International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 10 Issue: 04 | Apr 2023 www.irjet.net p-ISSN: 2395-0072

Autonomous Driving Scene Parsing

Neha Gupta[1], Mohan Krishna Gupta[2], Manan Mehrotra[3], Prakhar Arora[4], Rajul Bhatnagar[5]

[1]Assistant Professor of Computer Science and Engineering, MIT College, Uttar Pradesh, India
[2]Student of Computer Science and Engineering, MIT College, Uttar Pradesh, India
[3]Student of Computer Science and Engineering, MIT College, Uttar Pradesh, India
[4]Student of Computer Science and Engineering, MIT College, Uttar Pradesh, India
[5]Student of Computer Science and Engineering, MIT College, Uttar Pradesh, India

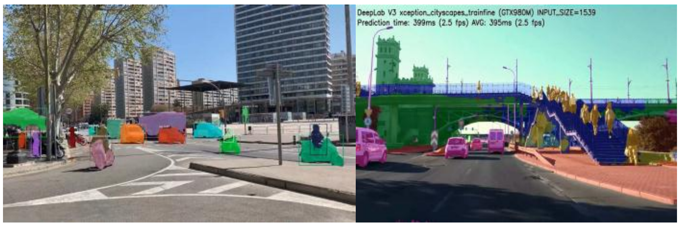
---***---

Abstract - *In this research paper, we give the semantic segmentation of driving scenes in unconstrained conditions. The focus of previous approaches has been on constrained environments but we focus on the unconstrained environment of Indian roads. We have used Indian Driving Dataset (IDD[5]) which consists of 182 drive sequences on Indian roads. To perform semantic segmentation, we have used U-Net[6] model which is completely convolutional neural network, we have made some slight adjustments in the architecture for our purposes.*

***Key Words*: Deep learning, classification, convolutional neural networks, semantic segmentation, OpenCV.**

1.INTRODUCTION

Autonomous Driving depends on information that has been processed from several sensors that are positioned above the car. These sensors enable the vehicle sense its surroundings, comprehend traffic scenes, and manage its motions, acting as its sense of hearing and sight. High-resolution cameras, radar, and Light Imaging Recognition and Ranging (LiDAR) are the main types of sensors used to classify objects using feature extraction and measure their distance from other objects using radio waves and illumination in order to create a three-dimensional (3D) image of the object. Environment. Several types of extra sensors, such as infrared, sonar, micro-radar, ultrasonic, and short-range sensors, have been fitted for autonomous vehicles to prevent collision with road obstacles. Similar to this, autonomous vehicles employ vision sensors to enable them to comprehend the visual components of their environment. Lane detection, traffic light and road sign analysis, pedestrian and vehicle detection, among other visual display and understanding problems, are all part of autonomous driving. By gathering this data, automated vehicle behaviours like lane changes, braking, and turning manoeuvres can be better and safer instructed.



(a) Instance Segmentation (b) Semantic Segmentation

Examples of segmented images for an autonomous vehicle to aid in scene analysis include (a) an instance segmentation exemplar where various objects from nearly identical classes are segmented into various colours with their own boundary pixels, and (b) a semantic segmentation example where objects are illustrated with a single colour without any distinction.

Vision sensor data is arguably regarded as the most reliable source of information among those gathered for vehicle decision-making. As a result, this field of study has received much research and application in Intelligent Transportation Systems (ITS), primarily from the perspectives of machine learning and deep convolutional neural networks (CNN). By iteratively collecting model characteristics from the input image and optimizingly achieving better representations, deep CNNs are neural networks with a variety of functional layers for image processing. Similar techniques are utilised for scene analysis from vision data, where deep CNN is used to real-time photos, for example, to determine where a pedestrian is and how far away it is from an autonomous vehicle. In contrast to this streamlined overview of computer-generated landscapes, the complicated models that are now being proposed are capable of generating multiple labels like pedestrians and vehicles as well as the localization.

2. LITERATURE REVIEW

To understand the development of autonomous driving research in recent years, it is necessary to organise a literature review to understand the various application areas by which autonomous driving has developed, as well as to recognise research gaps. Thus, the research process, approaches and findings of the literature review are introduced in the next sections.

2.1. Xinyu Huang, Xinjing Cheng, "The ApolloScape Dataset for Autonomous Driving"

They offer a huge, detailed dataset of street maps in this project. This dataset includes video sequence with instance-level annotations, 2D/3D annotation and location information, various annotated lane markings, and scenarios that are more sophisticated than those in previous datasets..

In the future, we'll start by increasing our dataset to one million labeled video frames with a wider range of weather, such snow, rain, and fog. In order to produce depth maps and panoramic photos in the near future, we also intend to attach stereo cameras and a panoramic camera system. For moving objects, the present version still lacks detailed information. For both stationary backgrounds and moving objects, they would like to produce full depth information..

2.2. Juan Rosenzweig, Michael Bartl, "A Review and Analysis of Literature on Autonomous Driving"

Autonomous vehicles are already within reach; for instance, Audi and Mercedes (Bartl, 2014) declared that their highly automated features were virtually ready for production. As a mirror of the news report, we can gradually see how this innovation manages to enter our daily lives, with examples such as cars taking a blind man for tacos in 2013,coast-to-coast excursions in 2016 (CNN), an Italy-to-China trip in 2014 (Broggi, et al.), and 6 million miles previously covered by Google (Urmson , 2015). However, despite being largely ready, the major competitors in the race, including Audi, Bentley, Chevrolet, Ford, Automakers, Volvo, Lexus, Lincon, Lamborghini, Nissan, Tesla, and Volvo, are attempting to integrate it gradually into their vehicles.

2.3. Markus Hofmarcher & Thomas Unterthiner, "Visual Scene Understanding for Autonomous Driving Using Semantic Segmentation"

In comparison to the ENet segmentation network, a popular network architecture for effective semantic segmentation on embedded devices, they rated the network favourably. There are comparable, more recent architectures, but none have been compared.

They presented a tiered approach, where only specific components are educated in an end-to-end manner, as a supplement to full deep learning, which lacks explainability. A human specialist with thorough understanding of the system may decipher the definition of intermediate nodes, which are outputs through one layer and inputs to another. A potent method that offers a reduced dimensional abstraction of incoming video signals is semantic segmentation.

2.4. Manikandan Thiyagarajan, "Self-Driving Car"

This project focuses on making improvements in pedestrian safety and commuting , drastically reducing accidents and human mistakes through continuous system learning because self-driving cars are the primary advancement in the automatable industry in the future. This initiative will revolutionise how people with disabilities and blind people who can drive themselves are transported. Mobile applications can be created using our solution as the foundation, allowing users to summon a vehicle through an app and, once the law is passed, manufacture a completely autonomous vehicle .

2.5. Girish Varma, Anbumani Subramanian, Anoop Namboodiri, Manmohan Chandraker, "IDD[5]: A Dataset for Exploring Problems of Autonomous Navigation in Unconstrained Environment"

They provide as a fresh dataset for researching issues with autonomous navigation in unstructured driving situations. They point out a number of flaws in the current data sets, including the inability to discriminate between safe and harmful roadside regions, the need for more labels for cars, and a label hierarchy that lessens ambiguity. They examine the dataset's class imbalance and label statistics. With relation to previous semantic segmentation datasets, we also look at the domain discrepancy properties. Our semantic segmentation datasets are more diverse than other datasets since they were collected in India, where there are disparities in background categories and road user looks. This not only presents intriguing challenges to the state of the art in semantic segmentation, but it also represents the first attempt to address problems linked to autonomous driving to our knowledge.

2.6. Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-Net[6]: Convolutional Networks for Biomedical Image Segmentation.

The u-net[6] architecture performs exceptionally well in a variety of biomedical segmentation applications. It just requires a small number of annotated photos and has a very manageable training time of only eight hours on an NVidia Titan GPU because of data augmentation using elastic deformations (5GB). Depending on Caffe and a trained network, we offer a whole implementation. We are sure that many such jobs can be completed with ease using the u-net[6] design.

Table -1: Table showing comparative study of 5 Research Papers:

PAPER TITLE	AUTHOR(S)	IDEA	FEATURES/TECH	CONCLUSION
1. Xinyu Huang, Xinjing Cheng,"The ApolloScape Dataset for Autonomous Driving"	Xinyu Huang, Xinjing Cheng	Present a large comprehensive street view dataset.	The dataset includes scenarios that are more complicated than those in existing datasets, as well as 2D/3D annotation and location data, a variety of annotated lane markings, and video frames labelled at the instance level.	We will raise the total size of our dataset to one million annotated video frames with a wider range of weather, such as snow, rain, and fog.
2. Juan Rosenzweig, Michael Bartl, "A Review and Analysis of Literature on Autonomous Driving"	Juan Rosenzweig, Michael Bartl	Examples include driving from Italy to China in 2014, 600,000 miles previously covered by Google, and autos bringing a blind guy for tacos as early as 2013 (Google, 2013). (Urmson, 2015).	A prototype for a self-driving automobile is being developed that combines a number of technologies, including lane detection, disparity maps for measuring distances between the car and other vehicles, and support vector machine classification algorithms for anomaly detection. With our data collection, very good accuracy.	The system's goal is to use innovative methods to increase the reliability of self-driving technology. The use of a self-driving car that can make informed decisions is the main focus. The system is regarded as a prototype vehicle that has cameras and sensors to understand its surroundings.
3. Markus Hofmarcher & Thomas Unterthiner, "Semantic Segmentation for Visual Scene Comprehension in Autonomous Driving"	Markus Hofmarcher & Thomas Unterthiner	Evaluated the network compares favourably against the ENet segmentation	ENet (Efficient Neural Network). It offers real-time, pixel-by-pixel semantic segmentation capabilities.	A self-driving car that requires little or no human involvement to navigate traffic and other challenges while detecting its surroundings. It will benefit those with disabilities and the blind.
4.Manikandan Thiyagarajan,"Self-Driving Car"	Manikandan Thiyagarajan	Examples of cars carrying a blind man for tacos as early as 2012 (Google, 2012), coast-to-coast road trips (CNN, 2015), traveling from Italy to China (2013), and the 700,000 miles already driven by Google (Urmson, 2014).	Drivers who are blind are independent. Our device can serve as the foundation for mobile applications that let the owner summon the vehicle through the app and construct a completely autonomous vehicle.	A self-driving vehicle that can recognise its surroundings and negotiate traffic and other obstacles with little to no assistance from humans. The blind and those with disabilities will profit from it. It will help differently abled people and blind people as well.
5. Girish Varma, Anbumani Subramanian "A Dataset Evaluating Problems of Autonomous Navigation in Unconstrained Environment"	Girish Varma, Anbumani Subramanian, Manmohhan Chandra	Examples include driving a blind man to tacos in a car as early as 2013 (Google, 2013), taking road trips from coast to coast in 2016 (CNN), and going from Italy to China in a car as early as 2013 (Google, 2013). (Urmson, 2015)	Distinguish between safe and harmful roadside regions, add labels for vehicles, and create a label hierarchy that reduces uncertainty. These are only a few of the flaws in the current datasets.	Distinguish between safe and risky zones along the road, the need for extra labels for cars, and a tag hierarchy that reduces uncertainty. These are only a few examples of the limitations of existing datasets.

Fig. Semantic Segmentation

3. PROBLEM STATEMENT

The challenging robotic task of autonomous driving requires vision, planning, and execution in a situation that is always going ahead. This process requires precision for safety reasons. Semantic segmentation provides pixel-accurate detection of objects in a driving scene. Most of the work has been done on the structured conditions as there are lot of data sets available for structured conditions. But the environment of Indian road is unconstrained and it is unstructured. Most of the datasets available are not on Indian roads they are more structured compared to the scenarios on Indian roads. So, in order to handle these unconstrained conditions, we have utilized Indian Driving Dataset. Past approaches have not completely utilized this dataset.

A dataset called IDD[5] consists of 11,000 photos with 33 classes that were gathered from 180 driving situations on Indian roadways. This dataset offers the information needed for an unrestricted environment, and it can be utilised for semantic segmentation as well as item recognition. The collection includes road scenes that were recorded in Bangalore, Hyderabad, and the surrounding areas of these cities. While some of the photographs are 720p, the majority are 1080p.

4. PRESENTLY AVAILABLE SOLUTIONS

Modern developments in machine learning and computer vision have permit advancements in autonomous navigation. Although, most of the already present research addresses on European driving situations, while very less progress has been made for the Indian driving situations.. Our contribution points to achieve a complete spatial understanding for the Indian driving situations. Our central focus on semantic segmentation.

The topic of automated vehicles has advanced significantly in recent years, attracting a great deal of interest and branching out into numerous sub-fields that address every facet of a self-driving vehicle. Vehicle-to-vehicle communication, energy storage systems, sensors, safety equipment, and more are a few examples. Among these is the difficult Computer Vision (CV) task of "scene understanding," which deals with the processing of unprocessed environmental data to produce a representation of the scene in front of the car that permits later interaction with the surrounding environment. Scene understanding involves interpreting the scene that has been viewed using a network of sensors to perceive, analyse, and develop the scene. It includes a range of difficult tasks, including image classification and more complex ones like object identification and semantic segmentation . The first task entails giving the input image a global label; however, because diverse items in the environment must be located, it has limited application in an autonomous vehicle scenario. The second assignment gives a more thorough description, finds all things that have been identified, and classifies them. The final assignment, which is the most difficult, is for giving each pixel in the input image a class. This challenge provides a detailed semantic description, so understanding the scene is the primary objective of the pre-processor, which calls for sophisticated machine learning architectures.

5. CONCLUSION

Not Satisfied with the previous approaches which tend to focus on constrained conditions. This drives us to focus on unconstrained conditions using IDD[5] dataset.

For Semantic Segmentation, we have used U-Net[6]. It was introduced in 2015 by Olaf Ronneberger, Philipp Fischer, and Thomas Brox. It is a fully convolutional architecture. It comprises encoder and decoder. We have implemented the architecture in PyTorch. We have slightly modified this architecture, after every two layers of convolutional layer we have added a Batch Normalization layer and later ReLU activation. Instead of up-conv layer, we have used Transpose Convolutional layers. Also, we have used padding to ensure that output of every layer in encoder is of same size as in corresponding input of decoder layer.

ACKNOWLEDGEMENT

Our mentor Assistant Professor Mrs. Neha Gupta has my warmest gratitude for her encouragement and assistance throughout the review paper. We learned about recent developments in this subject. We owe her a debt of gratitude. After completing this review paper, we have gained a great deal of confidence.

REFERENCES

[1] Xinyu Huang, Xinjing Cheng, Qichuan Geng, ""The ApolloScape Dataset for Autonomous Driving".

[2] Juan Rosenzweig, Michael Bartl, "A Review and Analysis of Literature on Autonomous Driving" in THE MAKING-OF INNOVATION, E-JOURNAL makingofinnovation.com, OCTOBER 2015.

[3] Markus Hofmarcher & Thomas Unterthiner, "Visual Scene Understanding for Autonomous Driving Using Semantic Segmentation" in LIT AI Lab & Institute for Machine Learning, Johannes Kepler University Linz, A-4040 Linz, Austria.

[4] Manikandan Thiyagarajan, "Self-Driving Car" in International Journal of Psychosocial Rehabilitation · April 2020.

[5] Girish Varma, Anbumani Subramanian, Anoop Namboodiri, Manmohan Chandraker, "DD: A Dataset for Exploring Problems of Autonomous Navigation in Unconstrained Environment" http://idd.insaan.iiit.ac.in/.

[6] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation.