2018 IEEE Intelligent Vehicles Symposium (IV)
Changshu, Suzhou, China, June 26-30, 2018

CyDet: Improving Camera-based Cyclist Recognition Accuracy with Known Cycling Jersey Patterns

Alexander Masalov[1], Jeffrey Ota[2], Heath Corbet[3], Eric Lee[3], Adam Pelley[3]
[1]Winkam, [2]Intel Corporation, [3]Specialized Bicycles

Abstract— **In this work, we propose CyDet, a hybrid approach to detection and classification based on human knowledge of the scenario context and statistical machine learning approaches to object detection, classification, and tracking. This approach enabled us to experiment with five different methods to maximize the accuracy of cyclist detection. These methods included detecting a cyclist as a single object, utilizing different classifiers for different view angles, using HOG and LBP with fewer descriptors and features, using a machine learning layer for classifier training, and uniquely, custom designed jerseys by Specialized Bicycles. By working closely with Specialized, we were able to ensure that designs were both optimized for computer vision algorithm detection and productizable, and we were able to train our algorithms to classify the known designs as cyclists which further improved our accuracy, especially when the cyclist is partially occluded. Our experiments show that our approach achieves state of the art accuracy on the KITTI [1] benchmark, and we establish a baseline detection rate with our approach on a new, open Specialized Bicycles Cyclist Detection Dataset that include cyclists wearing jerseys with known patterns.**

I. Introduction

Historically, traditional advanced driver assistance algorithm system (ADAS) research was primarily focused on highway driving scenarios and specific collision scenarios which involved pedestrians or other vehicles. These were targeted for what is now known as "Level 2: Partial Automation" in the Society of Automotive Engineers (SAE) automation levels recognized by the United States Department of Transportation. When artificial intelligence (AI) or deep learning computer vision approaches accelerated the possibility to jump to Levels 3, 4, and 5, the dependence on training data became more important than the first principles of the traditional approach to computer vision algorithm design. Other than Volvo's cyclist detection system that utilized both the camera and radar specifically for their autobraking system [2] and the Cho et al approach presented at IV 2010 [3], camera-based cyclist detection have largely been overlooked in what has become a race to Levels 4 and 5 ("high" and "full" automation respectively).

Recent reports of current state of the art systems almost ignoring cyclists [4] and creating dangerous situations while in proximity with cyclists has raised the alarm in the industry and in the popular press to focus more on cyclist detection [5] [6].

The importance of detecting a cyclist and classifying it correctly as a cyclist and not as a generic object is encapsulated in the future free space prediction step in autonomous driving. Typically, after objects are detected and classified, an environment model surrounding the car representing free and occupied space is built utilizing what is known as an occupancy grid [7]. Radar, LiDAR, ultrasound, and camera-based information including an object list is combined into the occupancy grid to determine what space might be free to drive in and what might be occupied by solid objects, which the car would not want to hit. Once the current state of the occupancy grid is built, a prediction of what the occupancy grid could evolve to is generated to determine what potential free space will be available in the future and therefore, allow the path planning step of autonomous driving select the safest possible future route to take that minimizes the possibility of a collision. This step is highly dependent on the accurate detection and classification of the objects in the scene.

The reason for this is that different objects such as a pedestrian versus a cyclist has the potential to occupy very different paths in the future. Pedestrians are fairly predictable and have a limited range they can be in the next five seconds for example. However, cyclists, depending on their current speed, can quickly turn left or right at anytime which makes the potential space they occupy very different than the pedestrians which means the occupancy grid will be updated very differently. Therefore, in order to increase safety for cyclists, it is imperative to increase the accuracy in detecting and classifying cyclists, so that the autonomous car correctly selects the safest path that minimizes the potential collision with a cyclist.

The rest of the paper discusses related work then introduces the thinking behind and the approach we took with CyDet. We share our experimental results with the KITTI dataset and with a new Specialized Bicycles Cyclist Detection Dataset with and without the custom cycling jerseys by Specialized and discuss what future work we plan to do beyond the research discussed in this paper.

II. Related Work

With LiDAR considered one of the mandatory sensors for autonomous vehicles, object recognition research using only 3D point cloud LiDAR data such as VoxelNet [8] and other multi-sensor fusion-based algorithms have been more recently explored. In our research, we focus only on maximizing the capabilities of a monocular camera such as that by Cho et al. [3] performed in 2010. In Cho et al, their fast bicyclist detector utilized a Histogram Oriented Gradients (HOG) based detector for frontal, rear, and right/left side views. They then utilized a cascade of classifiers including a linear Support Vector Machine (SVM) and AdaBoost for the classification step. This thorough, pre-machine learning revolution approach utilized the most state-of-the art traditional computer vision approaches at the time yet it had a hit rate of only 65% and had many false positives.

One of the more recent monocular camera approaches is Deep3DBox [9], which utilizes modern deep learning techniques and attempts to predict which way moving vehicles are facing and inferring a 3D box around each object identified in a 2D image. However, even this novel approach only achieves an identification rate of 74%.

Other numerous approaches focus on improving the accuracy of individual people and bicycles and then use proximity and geometry rules to conclude if the person is a pedestrian or a cyclist, but those suffer from many false positives as people standing near bicycles in urban settings are identified as cyclists.

The first stage of our method focuses on a hybrid approach that combines HOG [10] and a local binary pattern (LBP) operator inspired by Timo Ojala et al. [11] using uniform 59 descriptors. We then optimized the blurring approach and utilized modern machine learning techniques to classify the cyclist as a whole object (bike plus cyclist). We train the classifier for all possible orientations of the cyclist and show improvements in detection rates and accuracy. [12] [13].

In this CyDet implementation we used a statistical machine learning layer using AdaBoost [14], which allowed us to achieve an ultra small model size of only 0.0283 MB. In the next iteration of this algorithm we are planning to test a CNN as an alternative machine learning (ML) layer in a hybrid approach.

As Zdenek, Kalal et al. demonstrate [15], tracking against unstable backgrounds in video is a challenging task. CyDet uses a simple tracking method optimized for unstable background video, which allows us to achieve an increase of 5–10% True Positive Rate (Recall) as well as minimize recognition delay due to tracking to just 2 to 4 frames.

To run CyDet tests we utilized the KITTI raw data set [1] as it is currently the standard for detection classifier tests in autonomous driving. We tested CyDet accuracy results against the CNN detector YOLO [16] and SqueezeDet [17].

Finally, working with Specialized Bicycles, we created known cyclist jersey patterns that were designed to be easy to see by computer vision filters to increase the accuracy further, especially in situations where the bicycle is partially occluded and only the jersey can be seen clearly. Specialized Bicycles collected video with these jerseys being worn, and the video dataset will be open sourced to the community to enable further research that utilizes prior knowledge of clothing patterns on cyclists.

III. Method Description

To solve the issue of accurate cyclist recognition for increased safety, the CyDet cyclist detection classifier is proposed in this research. In the CyDet cyclist detector we use our Winkam 99.9 Hybrid method that incorporates both hand-crafted knowledge layers as well as ML layers (see

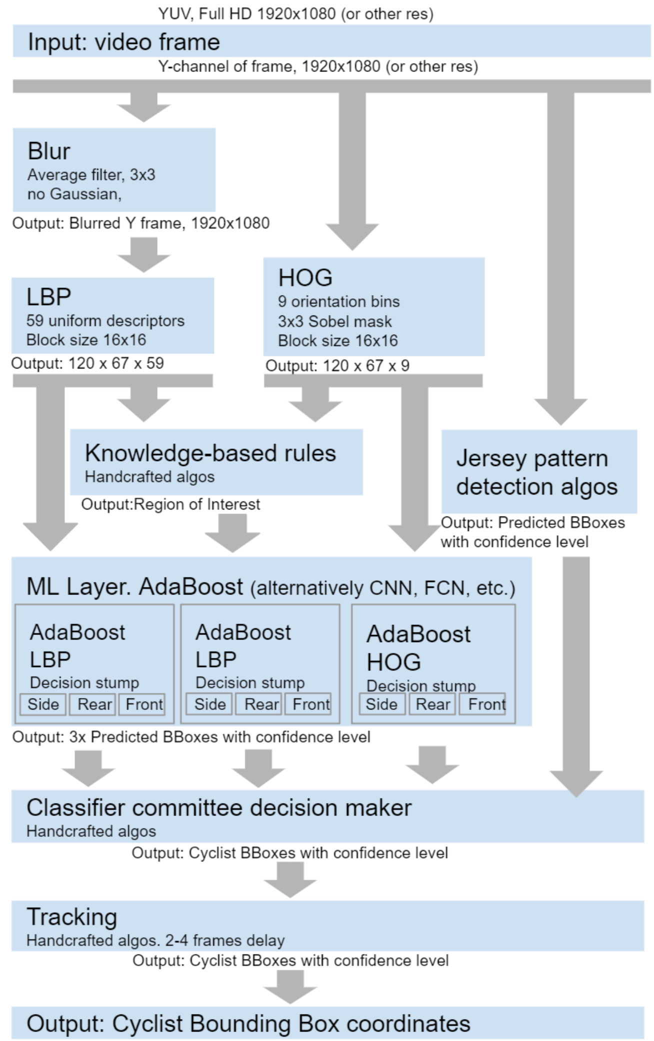

Fig. 1. CyDet pipeline utilizing WINKAM^{TM} 99.9 Hybrid pipeline model

Figure 1). We used a hybrid approach as it gives us opportunity to utilize smaller data sets, reduces risk of overfitting, increases classifier detection robustness, and results in small model sizes that have high real-time computation speed. All this allowed us achieve state of art accuracy higher than standard CNN methods. See Figures 8, 9 and Tables I, II.

To solve the issue of high accuracy detection in scenarios where the cyclist or bicycle is partially obscured, we propose using a known graphic pattern printed on the cyclist jersey. The pattern was intentionally designed by Specialized to ensure higher visibility and detection accuracy. The CyDet classifier is trained to recognize known patterns and uses detection of the pattern to supplement its own decisions in conditions of reduced visibility. These methods are utilized as they allow us to achieve state-of-the-art cyclist recognition accuracy and maintain detection precision in settings where the cyclist or bicycle are obscured or occluded.

Key challenges we solve with CyDet. Conditions of obscurity and poor visibility constitute the main challenge to recognition. Below are the most common types of reduced visibility conditions.

1) Physical occlusion by vehicles, other cyclists, lamp posts, traffic lights, signpost, trees, shrubbery.
2) Reduced visibility due to unfavorable light conditions such as glare, shadows, twilight.
3) Reduced visibility due to cyclist being located far from the camera.

Ultrasmall model size and real-time performance. The size of the CyDet model is 0.0283 MB which is 279 times smaller than SqueezeDet's 7.9 MB [17]. The ultrasmall model also enables cost-effective implementation to FPGAs which allows for extremely energy efficient real-time detectors. Moreover, ultrasmall model size allows more control over classifier logic, enabling humans to directly review and influence classifier decisions based on context. Therefore, it provides the opportunity to optimize algorithms for both higher accuracy under different weather, light, and obscurity conditions and real-time performance.

On a 35W Intel Core i7-7820HQ, 16GB RAM, Ubuntu 16.04 64-bit system, CyDet's performance is:

- 60 FPS (+/-15%) 1242x374 KITTI format video
- 40 FPS (+/-15%) 1280x720 HD video
- 15 FPS (+/-15%) 1920x1080 FullHD video

Detecting a cyclist as a single object. To increase recognition quality, the cyclist is recognized as a single object as opposed to a compound object comprised of a bicycle and human. This method enables the emphasis of cyclist-specific recognition and helps eliminate false positives with parked bicycles or pedestrians standing nearby. This method also works best for scenarios where several cyclists are present.

LBP and HOG synergy. LBP transforms gray-scale input image into image of descriptors by comparing each pixel of input image with its neighbor. HOG transforms gray-scale input image into image of gradients. HOG calculates orient and magnitude of gradient for each pixel. The statistics of LBP and HOG descriptors are used as features vector for classification.

HOG is a natural partner to LBP [18] as it uses not only qualitative brightness changes, i.e. direction, but also quantitative, i.e. magnitude. LBP conversely does not use magnitude, and as a result LBP exhibits little dependency on lighting, but it can detect any pattern with low noise. LBP identifies textures in the frame well while HOG can easily spot road signs, building edges or street lamps. The image with a cyclist after LBP and HOG transformations are shown in Figures 2 and 3.

LBP uniform scheme is used with 59 bins [11] [19], 16x16 blocks, with a 16pix stride. A small quantity of descriptors, as opposed to classic LBP, is chosen to decrease model size and lower the risk of overfitting. The advantage of LBP is resilience to changes in light conditions, both in terms of brightness and spectrum. We apply LBP to grayscale images.

A special feature of HOG [10] [18] is its small number of descriptors, of which there are 9, 16x16 blocks, with a 16pix stride. As a result, the number of features (the size of the feature map) is low, which in turn also lowers the risk of classifier overfitting and enables the use of a smaller model size.

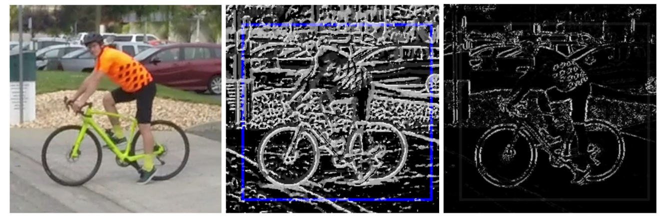

Fig. 2. Specialized Jersey with Diamonds pattern. FullColor, after LBP and HOG transformation. Side view.

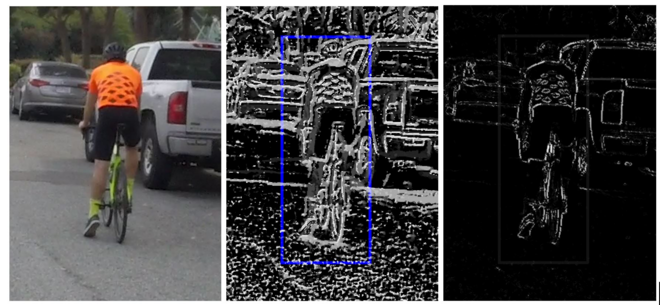

Fig. 3. Specialized Jersey with Diamonds pattern. FullColor, after LBP and HOG transformation. Rear view.

Using the ML layer for classifier training. AdaBoost is selected as the ML classifier model as it allows easy control over classifier size via setting a certain number of iterations and thus limiting overtraining and significantly reducing training time costs. Decision Stump is used as a weak classifier in AdaBoost. We train the AdaBoost HOG and AdaBoost LBP classifiers separately.

An alternative ML layer can also be utilized such as SVM, CNN, or other methods of statistical learning.

Utilizing different classifiers for different view angles. We utilized separate classifiers for different camera view angles such as the side, rear, and front of the cyclist [13]. This method allows us to run higher quality tests on different scenarios and target the improvement of specific classifier parts using additional samples.

Tracking. For cyclist recognition for autonomous driving we must consider real road conditions. It is therefore more appropriate to the train and test the classifier on video footage with an unstable background as opposed to separate images. Therefore, the CyDet basic detection is built on separate images and the final tracking layer considers previous frame features to increase recognition quality [15].

An extra tracking layer enables higher recognition accuracy. Following frame feature readings increases both Recall (TPR) and Precision (PPV). If the classifier detects an object in frame i, but not in i+1 and i-1, the detection event for frame i is canceled.

This tracking approach leads to minor recognition delays of 1-n frames as the object in frame i will be recognized only when i+1 (i+n) frame is processed. In CyDet we used

Fig. 4. Specialized Jersey with Diamonds pattern. FullColor. Rear, side, front view

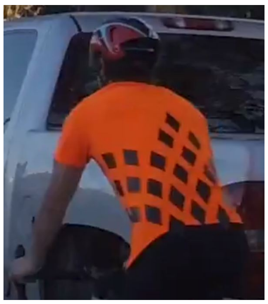

Fig. 5. Specialized Jersey with Diamonds pattern. FullColor. Rear

n = 4. Therefore minor delay in cyclist detection from 2-4 frames may occur.

Using special jersey graphics patterns. The custom Specialized jersey graphics in Figures 4 and 5 were used to train a classifier to recognize cyclists in scenarios and contexts of reduced recognizability. Even if the bicycle itself is out of view, the object can be recognized as a cyclist provided they are wearing a specially patterned jersey.

We tested and scaled different types of jersey designs to find graphic patterns and colors that yield the highest cyclist recognition accuracy improvement. In this article we demonstrate test results for one of the chosen jersey graphic patterns that provides a significant positive impact on cyclist recognition accuracy

For this paper, the diamonds pattern jersey design by Specialized Bicycles was chosen. Winkam's pattern detection algorithm was used for the chosen jersey pattern texture. Diamonds are themselves descriptors. Given sufficient training data, diamonds statistically influence classifiers trained on cyclist bounding boxes with known jersey patterns.

Winkam's jersey graphics pattern detection algorithm was added as a separate module. Additional jersey patterns can be added to this module without having to retrain all the other classifier parts. Thus this module is independent and serves as a second classifier.

The jersey graphics pattern detection algorithm extracts features from the gray-scale image of the jersey pattern. Then hand-crafted rules are used to recognize the cyclist in jersey with Specialized pattern. The algorithm then outputs bounding boxes with coordinates of where the jersey pattern was recognized in the frame.

IV. EXPERIMENTS

The aim of our research was to provide evidence that jersey patterns, regardless of color, increase recognition quality of cyclists by self-driving cars. The research experiment process was split into three stages:

1) **Prepare validation data sets.** For testing and comparative analysis we prepared two validation datasets. These datasets present detailed city riding scenarios where a self-driving car may collide with a cyclist.
    a) The KITTI dataset [1] that was developed specifically for evaluation of recognition quality by self-driving cars on city streets.
    b) And a new Specialized Cyclist Detection Dataset where one of the cyclists is wearing a Specialized graphic pattern jersey specially prepared for tests.
2) **Cyclist classifier testing preparation.** We then selected SqueezeDet [17] and Yolo [16] for comparison to our new CyDet classifier. CyDet was pre-trained using part of the chosen data sets. SqueezeDet and Yolo's default distribution were utilized in this paper.
3) **Testing.** Finally, we conducted tests and collected results from CyDet, SqueezeDet, and Yolo on the KITTI and Specialized datasets by measuring recognition accuracy both for all cyclists and for cyclists wearing the special graphic pattern jersey.

One of the key challenges and limitations of the experiments was the selection of the best validation dataset. To return valid results, a validation dataset (in vitro) should represent diverse and true to life road conditions (in vivo), such as different cyclist riding scenarios, different cyclist view angles, colors, light conditions, distances between cyclists and self-driving vehicles and cyclists obscured by different objects. It is also important to run tests on video as opposed to images as in reality, all objects are in motion. However, effects such as glare, motion blur, or shaking of the camera may impede the process and make it more difficult to split the data into training and testing datasets.

To address real-life situations with cyclists on the road, we based our experiments on real video footage of cyclist road scenarios designed and recorded by Specialized. For this, we used grayscale to eliminate classifier color overfitting while training. We chose the KITTI raw dataset as the baseline validation dataset, which allowed us to use video of cyclists in real-life conditions and also to compare our results against top cyclist detectors without training.

The second big challenge was using a variety of accuracy indicators for different scenarios for special jersey pattern usage. In the scope of this paper we chose to limit our test results to street clothes vs. the Specialized jersey in 3 groups by level, namely Easy, Moderate and Hard. In future work, we plan to describe in more detail, the different aspects of how the custom Specialized jerseys can affect cyclist recognition accuracy.

The two datasets we utilized are described in detail below.

Specialized Cyclist Detection dataset. We introduce the new, publicly available, Specialized Cyclist Detection dataset

Fig. 6. Specialized Dataset frame.Orange jersey with black diamonds. Plain orange jersey. Street clothes

0% 25% 50%

Fig. 7. Cyclist obscurity levels

which includes video footage of several cyclists in various road scenarios in Morgan Hill, California in the southern end of the Silicon Valley.

Cyclist ground truth labeling was done by manual processing of each frame and assigning each frame to one of 2 groups by cyclist gear: Street Clothes (SC) or Jersey with Pattern (JP). See Figure 6.

- Video specs: 1920x1080, 47.95 FPS, grayscale used (original video is full color)
- Total number of frames: 62702
- Total number of frames with cyclists: 18370 (Easy 10694, Moderate 6589, Hard 1087)
- Total number of frames of cyclists with street clothes: 12575 (Easy 7738, Moderate 4581, Hard 1087)
- Total number of frames of a cyclist with a known jersey pattern: 5795 (Easy 2956, Moderate 2008, Hard 831)

KITTI tracking dataset. [1] We used parts with cyclists 0001, 0002, 0005, 0011, 0014, 0051, 0056, 0060, 0084, 0091, and 0093 of the KITTI raw dataset for training and validation. We augmented the labeling by manual processing of each frame and assigning each supplemental frame tags. Video specs:

- 1242x375, 10 FPS, grayscale used (original video is full color)
- Total number of frames: 2844
- Total number of frames with cyclists: 633 (Easy 480, Moderate 105, Hard 48)
- Total number of frames of cyclists with standard gear: 633
- Total number of frames of cyclists with known jersey patterns: 0

The classifiers we utilized for this experiment were the following:

- **YOLO.** We used the as is version of YOLO [16] to conduct tests on both datasets.
- **SqueezeDet.** For validation and tests using the new Specialized dataset, we used the current pretrained version of SqueezeDet[17]. As SqueezeDet is optimized for a resolution of 1242x375, we converted the resolution of the Specialized dataset for the SqueezeDet test.
- **CyDet.** CyDet was trained using data from both the KITTI tracking training data set and the Specialized training datasets.

Now that we have two fully labeled datasets and the selected detectors and classifiers, we split the training and test sets in the following manner.

- 50% of the total frames are used for training and validation and 20% for testing-only.
- 25% of the total cyclist frames serve as objects for training and 25% of the total cyclist frames are used for validation.
- Among false positive examples (frames without cyclists) 25% of total frames without cyclists are used for training.
- The separation is randomized.

For Recognition Quality Indicators, we define cyclist detection accuracy to be shown in Average Precision in Tables I and II, and Recall (TPR) relative to obscurity levels in Figures 8 and 9. We define our terms as the following:

- **Average Precision (AP)** = Area under Precision-Recall curve.
- **Recall = True Positive Rate(TPR)** = TP / P = Percentage of total cyclists correctly recognized by the classifier
- **Precision** = Positive Predictive Value (PPV) = TP / (TP + FP) = Ratio of correct classifications to total classifications.
- False Positive Rate (FPR) = FP / N = Ratio of incorrect classifications to total number of frames.
- P = Number of cyclists in all frames
- TP = Number of correctly recognized cyclists
- FP = Number of falsely recognized cyclists
- Please see obscurity level definition in Figure 7

Our results in Table I and Table II and graphs on Figures 8 and 9 show that the CyDet classifier utilizing the 99.9 Hybrid Architecture achieves state-of-the-art cyclist recognition accuracy. On the KITTI data set, CyDet achieved an Average Precision (AP) of 88.12% and Recall (TPR, True Positive Rate) of 82.17% with a model size of 0.0289 MB processing at 60 FPS on a 35W Intel Core i7.

Using cyclist jersey patterns custom designed by Specialized and our Winkam jersey pattern detection algorithms in CyDet, we increased Average Precision 5–15% independent of jersey color. Importantly, when also wearing the known jersey pattern that CyDet was trained to identify, Recall detection accuracy increased by 14–45% relative to wearing street clothes independent of jersey color. In future work, we plan to continue to assess the effect of other Specialized

Method	Modality	Cyclist		
		Easy	Moderate	Hard
HC-baseline [8]	LiDAR	55.35	36.07	34.15
VoxelNet [8]	LiDAR	67.17	47.65	45.11
YOLO [16]	Mono	16.71	16.03	8.33
SqueezeDet+ [17]	Mono	87.6	80.3	78.1
CyDet	Mono	**88.12**	**80.92**	**75.47**

TABLE I

PERFORMANCE COMPARISON OF CYCLIST DETECTION: AVERAGE PRECISION (IN %) ON KITTI VALIDATION SET.

Method	Modality	Cyclist		
		Easy	Moderate	Hard
YOLO (without training) [16]	Mono	47.87	38.97	32.48
SqueezeDet (without training) [17]	Mono	53.54	39.53	39.01
CyDet (street clothes)	Mono	**76.56**	**70.09**	**51.65**
CyDet (Jersey with pattern)	Mono	**81.10**	**77.85**	**77.4**

TABLE II

PERFORMANCE COMPARISON OF CYCLIST DETECTION: AVERAGE PRECISION (IN %) ON SPECIALIZED VALIDATION SET.

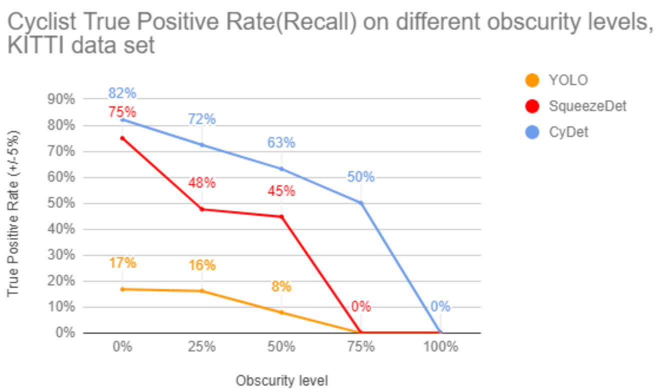

Fig. 8. KITTI. True Positive Rate(Recall) on Different obscurity levels in %

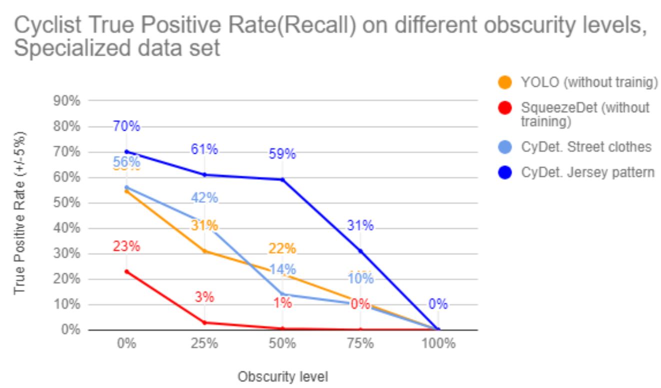

Fig. 9. Specialized data set. True Positive Rate(Recall) on Different obscurity levels in %

jersey patterns. We additionally plan to conduct rigorous tests using additional video datasets under similar conditions but different scenarios.

V. CONCLUSION

To improve cyclist detection and classification accuracy to autonomous driving perception systems, we proposed CyDet, a hybrid cyclist detection and classification method based on a combination of human knowledge of the scenario context and statistical machine learning approaches to object detection, classification, and tracking. We introduced and utilize the new Specialized Cyclist Detection Dataset that includes a known custom jersey design in order to demonstrate the improvement of accuracy in cyclist detection when the algorithm has knowledge of cyclist jersey patterns. Our results demonstrate that CyDet is small, fast, and accurate. On all of these metrics, our tests show that CyDet combined with known cycling jersey patters advance the state-of-the-art in camera-based cyclist detection and classification.

REFERENCES

[1] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? The KITTI vision benchmark suite," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.

[2] D. Woods, "Volvo car group reveals world-first cyclist detection with full auto brake in geneva," March 2013, https://www.media.volvocars.com/ca/fr-ca/media/pressreleases/48219.

[3] H. Cho, P. E. Rybski, and W. Zhang, "Vision-based bicyclist detection and tracking for intelligent vehicles," in *2010 IEEE Intelligent Vehicles Symposium*, June 2010, pp. 454–461.

[4] M. Ingram, "Tesla's autopilot tech is a danger to cyclists, robotics expert says," May 2017, http://fortune.com/2017/05/29/tesla-danger-cyclists/.

[5] P. Fairley, "The self-driving car's bicycle problem," Jan 2017, https://spectrum.ieee.org/cars-that-think/transportation/self-driving/the-selfdriving-cars-bicycle-problem.

[6] K. Murnane, "Cyclists may benefit the most and be the greatest challenge for self-driving cars," February 2017, https://www.forbes.com/sites/kevinmurnane/2017/02/13/cyclists-may-benefit-the-most-and-be-the-\greatest-challenge-for-self-driving-cars/#676b68bb6691.

[7] S. Steyer, G. Tanzmeister, and D. Wollherr, "Object tracking based on evidential dynamic occupancy grids in urban environments," in *2017 IEEE Intelligent Vehicles Symposium (IV)*, June 2017, pp. 1064–1070.

[8] Y. Zhou and O. Tuzel, "VoxelNet: End-to-End Learning for Point Cloud Based 3D Object Detection," *ArXiv e-prints*, Nov. 2017.

[9] A. Mousavian, D. Anguelov, J. Flynn, and J. Kosecka, "3D bounding box estimation using deep learning and geometry," *CoRR*, vol.

abs/1612.00496, 2016. [Online]. Available: http://arxiv.org/abs/1612.00496
[10] H. Jung, Y. Ehara, J. K. Tan, H. Kim, and S. Ishikawa, "Applying MSC-HOG feature to the detection of a human on a bicycle," in *2012 12th International Conference on Control, Automation and Systems*, Oct 2012, pp. 514–517.
[11] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, Jul 2002.
[12] X. Li, F. Flohr, Y. Yang, H. Xiong, M. Braun, S. Pan, K. Li, and D. M. Gavrila, "A new benchmark for vision-based cyclist detection," in *2016 IEEE Intelligent Vehicles Symposium (IV)*, June 2016, pp. 1028–1033.
[13] A. González, D. Vázquez, A. M. López, and J. Amores, "On-board object detection: Multicue, multimodal, and multiview random forest of local experts," *IEEE Transactions on Cybernetics*, vol. 47, no. 11, pp. 3980–3990, Nov 2017.
[14] K. B. Ge, J. Wen, and B. Fang, "Adaboost algorithm based on MB-LBP features with skin color segmentation for face detection," in *2011 International Conference on Wavelet Analysis and Pattern Recognition*, July 2011, pp. 40–43.
[15] Z. Kalal, J. Matas, and K. Mikolajczyk, "Online learning of robust object detectors during unstable tracking," in *2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops*, Sept 2009, pp. 1417–1424.
[16] J. Redmon, S. K. Divvala, R. B. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," *CoRR*, vol. abs/1506.02640, 2015. [Online]. Available: http://arxiv.org/abs/1506.02640
[17] B. Wu, F. N. Iandola, P. H. Jin, and K. Keutzer, "Squeezedet: Unified, small, low power fully convolutional neural networks for real-time object detection for autonomous driving," *CoRR*, vol. abs/1612.01051, 2016. [Online]. Available: http://arxiv.org/abs/1612.01051
[18] R. Brehar and S. Nedevschi, "Pedestrian detection in infrared images using HOG, LBP, gradient magnitude and intensity feature channels," in *17th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, Oct 2014, pp. 1669–1674.
[19] J. Xu, Q. Wu, J. Zhang, and Z. Tang, "Fast and accurate human detection using a cascade of boosted MS-LBP features," *IEEE Signal Processing Letters*, vol. 19, no. 10, pp. 676–679, Oct 2012.