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, how the 27 semantic classes are labelled.

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), which tells us how the how close the model's output are. It also measures metrics like pixel accuracy.

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.

IMPLEMENTATION

1. Setting Up the Environment

We used Windows 11 with Python 3.10.2 to build the project. We began the implementation on Google Colab because it's an excellent cloud environment for running deep learning models. We enabled GPU acceleration to speed up the training process. In order to keep things organized and ensure that we didn't lose work between sessions, we also mounted Google Drive. There, we placed the dataset and model checkpoints so that if necessary, we could pick up where we left off.

2. Preparing the Dataset

We employed the IDD Segmentation Part 2 dataset on Kaggle that contains urban road images and corresponding grayscale segmentation masks. Every pixel in the mask corresponds to a particular object such as a road, car, person etc. . We structured everything into folders and utilized a custom data loader to ensure that every image was paired with the correct mask.

3. Creating Segmentation Labels

All images and masks were resized to 256×256 for consistency. The mask values were kept unchanged to preserve labels, and we used data augmentation to mimic different lighting and scene conditions.

- Resize all input images and masks to 256x256 resolution.
- Normalize pixel values (e.g., scale to 0–1 or mean-std normalization).
- Convert masks to class labels (0–25), ensuring correct mapping.

4. Model Selection and Configuration

For the semantic segmentation task, the DeepLabV3+ architecture was chosen due to its strong performance in preserving object boundaries and capturing context at multiple scales. The model utilized a ResNet50 backbone pre-trained on ImageNet. To suit the IDD dataset, which includes 26 semantic classes, the final classifier layer of the model was modified to produce 26 output channels.

5. Training Strategy

The dataset was divided into training and validation subsets. Training involved passing the preprocessed images through the model in mini-batches. The CrossEntropyLoss function was used to calculate the error between the predicted and true class labels. The Adam optimizer was used to update model weights during backpropagation. The training loop was run for a few epochs initially to verify correctness, with plans for further training to improve accuracy.

- Loss Function: CrossEntropyLoss with ignore_index=255
- Optimizer: Adam, learning rate = 0.0001
- Platform: Trained on Google Colab GPU
- Epochs: 2 epochs
- Logging: Recorded training loss per epoch

6. Model Evaluation

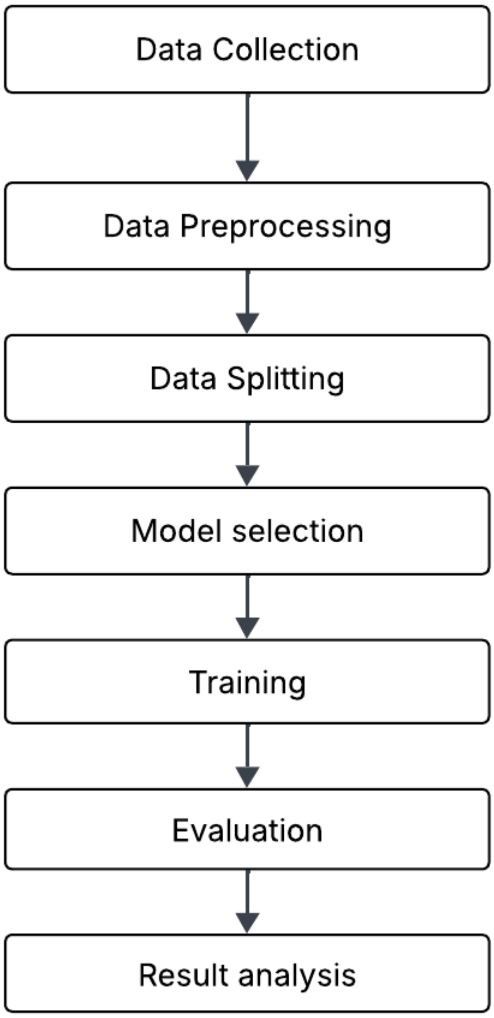
The evaluation was conducted on the validation set with all predictions and ground truths resized to 1280×720 resolution, in line with the AutoNUE benchmark format.

- Pixel Accuracy
- Mean IoU

These metrics indicate that the model was able to reliably distinguish between various elements of the road scene, such as vehicles, roads, pedestrians, and buildings, even in unstructured and cluttered environments.

7. Result Analysis

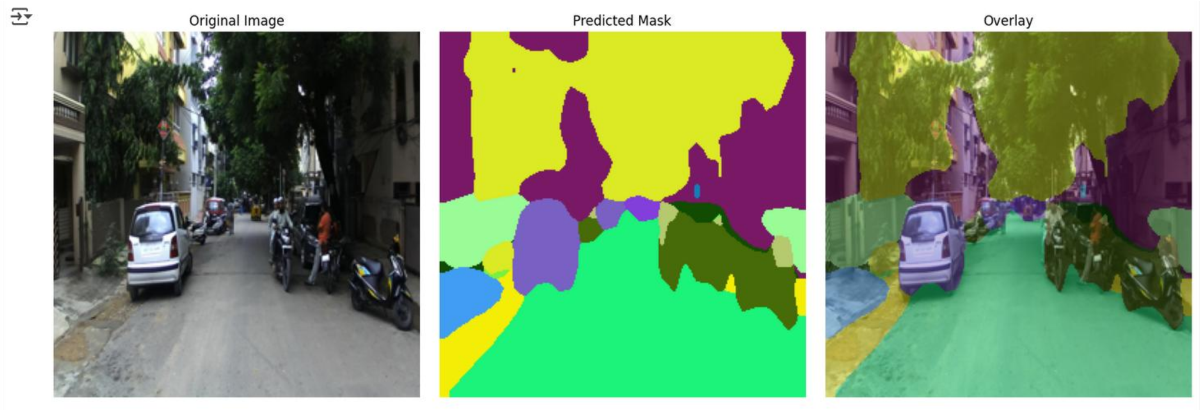
To check how well the model was performing, we placed the predicted masks over the original photographs. It helped us determine if the model was catching the correct areas, particularly in crowded street scenes such as the ones you get on Indian roads. The outputs looked great, with the model being able to capture small details well and maintain edges clean, which actually counts for a lot things such as autonomous technology or traffic patterns. Overlay predicted masks on original images with OpenCV's functionality. Assign unique colors to every class for easier interpretation.

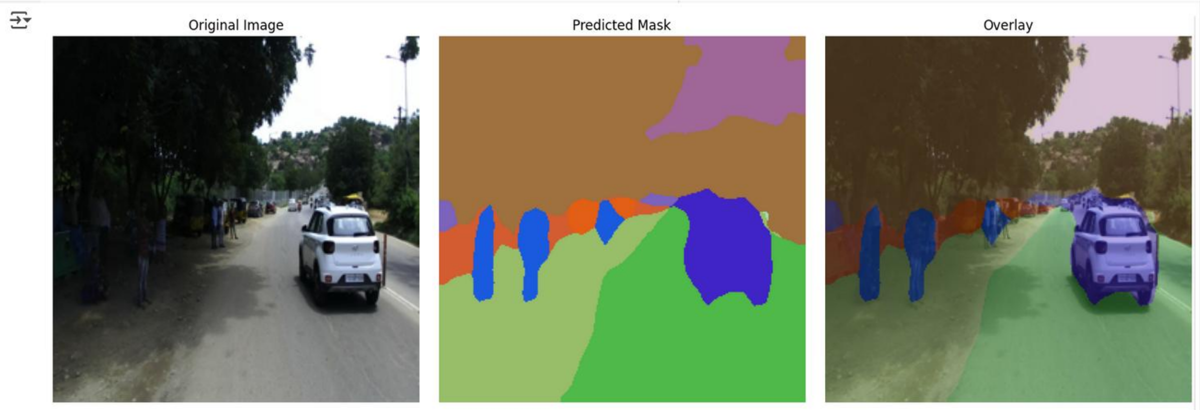


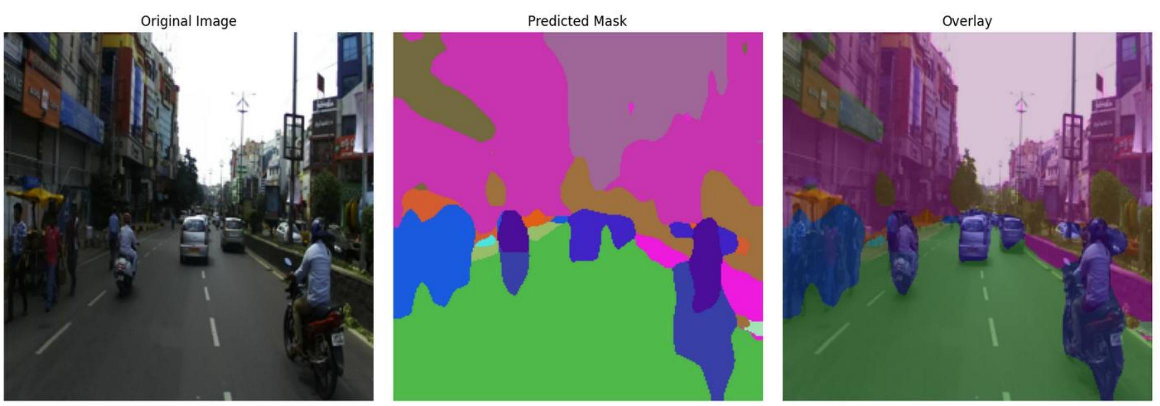
RESULT

We tested our DeepLabV3+ (ResNet-101) model on real images from the IDD Part-2 dataset using Google Colab (GPU). Visual results included:

- Left: Original road image
- Center: Predicted segmentation mask (26 classes)
- Right: Overlay of prediction on original image







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.