real Time Image and Video Semantic Segmentation For Self-Driving Cars”, Journal of Survey in Fisheries Sciences, 2023
5. Qusay Sellat, Sukant, Kishoro Bisoy, Rojalina Priyadarshini, Ankit Vidyarthi , Sandeep Kautish,Rabindra K. Barik “Intelligent Semantic Segmentation for Self-Driving Vehicles Using Deep Learning”, Hindawi, 2022
6. Sanchit Gautam, Tarosh Mathuria, Shweta Meena “Image Segmentation for Self-Driving Car”, International Conference on Intelligent Technologies (CONIT), 2022
7. M. Hosein Hamian ,A. Beikmohammadi, A. Ahmadi, B. Nasersharif “Semantic Segmentation of Autonomous Driving Images by the Combination of Deep Learning and Classical Segmentation”, International Computer Conference, Computer Society of Iran, 2021
8. M. Siam, M. Gamal, M. Abdel-Razek, S. Yogamani, M. Jagersand, “A Comparative Study of Real-Time Semantic Segmentation for Autonomous Driving”, IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2018

References

9. Jingdong Wang, Ke Sun, Tianheng Cheng, Borui Jiang, Chaorui Deng, Yang Zhao, Dong Liu, Yadong Mu, Mingkui Tan, Xinggang Wang, Wenyu Liu, and Bin Xiao, “Deep High-Resolution Representation Learning for Visual Recognition”, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, arXiv, 2019
10. Cityscapes Dataset Images Dataset[dataset], www.cityscapes-dataset.com, accesed November 2023, Cityscapes Dataset
11. CamVid Dataset Images Dataset[dataset], ,accesed October 2023, CamVid Dataset

12. Image Difference between image-classification-vs-semantic-segmentation-vs-instance-segmentation
13. Image Convolution Neural Network

Thank You