Case Review: Deep learning-based Lane Detection in Foggy weather Conditions

Introduction

Like human drivers, self-driving vehicles can have trouble seeing the inclement weather. Many autonomous cars rely on lidar and radar technology, which works by bouncing laser beams off surrounding objects to give a high-resolution 3D picture on a clear day but does not do so well in fog. Cameras can be thrown off by lighting conditions or snow-covered signs and lane markings. Driving in foggy weather conditions has been a significant safety concern for many years due to poor driver lane-keeping performance.[1]

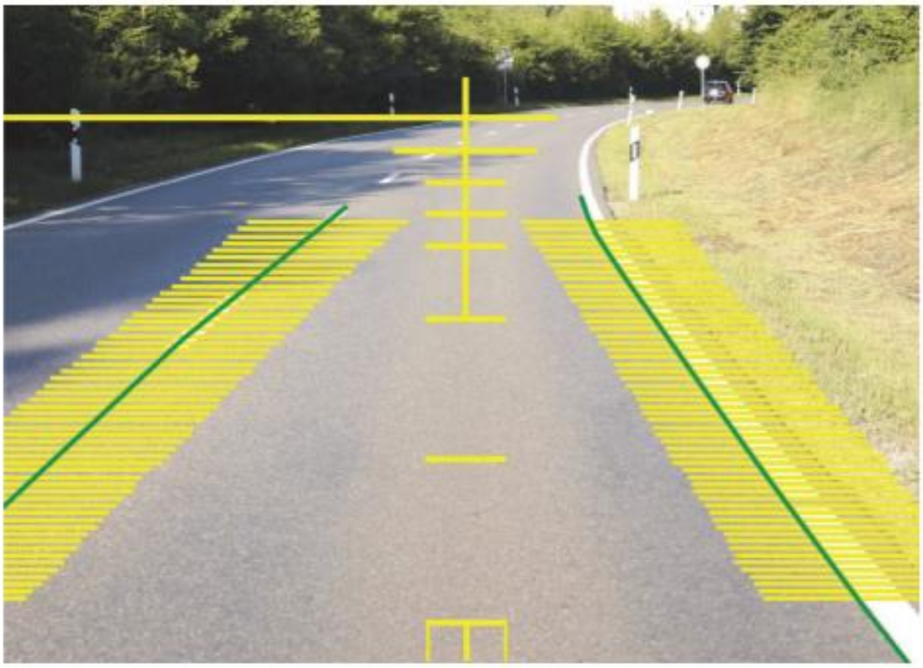
Autonomous vehicular navigation that uses visible light has one of the chief obstacles, an inability to handle misty driving conditions. Most cars rely on Advanced Driver Assistance Systems (ADAS), which offers lane keep assist and steering assist. The ADAS driving applications like Lane-keeping assist, Lane Departure Warning, Lane change assistance, and Adaptive Cruise Control are the core use cases of this lane detection technology. Deep learning models with real-time lane detection algorithms can generate undesired effects in unfavorable climatic conditions like foggy weather. The primary aim of these algorithms is to be able to provide accurate predictions of lanes even in adverse weather and hence a lot of research is ongoing to make the lane detection robust and repetitive.

Background

As the automotive industry progresses towards full-scale autonomous driving it has become important to develop accurate methodologies that function in varied weather conditions [2]. While classical approaches of image processing have been applied for this initially, the rise of convolution neural nets has accelerated the development and deployment process in the industry. Many cars now come with lane departure warning systems that alert the driver when the car is unintentionally crossing the lane markings. This system totally depends on the front camera fixed on the front of the car; This has driven the research to find a robust lane prediction algorithm that can accommodate all types of weather conditions.

Autonomous car companies such as Tesla, NVIDIA, Waymo have faced their fair share of speed bumps while tackling the issues with navigating in harsh weather. A Professor at UC San Diego[3] has used an inexpensive approach of fusing radar and lidar as radars are cheap. His team conducted tests using simulation of clear days and nights and scenarios where there is a lot of fog. The lidar with radar system performed better than the lidar alone system. NVIDIA lab is using AI with the aim to be improving autonomous vehicle perception. They have trained a Deep Neural Network to detect moving and stationary objects as well as accurately remove the false detections.

In December 2020 feature [4] consists of primary weather and road conditions that can hinder the functionality of lane-keeping assist included in the 2014-2020 CLA-Class, 2011-2019 B-Class MPV, and 2014-2020 B-Class. The issue talks about several environmental factors that can cause poor sensor performance in times of heavy fog or rain. These implicit errors are thus a weak point that needs continuous improvement through Deep learning methods



The multifunction camera and radar sensor system control unit each collect data on the lane markings separately, then combine their analyses to determine the relative position of the vehicle in the lane. The system cannot function where lane markings are absent or visibility is obscured.

Figure 1: Mercedes Lane keep assist using multi-sensor fusion

Problem Description

With the low visibility of roads due to the foggy climate, the camera is unable to capture a clear image of roads. On a clear day, the light’s return time accurately indicates the distances of the objects that reflected it, but fog causes light to scatter and bounce in random ways. In foggy weather, most of the light that reaches the camera’s sensor will have been reflected by airborne water droplets which may not be the actual objects that need to be avoided. And even the light that does reflect potential obstacles will arrive at different times, having been deflected by water droplets on both the way out and the way back. We attempt to implement working with a dataset that essentially captures these realistic scenarios and training a deep learning model using advanced computer vision and convolutional neural network frameworks.

Existing Approaches

Lane detection suffers as the road illuminations change, when there is damage on the lane markings and when the weather affects visibility. A review published in the IEEE journal [5] identified the research gap in existing techniques. In [6], they propose HSV color space-based detection algorithm that uses Progressive Probabilistic Hough Transform to detect lane and lane markings. This method does not perform well under foggy conditions and when the lane markings are not in good condition. In [2], the images are classified and processed differently according to their illuminations. According to the whole gray mean edge detection filters are applied to extract lane pixels.

On the other hand, many neural networks have been applied to [5] predict lane markings. Using neural networks increases the cost of processing time in real-time applications. We see that many networks have chosen datasets that are made mostly of cityscape data that involves minimum instances of harsh weather conditions. [7] describes a robust Deep Convolution Neural Net that is a Spatial-

Temporal Method with segmentation and structural analysis. This is the latest effort in the lane detection algorithms and has robust performance against weather conditions but suffered under disturbances to the camera and when the lane markings were too worn out.

Lane-DeepLab is one of the methods that addresses this problem, it is a lane segmentation detection method that applies the DeepLab3+ state of the art deep learning model for semantic image segmentation model. The fully convolutional network FCN assigns semantic label at the pixel level using an end-to-end encoder decoder network architecture. The backbone of this architecture is ResNet101 for feature extraction and atrous spatial pyramid pooling ASPP is employed in the encoder part. This technique down samples the spatial resolution of the input, developing lower-resolution feature maps and up samples the feature representation to fully resolution segmentation map. The down sampling is essentially developing lower-resolution feature mappings which are learned to be highly efficient at discriminating between classes of the image data [8]. This method involves up sampling the resolution of feature map whereas pooling operations involve down sampling of the resolution by summarizing a local area with a single value. This can be averaging or argmax value in the given area of pixels. Another way of down sampling is transpose convolution which involves using single value from low-resolution feature map and multiply by the weights in our filter by this value, project these weighted values on output feature map. The encoder network commonly used is 13 convolutional layers in VGG16 network. Lane-DeepLab adds attention mechanism to the encoder part, redesign the ASPP module, and name it attention atrous spatial pyramid pooling (ARM-ASPP). After which the method uses feature fusion method in the decoder module to combine the high-level and low-level semantic information to obtain more abundant features. Additionally employing the Semantic Embedding Branch (SEB) to combine high-level and low-level semantic information to obtain more abundant features using Single Stage Headless module abundant features could be extracted, thus making the lane detection technique more robust.

SegNet is an improved model of deep convolutional encoder-decoder architecture for image segmentation. This core trainable segmentation engine consists of the encoder-decoder network as mentioned above with an additional pixel-wise classification layer. The novelty of SegNet lies is in the way the decoder up samples its lower resolution input feature map. The decoder uses pooling indices computed in the max-pooling step of the corresponding encoder to perform non-linear upsampling, eliminating the need for learning to up sample. The upsampled maps are sparse and are then convolved with trainable filters to produce dense feature maps. In comparison with the widely adopted FCN and with the well-known DeepLab LargeFOV, DeconvNet architectures. SegNet provides good memory versus accuracy trade-off involved in achieving good segmentation performance. SegNet is motivated by scene understanding applications. Hence, it is designed to be efficient both in terms of memory and computational time during inference. It is also significantly smaller in the number of trainable parameters than other competing architectures. There is a controlled benchmark of SegNet and other architectures on both road scenes and SUN RGB-D indoor scene segmentation tasks.[9]

Proposed Solution

Ground truth data generation: We chose to implement a robust lane detection algorithm on the synthetically generated foggy dataset [10]. This synthetically generated data is from the Cityscapes dataset. This is our base dataset that will be used to generate ground truth for the training of the neural network. We will use heuristic approaches to generate the ground truth data which will reduce manual annotations required for training the neural network.

Neural Network Implementation: Semantic Segmentation based lane detection approach does not address the problem of realistic scenarios consisting of high speed, severe occlusion, and extreme lighting conditions segmentation-based shape description is inefficient for lane lines due to the excessively high degree of freedom. To address this problem, we attempt to solve this using top-to-down lane detection on framework that detects the lane instances first and then instance-wisely predict the shapes. Below is the table of benchmark Convolutional Neural networks that use instance-wise prediction of lanes which we will be referring to in order to make our network robust using foggy dataset[9]

Category	Total	Normal	Crowded	Dazzle	Shadow	No line	Arrow	Curve	Cross	Night	FPS	GFlops(G)
SCNN [28]	71.60	90.60	69.70	58.50	66.90	43.40	84.10	64.40	1990	66.10	7.5	328.4
ERFNet-E2E [41]	74.00	91.00	73.10	64.50	74.10	46.60	85.80	71.90	2022	67.90		
FastDraw [29]		85.90	63.60	57.00	69.90	40.60	79.40	65.20	7013	57.80	90.3	
ENet-SAD [12]	70.80	90.10	68.80	60.20	65.90	41.60	84.00	65.70	1998	66.00	75	**3.9**
UFAST-ResNet34 [30]	72.30	90.70	70.20	59.50	69.30	44.40	85.70	69.50	2037	66.70	175.0	
UFAST-ResNet18 [30]	68.40	87.70	66.00	58.40	62.80	40.20	81.00	57.90	1743	62.10	**322.5**	
ERFNet-IntRA-KD [11]	72.40										100.0	
CurveLanes-NAS-S [39]	71.40	88.30	68.60	63.20	68.00	47.90	82.50	66.00	2817	66.20		9.0
CurveLanes-NAS-M [39]	73.50	90.20	70.50	65.90	69.30	48.80	85.70	67.50	2359	68.20		35.7
CurveLanes-NAS-L [39]	74.80	90.70	72.30	67.70	70.10	49.40	85.80	68.40	1746	68.90		86.5
LaneATT-Small [32]	75.13	91.17	72.71	65.82	68.03	49.13	87.82	63.75	**1020**	68.58	250	9.3
LaneATT-Medium [32]	76.68	92.14	75.03	66.47	78.15	49.39	88.38	67.72	1330	70.72	171	18.0
LaneATT-Large [32]	77.02	91.74	76.16	69.47	76.31	50.46	86.29	64.05	1264	70.81	26	70.5
CondLaneNet-Small	78.14	92.87	75.79	70.72	80.01	52.39	89.37	72.40	1364	73.23	220	10.2
CondLaneNet-Medium	78.74	93.38	77.14	**71.17**	79.93	51.85	89.89	73.88	1387	73.92	152	19.6
CondLaneNet-Large	**79.48**	**93.47**	**77.44**	70.93	**80.91**	**54.13**	**90.16**	**75.21**	1201	**74.80**	58	44.8

Figure 2: Comparison of different lane detection benchmarks

References

[1] D. Pomerleau, “RALPH: Rapidly Adapting Lateral Position Handler,” *Intelligent Vehicles Symposium, Proceedings*, pp. 506–511, 1995, doi: 10.1109/IVS.1995.528333.

[2] N. Ma, G. Pang, X. Shi, and Y. Zhai, “An All-Weather Lane Detection System Based on Simulation Interaction Platform,” *IEEE Access*, vol. 8, pp. 46121–46130, 2020, doi: 10.1109/ACCESS.2018.2885568.

[3] “Researchers Working to Improve Autonomous Vehicle Driving Vision in the Rain - AI Trends.” https://www.aitrends.com/selfdrivingcars/researchers-working-to-improve-autonomous-vehicle-driving-vision-in-the-rain/ (accessed Feb. 23, 2022).

[4] “Understanding Mercedes-Benz Lane Keeping Assist”.

[5] M. Jefri Muril, N. H. Abdul Aziz, H. Ab. Ghani, and N. A. Ab Aziz, “A Review on Deep Learning and Nondeep Learning Approach for Lane Detection System,” in *2020 IEEE 8th Conference on Systems, Process and Control (ICSPC)*, Dec. 2020, pp. 162–166. doi: 10.1109/ICSPC50992.2020.9305788.

[6] J. H. Kim, S. K. Kim, S. H. Lee, T. M. Lee, and J. Lim, “Lane recognition algorithm using lane shape and color features for vehicle black box,” *International Conference on Electronics, Information and Communication, ICEIC 2018*, vol. 2018-January, pp. 1–2, Apr. 2018, doi: 10.23919/ELINFOCOM.2018.8330549.

[7] M. M. Yusuf, T. Karim, A. F. M. S. Saif, M. Yusuf, and A. F. M. Saifuddin, “A Robust Method for Lane Detection under Adverse Weather and Illumination Conditions Using Convolutional Neural Network A Robust Method for Lane Detection under Adverse Weather and

Illumination Condition Using Convolutional Neural Network," 2020, doi: 10.1145/3377049.3377105.

[8] X. Liu and Z. Deng, "Segmentation of Drivable Road Using Deep Fully Convolutional Residual Network with Pyramid Pooling," *Cognitive Computation*, vol. 10, no. 2, pp. 272–281, Apr. 2018, doi: 10.1007/S12559-017-9524-Y/TABLES/3.

[9] L. Liu, X. Chen, S. Zhu, P. Tan, and A. Group, "CondLaneNet: a Top-to-down Lane Detection Framework Based on Conditional Convolution," May 2021, Accessed: Feb. 23, 2022. [Online]. Available: https://arxiv.org/abs/2105.05003v2

[10] C. Sakaridis *et al.*, "Semantic Foggy Scene Understanding with Synthetic Data".