Human Detection on Road Lane For Accurate Driverless Car System Using ML & Open CV

Amrita Verma, Dr. S R Tandan
Research Scholar, Associate Professor
Department of Computer Science Engineering
Dr. C. V. Raman University
Kargi Road, Kota, Bilaspur (C.G.) 495113

Abstract*— Machine Learning & Deep learning is the future technology for many industry. Automobiles industry is one of the most important user of that technology. Now a days we are talking about the smarty car which have the driverless system. Many of the company are working of this area where they try to remove the real driver from te car. Tesla is a company who is basically talk about the driver less car system. In this paper we present a most critical & important step of the driverless car which is known as human detection on the road. Here we use the swift ML based technique and based on that we proposed an algorithm which is able to detect the human in real time on the road.

Key Words: DriverlessRoad, Lane, ML, DL, CNN, Human Detection.

I. INTRODUCTION

Image processing is one of the principle drivers of mechanization, security and well being related utilization of the electronic business. Most picture preparing advances include a few stages like treat the picture as a two-dimensional flag and apply standard sign handling procedures to it. Pictures are additionally dealt with as 3D signals where the third measurement is the time or the z-pivot. Exceptionally proficient, low memory and solid arrangements can be accomplished by using Embedded Systems and Image handling to draw out the advantages of both for applications. Google is one of the billion dollar organizations who has shown its own driverless vehicle, a structure that gets rid of every single traditional control including the guiding wheel, and other amazing innovations. In their driverless vehicle, Google has included Image Processing, yet in addition numerous other stunning advancements and one of the most significant among them is Lidar, which means "Light Detection and Ranging". It comprises of a cone or puck-formed gadget that ventures lasers which skip off items to make a high-goals guide of nature continuously. Notwithstanding assisting driverless vehicles with seeing", "Lidar is utilized to make quick, exact 3D outputs of scenes, structures, social legacy locales and foliage. A portion of different advances incorporate Bumper Mounted Radar for crash evasion, Aerial that peruses exact geo-area, Ultrasonic sensors on back wheels which recognizes and evades snags, programming which is modified to decipher basic street signs and so forth. Aside from these, there are altimeters, whirligigs, and tachymeters that decide the exact situation of the vehicle and offers exceptionally precise information for the vehicle to work securely. The synergistic joining of sensors is one of the most significant factors in this self-ruling vehicle which incorporates the information assembled out and out by these sensors are ordered and deciphered by the vehicle's CPU or in manufactured programming framework to make a protected driving encounter. Aside from Google, numerous different organizations like Tesla, Audi, Uber have additionally built up their own driverless vehicles and have tried conceivably. Picture handling is one of the standard drivers of automation, security and prosperity related usage of the electronic business. Most picture planning progresses incorporate a couple of stages like treat the image as a two-dimensional banner and apply standard sign taking care of systems to it. Pictures are also managed as 3D signals where the third estimation is the time or the z-rotate Human detection is a logical next step after the development of successful face detection algorithms. However, humans have been proven to be a much more difficult object to detect because of the wide variability in appearance due to clothing, articulation and illumination conditions that are common in outdoor scenes. Recently, Dalal & Triggs [1] presented a human detection algorithm with excellent detection results. Their method uses a dense grid of Histograms of Oriented Gradients (HoG), computed over blocks of size 16×16 pixels to represent a detection window. This representation is proved to be powerful enough to classify humans using a linear SVM. Unfortunately, their method can only process 320×240 images at 1 FPS using a very sparse scanning methodology that evaluates roughly 800 detection windows per image. We speed up their method, while increasing the number of detection windows for evaluation from 800 to 12, 800. The improvement is achieved by combining the cascade of rejector approach that is extensively used for face detection [13, 11] with the HoG features. However, we discovered that the use of fixed-size blocks, advocated by Dalal & Triggs is not informative enough to allow fast rejection in the early stages of the cascade. We therefore design a much larger set of blocks that vary in sizes, locations and aspect ratios. We then use AdaBoost to select the best blocks suited for detection and construct the rejector-based cascade. This results in a near real-time human detection system that matches existing methods in terms of accuracy and significantly outperforms them in terms of speed. Vital writing overview related past research on leaf deficiency identification are given in II recognition based past work are given in segment ii though area III portrays explore issue and future degree philosophy and IMPLEMENTATION FOR THE PREVIOUS EXISTING APPROACHES. IV portrays philosophy and IMPLEMENTATION FOR THE PREVIOUS EXISTING APPROACHES. Trial results and its examination are given in area V. At long last, area VI closes the paper.

II. LITRECTURE REVIEW

There give off an impression of being two driving ways to deal with the issue of human location. It would be ideal if you allude to [3] for a definite overview. One methodology utilizes a solitary location window examination though the other methodology utilizes a sections based methodology. Inside every technique, various creators offer various highlights and various classifiers to handle the issue. Under the single-discovery window approach, the work of Papgeorgiou and Poggio [8] utilizes Haar-based portrayal, joined with a polynomial SVM. Crafted by Gavrila and Philomin [4] contrast edge pictures with a model dataset utilizing the chamfer distance. Viola et al. [14] stretched out their Haar-like wavelets to deal with space-time data for moving-human discovery. Others have adopted a sections based strategy that focuses on managing the extraordinary fluctuation in appearance because of body explanation. In this methodology, each part is identified independently and a human is recognized assuming a few or the entirety of its parts are introduced in a mathematically conceivable setup. Felzenswalb and Huttenlocher [2] utilize pictorial structure approach where an article is portrayed by its parts, associated with springs, and speak to each leave behind Gaussian subordinate channels of various scale and direction. Ioffe and Forsyth [5] speak to parts as projections of straight chambers and propose proficient approaches to steadily gather these sections into a full body get together. Mikolajczyk et al. [7] speak to parts as co-events of neighborhood direction highlights. The framework continues by distinguishing highlights, at that point parts furthermore, at last people are identified dependent on congregations of parts. Dalal and Triggs [1] utilized the single window approach with a thick HoG portrayal that was effectively utilized for object portrayal [6, 10, 7]. 3 The Dalal-Triggs Algorithm We start with a short depiction of the Dalal and Triggs calculation. Every discovery window is separated into cells of size 8×8 pixels and each gathering of 2×2 cells is incorporated into a square in a sliding design, so obstructs cover with each other. Every cell comprises of a 9-receptacle Histogram of Oriented Angles (HoG) and each square contains a linked vector of every one of its cells. Each square is along these lines spoken to by a $36 - D$ element vector that is standardized to a L2 unit length. Each 64×128 identification window is spoken to by 7×15 squares, giving a sum of 3780 highlights for every location window. These highlights are then used to prepare a direct SVM classifier. 4 A Fast Human Detection Framework The Dalal and Triggs calculation utilizes three key segments: (1) the utilization of HoG as an essential structure block, (2) the utilization of a thick framework of HoGs across the whole discovery window to give a decent portrayal of the identification window, and (3) a standardization venture inside each square that stresses relative conduct, as for the neighboring cells, instead of the total qualities. A significant factor, that is absent in their methodology, is the utilization of squares in various scales. They utilize minuscule block size (normally, 16×16 pixels) which may miss the "higher perspective", or worldwide highlights of the whole identification window. Surely, they report that adding blocks/cells of various scales would fairly improve the outcomes yet would additionally fundamentally increment the calculation cost. The catching of the "higher perspective" is in this manner depended on the thick set of limited scope blocks across the whole identification window. To quicken the discovery cycle we utilize a course of rejectors and use AdaBoost to pick which highlights to assess in each stage, where each component compares to one square. Notwithstanding, the little size of the squares is demonstrated to be a significant obstruction. We found that none of these little size blocks was sufficiently educational to dismiss enough examples to quicken the discovery cycle. Subsequently, we increment our element space to incorporate squares of various sizes, areas and angle proportions. Therefore, we have 5, 031 squares to browse, contrasted with the 105 squares utilized in the Dalal-Triggs calculation. Besides, we found that the initial not many phases of the course, that dismisses most of identification windows, really utilize huge squares and the little squares are utilized a lot later in the course. To help the quick assessment of explicit squares, as are picked by our AdaBoost-based component determination calculation, we utilize the essential picture portrayal to proficiently register the HoG of each square. It merits referencing that Viola et al. [14] utilize a comparable structure for identifying people in an observation climate, where individuals to be distinguished are little also, as a rule have an unmistakable foundation (street, divider, and so forth) Be that as it may, as asserted in their paper, the location execution extraordinarily depends on the accessible movement data. Notwithstanding, for the Dalal-Triggs' INRIA information base which contains incredibly muddled foundations and emotional enlightenment changes, the Harr-wavelet include accomplishes a much lower location precision than that of the HoG include.

III. RESEARCH ISSUE

In this section basically we talk about research gap which need to be solved,present research there is lots of issues are there which followings are:

1. Lack in Accuracy: As we know human detection is the most important process for driverless car but most of the approaches are nit up to the mark.
2. Time complexity: In existing solution time complexity is main challenge because human detection should be complete in few second.
3. False Detection: Most of the algorithms are not capable to track the exact presence of the human.
4. Time & Quality management issue: There is no any approach which is able to make justice with both parameters.
5. Accuracy: There is lack of accuracy in most of the previous existing approach

As per the previous research there is lots of research gap which need to be solved in near future

IV. METHADOLOGY & IMPLEMENTAION

In this section we talk about the basic leaf fault detection process, what kind of basic algorithm was used and what are the advance research is there. Here we did the complete comparative study and implementation of those approaches.

4.1 Qiang:

Histograms of Oriented Gradients (HoG) highlights to accomplish a quick and precise human identification framework. The highlights utilized in our framework are HoGs of variable-size impedes that catch notable highlights of people consequently. Utilizing AdaBoost for include determination, we recognize the fitting arrangement of squares, from a huge arrangement of potential squares. In this framework, we utilize the vital picture portrayal and a dismissal course which essentially accelerate the calculation.

Algorithm 1 Training the cascade

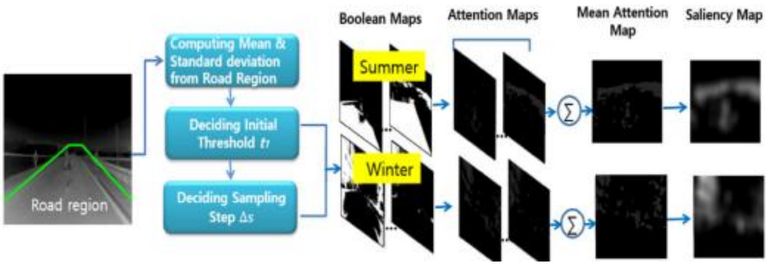
Input: F_{target}: target overall false positive rate
f_{max}: maximum acceptable false positive rate per cascade level
d_{min}: minimum acceptable detection per cascade level
Pos: set of positive samples
Neg: set of negative samples

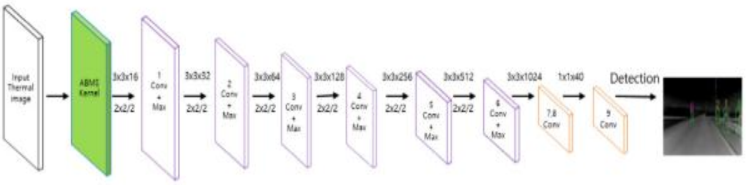
initialize: i = 0, $D_i = 1.0$, $F_i = 1.0$
loop $F_i > F_{target}$
$i = i + 1$
$f_i = 1.0$
loop $f_i > f_{max}$
1) train 250 (%5 at random) linear SVMs using Pos and Neg samples
2) add the best SVM into the strong classifier, update the weight in AdaBoost manner
3) evaluate Pos and Neg by current strong classifier
4) decrease threshold until d_{min} holds
5) compute f_i under this threshold
loop end
$F_{i+1} = F_i \times f_i$
$D_{i+1} = D_i \times d_{min}$
Empty set Neg
if $F_i > F_{target}$ then evaluate the current cascaded detector on the negative,i.e. non-human,images and add misclassified samples into set Neg.
loop end

Output: A i-levels cascade
each level has a boosted classifier of SVMs
Final training accuracy: F_i and D_i

4.2 Duyoung:

In this approach author use YOLO process and based on that here they use to train the human detection using the tensor flow approach and based on that they find the human detection.





Algorithm 1: Adaptive BMS

A set of Boolean maps $\mathbf{B} = \{\}$; A set of attention maps $\mathbf{A} = \{\}$; $\bar{A} \leftarrow 0$
(1) Input feature map $\emptyset(I)$ is generated from a thermal image
Compute the initial threshold t_1 and sampling step Δs using Equations (3) and (4)
For $i = 1$ to N // refer to the results in Figure 1(b), (c)
For $\theta = t_1$ to 255
B_i = **THRESH**$(\emptyset(I), \theta)$
$\bar{B}_i$ = **INVERT**(B_i)
Morphological opening to B_i and $\bar{B}_i$
Add B_i and $\bar{B}_i$ to **B**
End for
End for
(2) For $i = 1$ to N // refer to the results in Figure 1(d)
Set $A_i(x, y) = 0$ if all pixels of $B_i(x, y)$ are connected to the image borders
Morphological dilation to A_i
Normalization A_i
$\bar{A} \leftarrow \bar{A} + A_i$
End for
(3) $\bar{A} \leftarrow \bar{A}/\max_i A_i$ // refer to the results in Figure 1(e)
(4) S← $PostProcessin(\bar{A})$ // refer to the results in Figure 1(f)

4.3 Joko:

In this approach author first find the road area, and based on that they check if human cross that area so they check the human walking human model and based on that they decide the localization of human body and art final they give the final result regarding human on road.

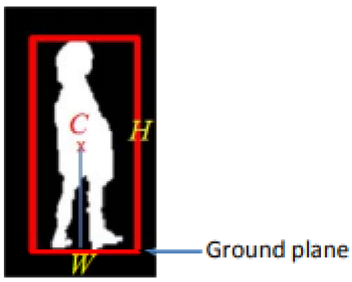


Fig. 4.1 Human Model

Fig. 4.2 Road Estimation

Fig. 4.3 Human Detection on Road

Proposed Approach:

Here we present a fast-processing step using the open cv and training base model here we use the swift ML approach. These are the followings steps which we follow for the human detection on the road:

1. Road Lane Detection
    a. Real Time Video Conversion on grayscale
    b. Color Classification
    c. Morphology
    d. Thresholding
    e. Road Lane Marking
2. Human Body Detection:
    a. Human Upper Body Data Base Creation
    b. Human Lower Body Data Base Creation
    c. Human Face Database Creation
    d. Upper Body Training
    e. Lower Body Training
    f. Face Training
    g. Deep Analysis On human Real time Footage Using Train Data
    h. Final Human Body Estimation
3. Mearing Road Lane & Human Body Detection:
    a. Real time Analysis of Both parameters
    b. Marking Road Lane
    c. Marking detected human body
    d. Signal generation based on human detection for driver less car.

V. Result & Analysis

In this section we introduce the relative investigation of all with past existing methodology. As per the driver less car these are the main feature which should be there, Here we use Accuracy, time & Quality as a parameter through that we got the comparative analysis.

Table 5.1 Comparative Analysis for Time

Approach	Time
Qiang	0.10 Sec
Duyoung	0.22 Sec
Joko	0.12 Sec
Proposed	0.067 Sec

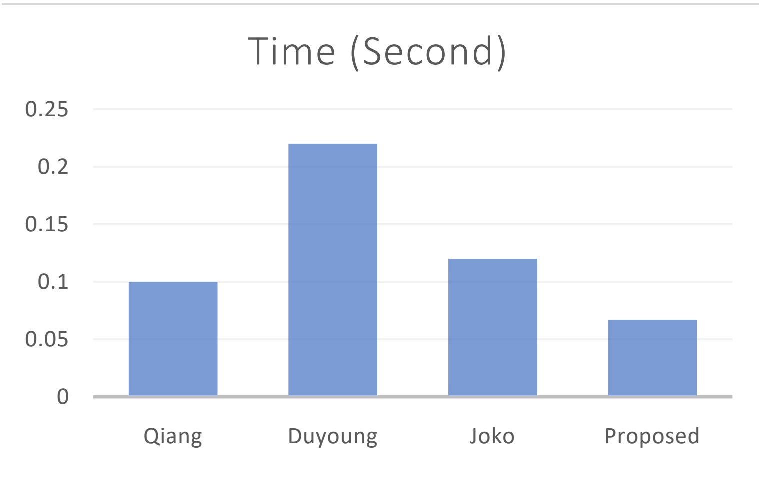


Fig. 5.2 Comparative analysis

Approach	Accuracy	Quality	Efficiency
Qiang	Mid	Low	Avg
Duyoung	Low	Mid	Mid
Joko	Avg	Mid	Avg
Proposed	High	Mid	High

As per the analysis we can see our proposed approach is far better than other approaches in terms of of the time & efficiency.

VI. Conclusion

In present era driverless car is the next future for everyone, as we know the cost of driver is too much if we are talking about the business level model and the most important things there is lots of dangerous part is involve if any human run any machine there is lots of chances any accident can happen. In this paper we present a most important part which is known as human detection on road. So here we use the Swift ML technique using this approach we are able to get a good result in terms of the efficiency & time complexity.

REFERENCES

[1] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. Coference on Computer Vision and Pattern Recognition (CVPR), 2005.

[2] P. Felzenszwalb and D. Huttenlocher. Pictorial structures for object recognition. Internation Journal of Computer Vision (IJCV), 61(1):55–79, 2005.

[3] D. M. Gavrila. The visual analysis of human movement: A survey. Journal of Computer Vision and Image Understanding (CVIU), 73(1):82–98, 1999.

[4] D. M. Gavrila and V. Philomin. Real-time object detection for smart vehicles. Coference on Computer Vision and Pattern Recognition (CVPR), 1999.

[5] S. Ioffe and D. Forsyth. Probabilistic methods for finding people. International Journal of Computer Vision (IJCV), 43(1):45–68, 2001.

[6] D. Lowe. Distinctive image features from scale-invariant keypoints. Internation Journal of Computer Vision (IJCV), 60(2):91–110, 2004.

[7] K. Mikolajczyk, C. Schmid, and A. Zisserman. Human detection based on a probabilistic assembly of robust part detectors. European Conference on Computer Vision (ECCV), 2004.

[8] C. Papageorgiou and T. Poggio. A trainable system for object detection. Internation Journal of Computer Vision (IJCV), 38(1):15–33, 2000.

[9] F. Porikli. Integral histogram: A fast way to extract higtograms in cartesian spaces. Coference on Computer Vision and Pattern Recognition (CVPR), 2005.

[10] J. M. S. Belongie and J. Puzicha. Shape matching object recongition using shape contexts. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 24(24):509–522, 2002.

[11] H. Schneiderman. Feature-centric evaluation for efficient cascaded object detection. Conference on Computer Vision and Pattern Recognition (CVPR), 2004.

[12] B. Scholkopf and A. Smola. Learning with Kernels Support Vector Machines, Regularization, Optimization and Beyond. MIT Press, Cambridge, MA, 2002.

[13] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. Coference on Computer Vision and Pattern Recognition (CVPR), 2001.

[14] P. Viola, M. Jones, and D. Snow. Detecting pedestrians using patterons of motion and appearance. International Coference on Computer Vision (ICCV), 2003.

[15] A. Prioletti, A. Mogelmose, P. Grislieri, M. Trivedi, A. Broggi, and TB.Meslund, "Part-based Pedestrian Detection and Feature-based Tracking for Driver Assistance: Real-Time, Robust Algorithms and Evaluation," IEEE Trans. ITS, Vol. 14(3), pp. 1346-1359, 2013.

[16] YW. Xu, D. Xu, S. Lin, XT. Han, XB. Cao and XL. Li, "Detection of Sudden Pedestrian Crossings for Driving Assistance Systems," IEEE Trans. SMC (Part B), Vol. 42 (3), pp.729- 739, 2012.

[17] N. Dalal, B. Triggs, "Histograms of Oriented Gradients for Human Detection," IEEE Proc. CVPR, pp. 886-893, 2005.

[18] J. Hariyono, V. D. Hoang, and K. H. Jo, Location Classification of Detected Pedestrian, Proc. ICCAS, Seoul, 2014. [5] C. Tomasi and T. Kanade, "Detection and Tracking of Point Features," International Journal of Computer Vision, vol. 9, pp. 137-154, 1991.

[19] T. Kobayasi, A. Hidaka, and T. Kurita, "Selection of Histograms of Oriented Gradients Features for Pedestrian Detection," Proc. ICONIP, pp. 598-607, 2008.

[20] X. Wang, X. Han, and S. Yan, "An HOG-LBP human detector with partial occlusion handling," Proc. CVPR, 2009.

[21] J. Hariyono, V.D. Hoang, and K.H. Jo, "Moving Object Localization using Optical Flow for Pedestrian Detection from a Moving Vehicle," The Scientific World Journal, Volume 2014, 2014.

[22] J.F.P. Kooij, N. Schneider, F. Flohr, and D.M. Gavrila, "Context-based Pedestrian Path Prediction," Proc. ECCV, pp. 618- 633, Zurich, 2014.

[23] V.D. Hoang, L.M. Ha, and K.H. Jo, "Hybrid Cascade Boosting Machine using Variant Scale Blocks based HOG Features for Pedestrian Detection," Neurocomputing, Vol.135, pp.357- 366, 7 2014.

[24] P. Dollar, C. Wojek, B. Schiele, and P. Perona. "Pedestrian detection: an evaluation of the state of the art," IEEE Trans.PAMI., Vol. 34(4), pp. 743 – 761, 2012.

[25] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollar. Focal Loss for Dense Object Detection. In ´ Conference on Computer Vision and Pattern Recognition (CVPR), 2017. 4

[26] Tong Liu, Zhaowei Chen, Yi Yang, Zehao Wu, and Haowei Li. Lane Detection in Low-light Conditions Using an Efficient Data Enhancement: Light Conditions Style Transfer. In Intelligent Vehicles Symposium (IV), 2020. 2, 7

[27] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C Berg. SSD: Single Shot Multibox Detector. In European Conference on Computer Vision (ECCV). Springer, 2016. 2

[28] Xingang Pan, Jianping Shi, Ping Luo, Xiaogang Wang, and Xiaoou Tang. Spatial As Deep: Spatial CNN for Traffic Scene Understanding. In AAAI, February 2018. 1, 2, 5, 6, 7

[29] Jonah Philion. FastDraw: Addressing the Long Tail of Lane Detection by Adapting a Sequential Prediction Network. In Conference on Computer Vision and Pattern Recognition (CVPR), 2019. 1, 2, 6, 7

[30] Fabio Pizzati, Marco Allodi, Alejandro Barrera, and Fernando Garc´ıa. Lane Detection and Classification using Cascaded CNNs. In International Conference on Computer Aided Systems Theory, 2019. 6

[31] Zequn Qin, Huanyu Wang, and Xi Li. Ultra Fast Structureaware Deep Lane Detection. In European Conference on Computer Vision (ECCV), 2020. 2, 6, 7

[32] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You Only Look Once: Unified, Real-Time Object Detection. In Conference on Computer Vision and Pattern Recognition (CVPR), 2016. 2

[33] Eduardo Romera, Jose M Alvarez, Luis M Bergasa, and ´ Roberto Arroyo. ERFNet: Efficient Residual Factorized ConvNet for Real-Time Semantic Segmentation. Transactions on Intelligent Transportation Systems, 19(1):263–272, 2017. 1

[34] Lucas Tabelini, Rodrigo Berriel, Thiago M. Paixao, Claudine ˜ Badue, Alberto F. De Souza, and Thiago Oliveira-Santos. PolyLaneNet: Lane Estimation via Deep Polynomial Regression. In ICPR, 2020. 1, 2, 6, 7, 8

[35] TuSimple. Tusimple benchmark. https://github. com/TuSimple/tusimple- benchmark. Accessed September, 2020. 1, 5

[36] Wouter Van Gansbeke, Bert De Brabandere, Davy Neven, Marc Proesmans, and Luc Van Gool. End-to-end Lane Detection through Differentiable Least-Squares Fitting. In ICCV Workshop, 2019. 6

[37] Hang Xu, Shaoju Wang, Xinyue Cai, Wei Zhang, Xiaodan Liang, and Zhenguo Li. CurveLane-NAS: Unifying LaneSensitive Architecture Search and Adaptive Point Blending. In European Conference on Computer Vision (ECCV), 2020. 2, 7

[38] Seungwoo Yoo, Hee Seok Lee, Heesoo Myeong, Sungrack Yun, Hyoungwoo Park, Janghoon Cho, and Duck Hoon Kim. End-to-End Lane Marker Detection via Row-wise Classification. In IEEE CVPR Workshop, 2020. 2, 6, 7

[39]Li, Zhenguo. "CurveLane-NAS: Unifying Lane-Sensitive Architecture Search and Adaptive Point Blending." (2020).

[40]Tabelini, Lucas, et al. "Keep your Eyes on the Lane: Attention-guided Lane Detection." arXiv preprint arXiv:2010.12035 (2020).

[41]Ko, Yeongmin, et al. "Key Points Estimation and Point Instance Segmentation Approach for Lane Detection." arXiv preprint arXiv:2002.06604 (2020).

[42] Pan, Xingang, et al. "Spatial as deep: Spatial cnn for traffic scene understanding." arXiv preprint arXiv:1712.06080 (2017).