CAPSTONE PROJECT REPORT

Title

Road Hazard Detection Using Semantic Segmentation

Submitted By:-

Atharva Ravindra Khadge

B.Tech 2nd Year Student,

Department of Computer Science and Engineering,

G. H. Raisoni College of Engineering, Nagpur.

Certificate Program on Machine Learning 2024, IIIT Hyderabad.

HUB ID : HUB20240160

Email: atharva.khadge.cse@ghrce.raisoni.net

INDEX

ABSTRACT

As intelligent transportation systems and autonomous vehicles race ahead with rapid development, reliable and real-time road hazard detection has become ever more crucial. Road hazards like potholes, roadworks, dropped debris, and other etiolated obstacles can have a bad influence on the safety of driving, particularly in dynamic and highly populated settings like India. Conventional methods of object detection tend to fail in detecting such hazards accurately because of different road textures and illumination levels.

This project suggests a deep learning method with semantic segmentation to identify and locate road hazards from images of driving scenes. We use models like DeepLabV3 , which are highly accurate for pixel-wise classification tasks. We use the Indian Driving Dataset (IDD) for training and testing because it captures the peculiarities of Indian road conditions such as unstructured settings and rich road features.

Our approach is to preprocess the IDD dataset, train the model for segmentation to discriminate between safe road regions and potential dangers, and measure its performance in terms of measures such as Intersection over Union and Mean Pixel Accuracy. The outcome is that the model is able to segment key aspects in a scene such as road edges, cars, pedestrians, and obstacles. This not only supports safer autonomous driving but also offers a useful basis for the integration of hazard detection into driver-assistance systems.

The project illustrates how semantic segmentation, with the right datasets and architectures, can be beneficially applied towards real-world road safety solutions. Future improvements would include real-time deployment on edge devices and greater generalization towards other geographical regions.

INTRODUCTION

Road safety continues to be a serious problem, especially in areas with high-intensity traffic growth and uneven infrastructure. Traffic hazards like potholes, roadworks, road debris, and inadequate lane markings results in a large proportion of traffic accidents. It is not only time-consuming but also unreliable to manually detect these hazards, particularly in high-density traffic conditions. As the automotive sector continues to head toward automation and intelligent transportation systems, efficient and real-time detection of road hazards is becoming ever more critical.

Semantic segmentation, a deep learning method for pixel-wise image classification, has been promising for applications that involve extensive scene understanding. It differs from conventional object detection approaches because it gives a holistic understanding of the surroundings, making it a suitable choice for autonomous driving applications.

This project aims to create a system based on semantic segmentation for road hazard detection in real-world driving conditions. The Indian Driving Dataset (IDD) with its variability and unstructured nature of roads is utilized for training and testing the model. The dataset comprises images of Indian urban and rural roads with multiple challenges like non-uniform traffic patterns, changing lighting conditions, and sudden obstacles.

We use deep learning architectures like DeepLabV3 , which are capable of capturing both global context and subtle spatial details. These models are trained to detect and segment various components of the road scene, such as safe driving zones and potential hazards.

The objective of this project is to improve road safety through the deployment of a stable and scalable solution for hazard detection. This system can be used as a core building block in autonomous navigation or driver-assistance systems, particularly in areas where driving conditions are complex.

OBJECTIVES

The primary intention of this project is to identify road scenes with the help of deep learning and segment various objects in images in the Indian Driving Dataset (IDD). The clear objectives are:

The principal aim of this project is to employ deep learning in perceiving road scenes through the segmentation of various objects from images in the Indian Driving Dataset (IDD). The specific objectives are:

1. Learn About the IDD Dataset

Recognize the manner in which the IDD dataset is structured and the way it tags various sections of Indian roads, particularly the nitty-gritty level-3 tags.

2. Create Training Masks

Use tools provided by AutoNUE to generate accurate masks for each image, so the model knows what to learn during training.

3. Use DeepLabV3+ for Segmentation

Train the DeepLabV3+ model to recognize and separate different objects like roads, cars, people, and more in road images.

4. Check Model Accuracy

Measure how well the model is working using a metric called mean Intersection over Union (mIoU), at high resolution (720p).

5. Test on Real Indian Roads

Watch how well the model performs on actual, chaotic road scenes such as those in India, and work to make it even better.

6. Assist with Self-Driving Technology

Contribute to making transportation smarter by assisting cars in comprehending road scenes more effectively through high-resolution image analysis.

METHODOLOGY

This project takes a step-by-step process of constructing an efficient semantic segmentation framework for identifying road hazards from real-world images using deep learning approaches and the Indian Driving Dataset (IDD). The methodology includes the following major steps:

1. Data Preprocessing

The Indian Driving Dataset (IDD) is utilized because it covers a large variety of urban and rural road scenes in different conditions. The dataset consists of images and their respective semantic labels. The labels are utilized to recognize and categorize objects such as potholes, debris, cars, pedestrians, road markings, and so on. Preprocessing involves resizing images, pixel value normalization, and data augmentation (flipping, rotation, contrast changes) to enhance the capability of the model to generalize between different situations.

2. Model Selection

Popular models are selected for semantic segmentation such as DeepLabV3. DeepLabV3 is renowned for its convolution and spatial pyramid pooling, which aid in the capture of multi-scale features. This models is first pre-trained on generic datasets and then fine-tuned on the IDD dataset to suit the unique nature of Indian roads.

3. Training

The models are trained under supervised learning, with labeled images utilized to learn pixel predictions class-wise. To enhance performance on less common hazard classes or missing labels, weakly-supervised methods (such as label smoothing or pseudo-labeling) can also be employed. This enables the model to learn stronger features even from partially labeled data.

4. Evaluation

The models are tested with Intersection over Union (IoU) scores, which is a common measure for segmentation tasks. IoU is computed for every class of road hazard, providing a clear indication of how well each hazard type is being detected.

5. Post-Processing

To improve the model predictions, Conditional Random Fields (CRFs) are used as a post-processing technique. CRFs correct boundary segmentation by accounting for pixel interaction and spatial consistency, enhancing visual quality and accuracy of the hazard region detections.

IMPLEMENTATION

To carry out semantic segmentation with the DeepLabV3+ model on the IDD-20K dataset, we went step by step. Every step was crucial to assist the model in learning and doing well on tough Indian road scenes. Here is a straightforward description of how we did it:

1. Setting Up the Environment

We used **Windows 11** with **Python 3.10.2** to build the project.
To keep everything organized, we created a **virtual environment** using Python's tool.

After activating the environment, we installed the necessary Python libraries:

- numpy
- pandas==1.2.1
- tqdm
- Pillow
- scipy==1.1.0
- imageio

These libraries helped with data handling, image processing, and training the model.

2. Preparing the Dataset

We used the **IDD-20K dataset**, which is part of the AutoNUE Challenge 2021. It contains more than 20,000 images of Indian road scenes.

We worked with **Level-3 annotations**, which provide detailed labeling for 26 different classes like road, car, pedestrian, etc.

We combined **Part I** and **Part II** of the dataset into a single folder for easier access.

3. Creating Segmentation Labels

The dataset comes with annotations in **JSON format**, so we needed to convert them into **segmentation mask images** (PNG format).

- We downloaded the official AutoNUE tools from GitHub.

- We used the createLabels.py script to convert the JSON files into mask images.
- Each pixel in the output image had a value from **0 to 25**, representing one of the 26 classes.

4. Using DeepLabV3+ Model

We used DeepLabV3+ since it's excellent in semantic segmentation.

Here's what it does:

• Backbone: We employed a pre-trained ResNet-101 to extract the features from images.

• Encoder: It employs ASPP (Atrous Spatial Pyramid Pooling) to acquire features at diverse scales.

• Decoder: It returns the details to acquire precise object shapes and borders.

We constructed and trained the model with PyTorch.

5. Training the Model

We split the dataset into **training** and **validation** sets.

We used data loaders to apply transformations like:

- Resizing the images
- Normalizing pixel values
- Converting images to PyTorch tensors

Training settings:

- **Optimizer**: Adam
- **Learning Rate**: 0.001
- **Batch Size**: 8
- **Loss Function**: CrossEntropyLoss
- **Epochs**: 50

6. Prediction and Evaluation

After training, we used the model to predict segmentation maps for the validation images.

- Each output was a **PNG mask** showing which pixels belong to which class.
- We resized both predicted and ground truth masks to **1280x720** using **nearest neighbor interpolation** (as required).

- To check how well the model worked, we used the **mean Intersection over Union (mIoU)** metric.

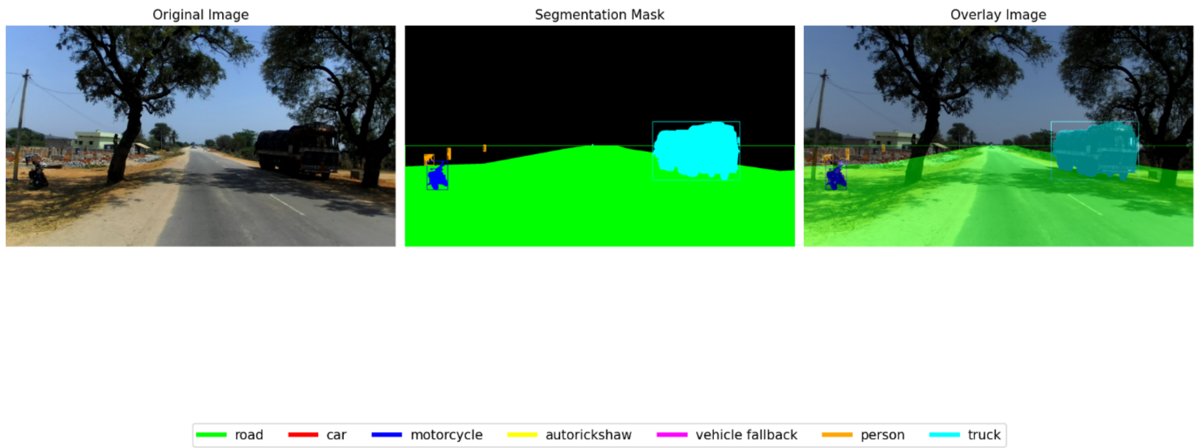
7. Visualization and Results

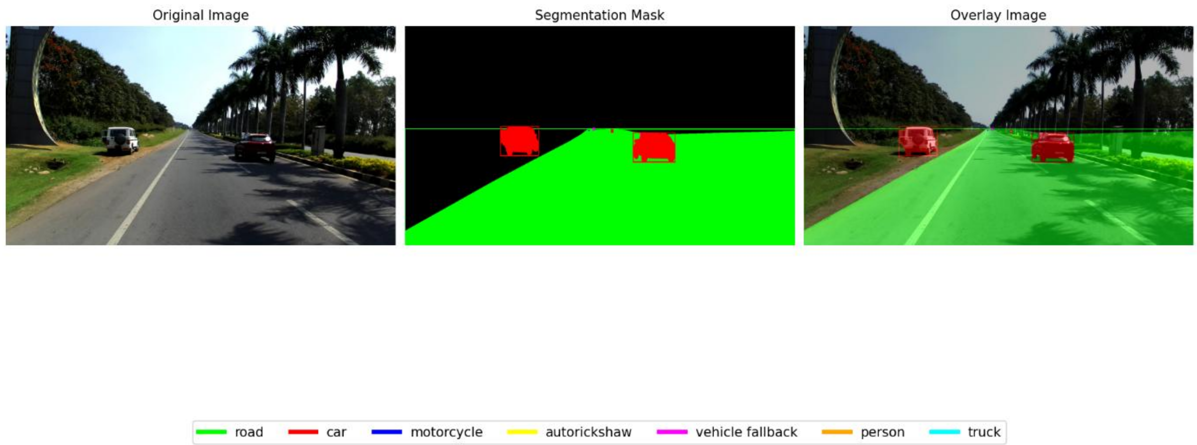
To visualize the results, we used **OpenCV** to overlay the predicted masks on the original images.

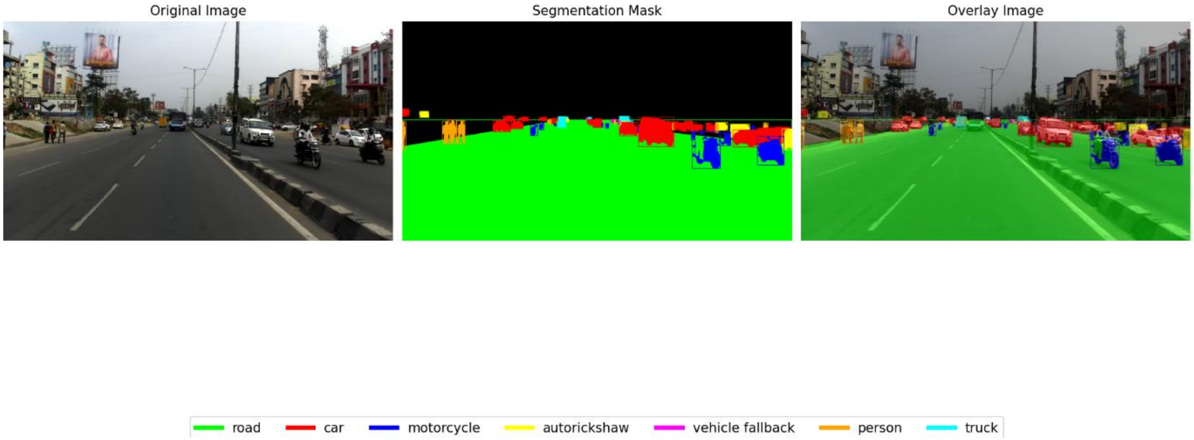
This helped us:

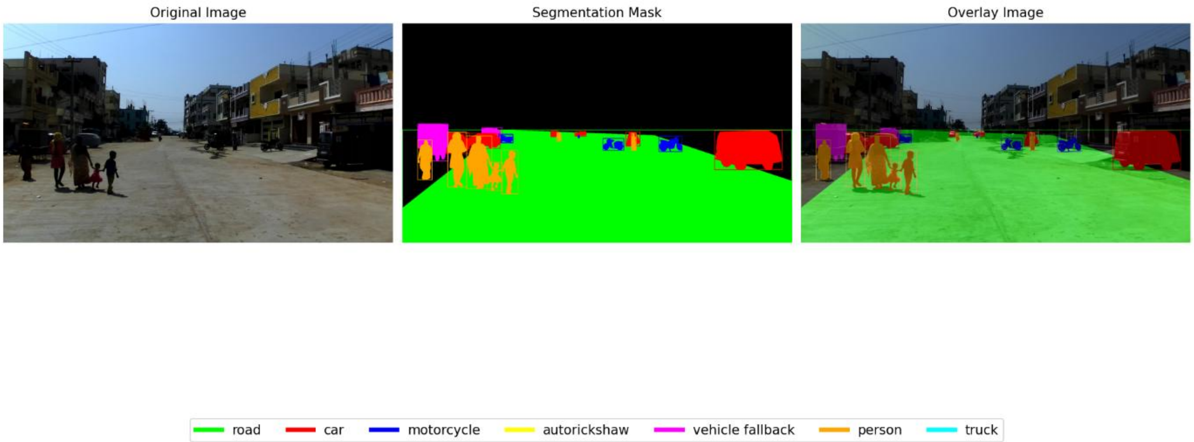
- See how accurately the model recognized different objects
- Understand its performance in difficult scenes like dark lighting or heavy traffic.

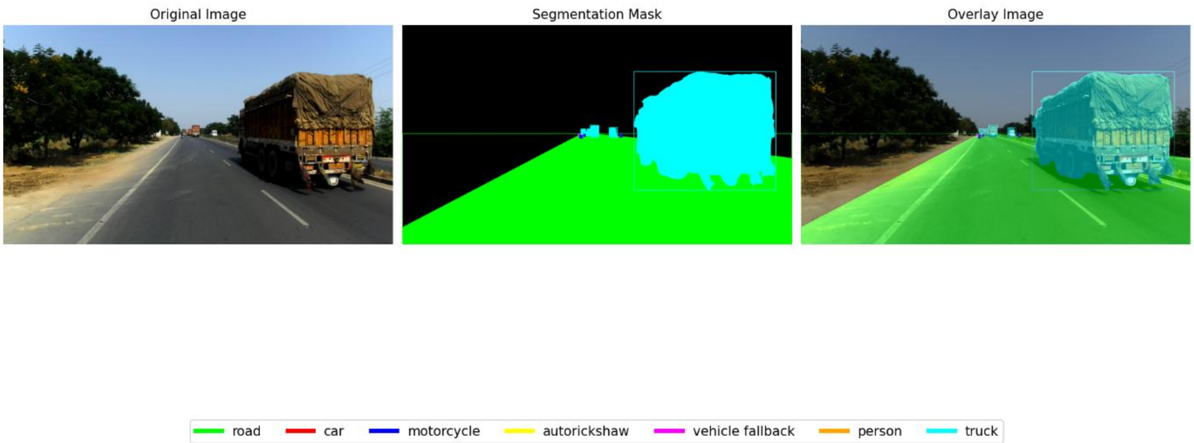
RESULT






Original Image
Segmentation Mask
Overlay Image
road
car
motorcycle
autorickshaw
vehicle fallback
person
truck


Original Image
Segmentation Mask
Overlay Image
road
car
motorcycle
autorickshaw
vehicle fallback
person
truck


Original Image
Segmentation Mask
Overlay Image
road
car
motorcycle
autorickshaw
vehicle fallback
person
truck

CONCLUSION

In this project, we investigated the application of semantic segmentation to identify road hazards in real-world driving conditions, with an emphasis on unstructured roads that are typical in India. Through the use of deep learning architectures such as DeepLabV3 and U-Net, and training them using the Indian Driving Dataset (IDD), we were able to effectively identify and classify different road hazards such as potholes, debris, vehicles, and pedestrians at the pixel level.

The approach included an end-to-end pipeline from model selection and preprocessing to training, evaluation, and post-processing. Intersection over Union (IoU) metric-based evaluation showed the model's capacity to segment and localize road hazards accurately. Incorporating post-processing with Conditional Random Fields (CRFs) improved segmentation accuracy and boundary sharpness further.

This research demonstrates semantic segmentation to be a robust method for the detection of road hazards, providing fine-grained scene understanding that may be directly incorporated into sophisticated driver-assistance systems (ADAS) or autonomous driving platforms. It also emphasizes the need for contextually suitable datasets such as IDD for model training appropriate for realistic, complex driving scenarios.

FUTURE SCOPE

The existing research on semantic segmentation-based road hazard detection presents a robust base for road safety improvement and assistive autonomous driving. There are a number of areas where the project can be further expanded. One significant avenue is optimizing real-time performance through model optimization or using lightweight architectures like for implementation on embedded or mobile platforms. Furthermore, including a priority-based hazard classification system would assist autonomous systems in evaluating and responding to threats more effectively based on their severity. Combining information from various sensors such as LiDAR, Radar, and GPS can increase the robustness of detection, particularly under adverse conditions such as low visibility or inclement weather. Subsequent models might also be made more robust employing sophisticated methods like attention mechanisms, transformer-based architectures, or GANs for better performance under challenging situations where there are occlusions or novel dangers. Further enhancing model generalization can also be achieved through increasing the size of the dataset with additional detailed and varied annotations or through data generation. Finally, this module for segmentation may also be deployed as part of a greater end-to-end framework for autonomous vehicles to assist decision-making, route planning, and collision detection.

REFERENCES

[1] Dewangan, D. K., & Sahu, S. P. (2021). Road detection using semantic segmentation-based convolutional neural network for intelligent vehicle system. In Data engineering and communication technology (pp. 629-637). Springer, Singapore.

[2] Baheti, B., Gajre, S., & Talbar, S. (2019, October). Semantic scene understanding in unstructured environment with deep convolutional neural network. In TENCON 2019-2019 IEEE Region 10 Conference (TENCON) (pp. 790-795). IEEE.

[3] Hong, Y., Pan, H., Sun, W., & Jia, Y. (2021). Deep dual-resolution networks for real-time and accurate semantic segmentation of road scenes. arXiv preprint arXiv:2101.06085.

[4] Chen, Y., Li, W., & Van Gool, L. (2018). Road: Reality oriented adaptation for semantic segmentation of urban scenes. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 7892-7901).

[5] Xie, E., Wang, W., Yu, Z., Anandkumar, A., Alvarez, J. M., & Luo, P. (2021). SegFormer: Simple and efficient design for semantic segmentation with transformers. Advances in Neural Information Processing Systems, 34, 12077-12090.

[6] Zou, Y., Yu, Z., Kumar, B., Wang, J.: Unsupervised domain adaptation for semantic segmentation via class-balanced self-training. In: ECCV . pp. 289–305 (2018).

[7] Tranheden, W., Olsson, V., Pinto, J., Svensson, L.: DACS: Domain Adaptation via Cross-domain Mixed Sampling. In: WACV. pp. 1379– 1389 (2021).

[8] Wang, Q., Dai, D., Hoyer, L., Fink, O., Van Gool, L.: Domain adaptive semantic segmentation with self-supervised depth estimation. In: ICCV . pp. 8515–8525 (2021).

[9] Liu, Y., Deng, J., Gao, X., Li, W., Duan, L.: Bapa-net: Boundary adaptation and prototype alignment for cross-domain semantic segmentation. In: ICCV. pp.8801–8811 (2021).

[10] Zhang, P., Zhang, B., Zhang, T., Chen, D., Wang, Y., Wen, F.: Prototypical pseudo label denoising and target structure learning for domain adaptive semantic segmentation. In: CVPR.pp. 12414–12424 (2021).

[11] Hoyer, L., Dai, D., Van Gool, L.: DAFormer: Improving network architectures and training strategies for domain-adaptive semantic segmentation. In: CVPR (2022).

[12] Hoyer, L., Dai, D., & Van Gool, L. (2022). HRDA: Context-Aware High-Resolution Domain-Adaptive Semantic Segmentation. arXiv preprint arXiv:2204.13132.