Lane Detection in Adverse Visibility Conditions for Autonomous Driving Assistance

Rohit Kumar Sharma
rsharma54@wisc.edu

Varun Batra
vbatra@wisc.edu

Vibhor Goel
vgoel5@wisc.edu

1 Problem Statement

Given the fact that we are headed towards autonomous driving vehicles, lane detection becomes one of the key aspects for it. Although, there has been some work in the domain of Lane Detection [Lee and Moon, 2018] [Wang et al., 2018] [Leng and Chen, 2010], real time implementation of robust lane detection algorithms is still missing for adverse visibility conditions. In fact, Escanilla et al [Hancock and Escanilla, 2018] at UW-Madison have done some novel work in the domain of lane detection last year. Over the years, several works on detection of lanes in adverse conditions from image scenes [Shibata et al., 2014] [Fu et al., 2017] [Shen et al., 2018] [Li et al., 2017] have been proposed. We wish to research and integrate the task of lane detection in conditions such as glare, night and crowd to provide assistance in autonomous driving vehicles. We are exploring the problem as form of multiple-lane detection, and not only single lane.

2 Motivation

Lane detection is one of the most important and preliminary tasks of autonomous driving. Good lane detection can lead to better automation and enhanced driving experience. If the lane detection is not done properly, there can be adverse effects in autonomous driving which may even lead to accidents. Therefore, the accuracy of lane detection is equally important in good conditions as well as in conditions with adverse visibility conditions such as bad lighting, glare, crowd and no clear demarcations of lanes.

A lot of work has been done in the domain of lane detection, especially in very good lighting conditions and clearly marked lanes. But, there seems to be no robust solution which works well under poor conditions such as glare, bad light, night. Therefore, we propose to achieve this task of lane detection in adverse conditions by employing existing state-of-the-art deep learning based Computer Vision algorithms.

3 Related Work

Several authors have proposed techniques for lane detection, especially with the advent of autonomous driving systems. [Kumar and Simon, 2015] present a comprehensive review of current state of art techniques in lane detection. [Lee and Moon, 2018] propose a robust vision-based lane detection algorithm for tracking lanes in real-time scenarios. They also focus on other aspects such as lack of lane marking clarity, which are especially relevant for our task.

To tackle the problem of lane detection in adverse conditions, one could think of a two step approach: first, de-noising the images to remove noise of haze, rain, shadow, snow, etc., and second, detect lanes in the de-noised images. [Bossu et al., 2011] proposes a computer vision based system for de-noising the rain and snow on image sequences. The authors use the approach of separating the foreground from background in image sequences using a Gaussian mixture model. In a sequence of images, the dynamic foreground represents the noise due to rain or snow. [Wang et al., 2017] discusses a method of removing snow with image decomposition.

[Parajuli et al., 2013] propose a method for robust lane detection in case of shadows, and discontinuous illuminations. Their approach is based on local gradient features and characteristic spectrum of lanes, and then using linear prediction to find the lane markings. [He et al., 2011] talk about haze removal in single images. This approach can be used for preprocessing of unclear images having glares and bad illumination. Further, there has been some research on removing glares from images ([Wan et al., 2017], [Yang et al., 2010]). Though, all this research has been done on normal images with small glare region, and not on road data which can have large sections of glares.

In the end, [Diamond et al., 2017] propose a way of combining de-noising approaches with the actual task (classification in their case) in a single end-to-end model. We try to achieve the same thing over the model proposed by [Pan et al., 2018], who have segragated the different kinds of lane detection images for different adverse conditions in data-set. But have not proposed any specific methods for improving performance for these individual conditions like shadow, night, bad lighting, glare, crowd etc..

4 Current Progress

4.1 Dataset

As originally proposed, we started with the task of lane detection in the scenarios of snow and/or rain. However, we couldn't find good datasets that have been annotated that could be used to proceed. We did find a picture-based dataset - [Nexar, 2017] , but this was not solving our purpose. Although this dataset had a lot of samples, the problem was that it did not have many adverse samples to provide meaningful training. Instead, we found another dataset - CULane [1]. CULane is a very large scale dataset for traffic lane detection with various levels of challenging images collected by the Multimedia Laboratory at the Chinese University of Hong Kong. It has been collected by cameras mounted in the rural and urban regions of Beijing. The dataset consists of more than 55 hours of videos with over 133,235 frames extracted and marked. These frames are also categorized into normal and 8 different challenging categories with adverse visibility conditions such as harsh sunshine, sharp curves, tunnels, water on the road, missing lane markings, etc. The images contains lane markings annotated with cubic splines. These markings also include cases when the lane markings are not clear due to obstructions by vehicles, etc. Figure 1 shows various adverse conditions along with the ground truth annotations for a few sample images from the dataset. The creators of the dataset have made it public to improve the lane detection task. The diversity of the dataset and the presence of multiple challenging scenarios makes it an ideal choice for the purpose of our project.

Figure 1: Adverse conditions annotations and their distribution in CULane dataset

4.2 Spatial-CNN model

After deciding the dataset, our next challenge was to build a model to perform lane detection. The creators of the CULane dataset have implemented a baseline model in [Pan et al., 2018]. Spatial CNN (SCNN) is a novel idea which incorporates spatial semantics of pixels across rows and columns in an image. This is important in the case of lane detection because of the relationships in the traffic lanes. There is a strong prior on the shape of pixels in lane detection task. Based on the construction, SCNN is suitable for such tasks with strong spatial relationship but less appearance clues such as traffic lanes. Figure 2 contrasts the performance of SCNN with VGG16 model [Simonyan and Zisserman, 2014] for some images with different difficulties.

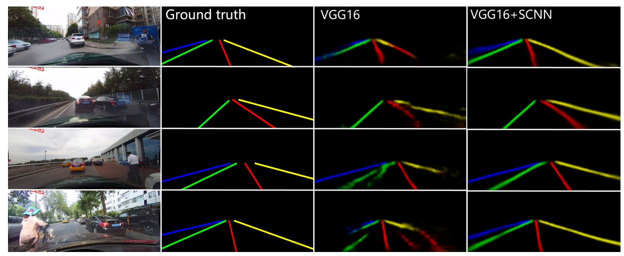

Figure 2: Spatial CNN vs other models

As of now, the state of the art performance in the task of lane detection is achieved by SCNN. For evaluating the performance, lane markings are considered as lines with 30 pixel width and Jaccard Similarity (Intersection-over-Union) is computed between the ground truth labels and predictions. Predictions with similarity higher than a threshold are considered as true predictions. For 0.5 threshold, the F1 scores of this model for various categories of the dataset are summarized in table 1 below:

Category	F1 Score
Normal	90.2
Crowded	71.9
Night	64.6
No line	45.8
Shadow	73.8
Arrow	83.8
Dazzle light	59.5
Curve	63.4
Crossroad	41.37
Total	71.3

Table 1: SCNN Results for various conditions

4.3 Our Idea

As it can be seen from table 1, even though SCNN model performs well in normal conditions, there seems to be a scope of improvement in other categories such as Night, Shadow or Dazzle light. Therefore, we want to tackle

[1] `https://xingangpan.github.io/projects/CULane.html`

one of the scenarios and improve the performance of lane detection in it. We have chosen "Glaze" condition and desire to raise the performance of lane tracking in this scenario. One sample "Glaze" image is shown in figure 3. We propose to tackle this in a two step procedure as of now. In the first step, we plan to perform deglazing using some Computational Imaging technique. Once we have retrieved an image with less glaze, we can apply the SCNN model to detect the lanes in the second step. We hope that this will improve the overall performance of lane detection.

Figure 3: Example of glare in dataset

4.4 Implementation

Parallelly we also started to setup the SCNN code base to perform the training. SCNN code is available on Github [2]. It is implemented in Lua based on Torch framework. It is also based on VGG16 model [Simonyan and Zisserman, 2014]. The pretrained SCNN parameters are also provided for testing. As per the hardware requirement suggestions by the developers of SCNN, we need 1 GPU with 3G memory for testing and 4 GPUs with 12G memory for training. It was a challenge to acquire GPUs and setup our environment with necessary tools installed. We are currently using Wisconsin Applied Computing Center's Euler which is a *Supercomputing Cluster* based on Slurm Workload Manager to train the model. We have our accounts ready, and the next step is to get Lua-Torch setup on these machines to run our code there. We require admin permissions for that and should get this access shortly. Due to the difficulties we have been facing to run the model, we started looking at an alternate implementation of the same model based on Tensorflow framework in Python [3]. With this Tensorflow implementation, we will also be able to perform distributed training easily. We were able to successfully setup the environment for this codebase in our local machines. We then used the pretrained model parameters and performed testing on few images. As we were running it on a CPU, the execution speed was very slow. Out of around 82000 images that were used for testing, the probability masses for lanes were produced for only about 800 in 8 hours time. Figure 4 shows a sample image on which SCNN model was applied to predict the lanes and figure 5 shows corresponding lane probability masses predicted by the model. We also tried training the model on a CPU just to get an estimate of how long it was taking by the time the GPUs were being set up. But even for around 100 training and testing images, the training wasn't successful. Therefore, without GPUs, the model cannot be trained or tested efficiently.

[2]https://github.com/XingangPan/SCNN

[3]https://github.com/cardwing/Codes-for-Lane-Detection

Figure 4: Original Image

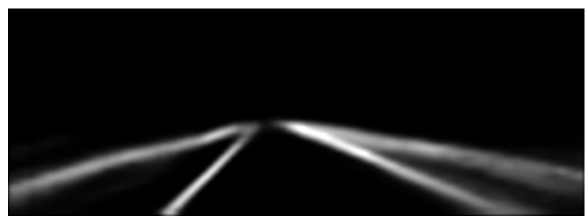

Figure 5: Probability distribution for lanes

5 Future Plan

Once the environment is setup on a machine with GPUs, we will run the SCNN model to perform training and testing. After we are able to perform training, our baseline model to evaluate the performance will be ready. After that, we will focus on our main problem as described earlier, which is focusing on one challenging condition ('glare') and improve the performance of lane detection in this scenario. For removal of dazzle light, few techniques have been proposed [Wan et al., 2017], [Yang et al., 2010]. Though, from what we see, the task of glare removal is not simple. The main challenge lies in detecting the portion of glare from the image, as it can be easily be confused with sunlight especially in road images. We have found some approaches combining interpolation and homomorphic filtering to remove the glare from normal images as shown in figures 6, 7, 8. However. from what we observe, these approaches have been tested on cases where glare region is very small and interpolation from nearby pixels is realistic, while in case of road dataset, the size of glare can be large as seen in fig 3. We plan to evaluate these procedures first, and see the improvements. [Chi et al., 2018] have also shown a method for reflection removal from single images which can also be looked upon.

Figure 6: Original Image

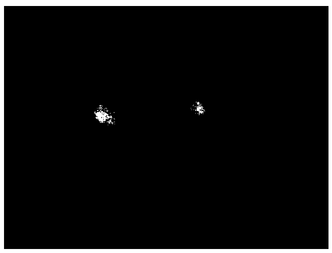

Figure 7: Glare Region

Figure 8: Final Result

There are two approaches to employ this task of dazzle light removal. The first method is to pre-process the image to detect the glare and filter it out. After filtering the glare and other noise from the image, we can then pass it to detect and track lanes. Another method is to perform lane detection end-to-end in one shot. To set up a baseline, we want to implement the first approach. Once that is done, if time permits, we will investigate on how to perform this task in one step end-to-end and evaluate both the approaches. If time permits, we also have a plan to handle different types of challenging conditions such as night and shadow creating separate models for these. For this, we propose to first classify the condition using some features such as average image intensity (night), HOG, saturation, etc,. and then apply the corresponding filtering technique to de-noise the image. The clean image can then be passed through the SCNN model to detect the lanes and hopefully improve performance. We can even combine these different models trained on different abnormalities to form an ensemble of models and evaluate for the final task. Though, we see that [Pan et al., 2018] have evaluated their model for both normal images and those in adverse conditions for lane detection, which has not been done previously, they do not do anything specific to handle these abnormalities which is what we aim to achieve through our work.

References

[Bossu et al., 2011] Bossu, J., Hautière, N., and Tarel, J.-P. (2011). Rain or snow detection in image sequences through use of a histogram of orientation of streaks. *International journal of computer vision*, 93(3):348–367.

[Chi et al., 2018] Chi, Z., Wu, X., Shu, X., and Gu, J. (2018). Single image reflection removal using deep encoder-decoder network. *arXiv preprint arXiv:1802.00094*.

[Diamond et al., 2017] Diamond, S., Sitzmann, V., Boyd, S., Wetzstein, G., and Heide, F. (2017). Dirty pixels: Optimizing image classification architectures for raw sensor data. *arXiv preprint arXiv:1701.06487*.

[Fu et al., 2017] Fu, X., Huang, J., Zeng, D., Huang, Y., Ding, X., and Paisley, J. (2017). Removing rain from single images via a deep detail network. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1715–1723.

[Hancock and Escanilla, 2018] Hancock, D. and Escanilla, N. (2018). Lane tracking in adverse visibility conditions. `https://dereklh4.github.io/lane_tracking_webpage/`.

[He et al., 2011] He, K., Sun, J., and Tang, X. (2011). Single image haze removal using dark channel prior. *IEEE transactions on pattern analysis and machine intelligence*, 33(12):2341–2353.

[Kumar and Simon, 2015] Kumar, A. M. and Simon, P. (2015). Review of lane detection and tracking algorithms in advanced driver assistance system. *Int. J. Comput. Sci. Inf. Technol*, 7(4):65–78.

[Lee and Moon, 2018] Lee, C. and Moon, J. (2018). Robust lane detection and tracking for real-time applications. *IEEE Transactions on Intelligent Transportation Systems*, 19(12):4043–4048.

[Leng and Chen, 2010] Leng, Y. and Chen, C. (2010). Vision-based lane departure detection system in urban traffic scenes. In *2010 11th International Conference on Control Automation Robotics Vision*, pages 1875–1880.

[Li et al., 2017] Li, R., Cheong, L. F., and Tan, R. T. (2017). Single image deraining using scale-aware multi-stage recurrent network. *CoRR*, abs/1712.06830.

[Nexar, 2017] Nexar (2017). Nexar challenge ii - vehicle detection in the wild using the nexet dataset. `https://www.getnexar.com/challenge-2/`.

[Pan et al., 2018] Pan, X., Shi, J., Luo, P., Wang, X., and Tang, X. (2018). Spatial as deep: Spatial cnn for traffic scene understanding. In *Thirty-Second AAAI Conference on Artificial Intelligence*.

[Parajuli et al., 2013] Parajuli, A., Celenk, M., and Riley, H. B. (2013). Robust lane detection in shadows and low illumination conditions using local gradient features. *Open Journal of Applied Sciences*, 3(01):68.

[Shen et al., 2018] Shen, L., Yue, Z., Chen, Q., Feng, F., and Ma, J. (2018). Deep joint rain and haze removal from single images. *arXiv preprint arXiv:1801.06769*.

[Shibata et al., 2014] Shibata, K., Takeuch, K., Kawai, S., and Horita, Y. (2014). Detection of road surface conditions in winter using road surveillance cameras at daytime, night-time and twilight. *International Journal of Computer Science and Network Security (IJCSNS)*, 14(11):21.

[Simonyan and Zisserman, 2014] Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

[Wan et al., 2017] Wan, R., Shi, B., Duan, L.-Y., Tan, A.-H., and Kot, A. C. (2017). Benchmarking single-image reflection removal algorithms. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 3922–3930.

[Wang et al., 2017] Wang, Y., Liu, S., Chen, C., and Zeng, B. (2017). A hierarchical approach for rain or snow removing in a single color image. *IEEE Transactions on Image Processing*, 26(8):3936–3950.

[Wang et al., 2018] Wang, Z., Ren, W., and Qiu, Q. (2018). Lanenet: Real-time lane detection networks for autonomous driving. *CoRR*, abs/1807.01726.

[Yang et al., 2010] Yang, Q., Wang, S., and Ahuja, N. (2010). Real-time specular highlight removal using bilateral filtering. In *European conference on Computer vision*, pages 87–100. Springer.