Autonomous Mooring towards Autonomous Maritime Navigation and Offshore Operations

Ajit Jha*, Dipendra Subedi, Per-Ove Løvsland, Ilya Tyapin, Linga Reddy Cenkeramaddi, Baltasar Lozano, and Geir Hovland
University of Agder, Norway
MacGregor Norway AS, Norway
*ajit.jha@uia.no

***Abstract*—Bollard is a vital component of mooring system. It is the anchor point for mooring ropes to be fixed in order to secure the vessel or ship. An algorithm that translates the segmented mask of bollard output from masked R-CNN along with bounding box and associated class probability to its corresponding edge co-ordinate and finally to the single reference point for efficient detection and classification of bollard towards autonomous mooring is presented. At first stage, Mask R-CNN framework is trained with custom built bollard. The model obtained from the training is inferred with real data resulting in instance segment of bollard. The segmented mask obtained contains relatively large amount of the data points representing the whole area of bollard, which typically is not desirable. In order to precisely localize the bollard with one reference co-ordinate, the proposed algorithm is applied to segmented mask. Firstly, it translates the segmented mask to only four co-ordinate points, where each point correspond to the edge of bollard. Further, from the edges, the reference point is estimated. This causes significant reduction in point of interest (POI) and has potential to reduce the error encountered during pose estimation of the bollard in 3D thus making the autonomous mooring more precise and accurate.**

***Index Terms*—Instance segmentation, autonomous mooring, Mask R-CNN, object detection and classification**

I. Introduction

Autonomous navigation of ships or vessels and related maritime operations contribute to the aim of sustainable transportation and trade. It has potential to reduce carbon footprint, perform various operations at adverse maritime climatic and environmental conditions while reducing operational cost, increasing reliability and making activities safer thus harmonizing three core aspects - economic growt, environmental protection and social inclusion. Thus with the aim of future innovation and safety, use of driverless means of transportation and automated operations has already gained wide attention. Advanced driver assistance system (ADAS) defines the technologies and concept that guide the development of autonomous vehicle and systems. ADAS features such as obstacle avoidance, collision detection and avoidance, object detection and classification, object localization, autonomous emergency bracking (AEB) [1] help transferring most of driver's task such as perception, planning and control to vehicle itself, with minimal or exceptional intervention. Having the ADAS features implemented, there are multiple intuitions about optimal control strategies and energy management as well as elimination of human factors and driver inefficiencies which could have a positive impact on safety; the dynamics, and the fuel consumption of the vehicle. Enabling ADAS features, can help ships in the deep-sea exploration, arctic circle exploration, fishery and oceanic food cultivation, autonomous navigation and other services even in different harsh conditions. Apart from the extended use that would not have been possible, it also enjoys added benefit in terms of environment, safety and economy that can be exploited by maritime industries.

ADAS is already an established technology in automobile industries encompassing functionalities mentioned above. In [1], authors explain the different building block of ADAS for autonomous vehicle. Magali [2] developed generic and modular architecture for maritime autonomous vehicle Robot System Onboard Architecture (RSOA). However, the maritime has not seen significant developments, at least in practise, in implementing ADAS features and towards autonomous ships.

Among the other activities necessary for autonomous navigation, mooring the ship has attained special attention. Traditionally, mooring involves tying rope to the mooring stations on deck to ensure that the ship is stationary. Mooring a ship or vessel either for on/offshore operation such as ship to ship transfer of load or docking the ship in harbour is affected by many forces such as wind, water current, tides, waves, loading conditions and interaction from other ships. So it is necessary to take those factors while mooring a ship. If the mooring ropes and trajectory are not used in accordance to the above mentioned factors, it can cause serious hazards to the crew, cargo, ship and the people nearby. One one hand autonomous mooring will increase the safety of the crew and the operator while significantly reducing the operation time thus making the entire process safe, economy and sustainable. Thus it is desirable to perform the mooring process autonomously.

When it comes to autonomous mooring, among the other procedures, moor detection, classification and localization becomes the primary objective. Vision based technique combined with machine leairning (ML) algorithm has proven to be one of the most suited and elegant way of object recognition and classification. There exists different ML algorithms and framework that can efficiently detect and classify the particular object in given image frame. The modern object detection algorithms

AJ, DS, IT, GH, LR, BL is with University of Agder; PL is with MacGregor Norway AS.

are based on two stage approaches. First stage generates the object locations and second stage classifies each object with in the image with a bounding box and associated class probability. Example of such algorithms and/or frame work are Region-based CNN (R-CNN) [3], Fast R-CNN [4], Faster R-CNN [5], Feature Pyramid Network (FPN) [6]. Despite its accuracy, there are several one stage detector suc as You Only Look Once (YOLO) framework [7], [8], Single Shot MultiBox Detector (SSD) [9] that give comparative results with less computation cost [10]. Despite the fact that algorithms in both categories successfully detect and classify the object with in a bounding box and associated class probability, Mask R-CNN goes one step further in providing the segmented mask with in bounding box [11]. The implication of this segmented mask is that now each pixel in the image is associated with particular class called instance segmentation. This forms the basis of the proposed algorithm that helps us pin point the reference co-ordinate of bollard using a single co-ordinate.

This paper is organized as follows. The proposed methodology including architecture, experimental set-up and algorithm and results is described in section II and finally section III closes the paper with conclusion, discussion and future work.

II. Methodology

A. Architecture

The concept of autonomous mooring is shown in Fig. 1. It consists of a robot on-board with multiple robotic arm (manipulators) equipped with sensors such as camera, lidar, radar or combination of multiple sensors. Those sensors perceive the surrounding and deliver data in different format e.g. image, point cloud and voltages respectively. Machine learning algorithm is employed to extract the relevant information from the heterogeneous data for control and planning of the robotic arm to place the mooring rope around the bollard. Based on this, the architecture of autonomous mooring can be classified into four abstract layer [12].

- Sensor: The choice of sensors are dependent upon the operating conditions. Camera based systems give 2D view of the surroundings in good lighting conditions with relatively better performance in terms of noise, resolution, and range. They can be used in object detection and classification with relatively less computation cost. Lidar gives us the 3D profile of the surrounding independent upon the lighting conditions with reduced performance in terms of noise. They are best suited for range estimation with increased resolution, object and edge detection, however they come at the increased computation cost. Apart from these two type of sensors, Radar also finds its application for range and velocity estimation. The optimal choice of sensor and sensor fusion is made depending upon the operating conditions. At preliminary stage only camera and lidar is used for the experiment.

- Perception: The raw heterogeneous data received from sensor is processed to extract related information about the surrounding. This could be processing data e.g. 2D image, 3D point cloud or voltages coming from different individual sensors or processing the data fused with multiple sensors. As explained earlier there exists extensive list of software, algorithms and framework at this layer to process the data from sensors. However, here we use the Mask R-CNN because of the added information (segmented mask) it gives as compared to other frameworks. Further, an algorithm is proposed and presented for post processing of the segmented mask to find single point of reference for the bollard under test and is explained in proceeding sections.

- Planning: This layer has the decision making capabilities and takes the decision based on the perceived information.

- Control: This layer ensures that the planned action is executed safely.

For brevity, all these abstraction is summarized in Fig. 2.

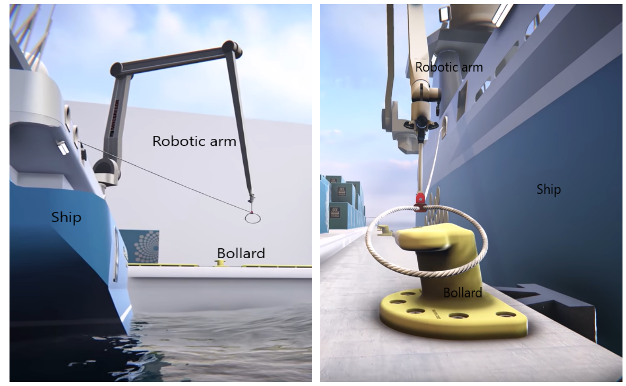

Fig. 1. Autonomous mooring. (left) conceptual presentation showing the relative position of ship, robotic arm and bollard; (right) exploded view (with permission from MacGregor Norway AS).

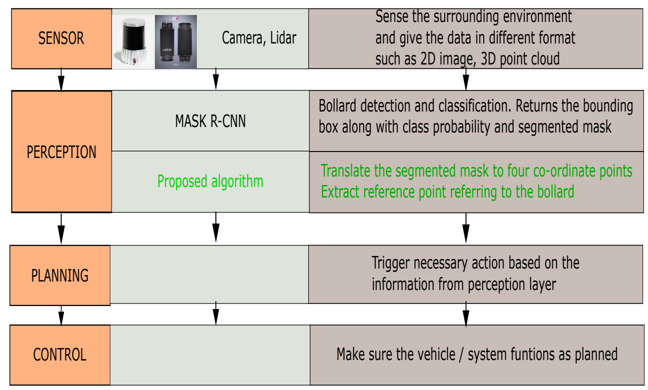

Fig. 2. Architecture of autonomous mooring and proposed algorithm at perception layer.

B. Experimental setup

Based on this description, the overall requirement translates to the detection and classification of bollard placed on the

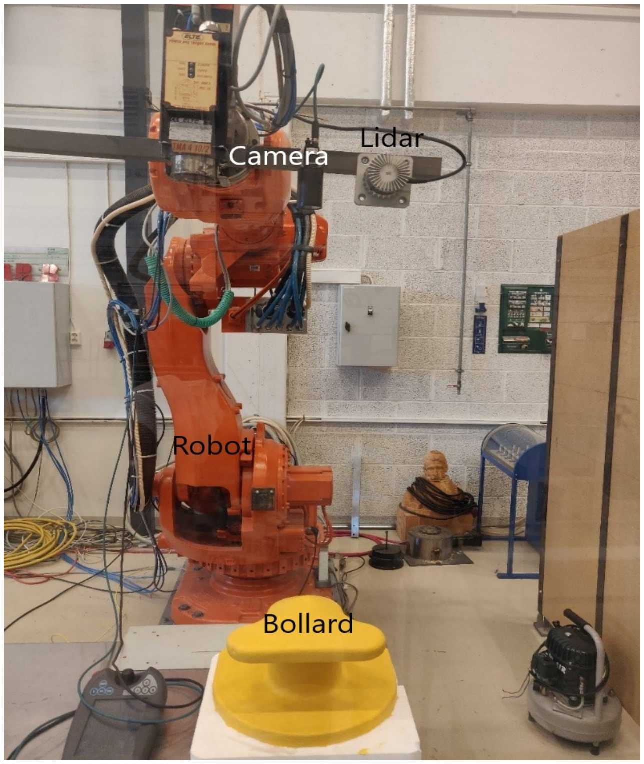

Fig. 3. Experimental setup for detection and classification of bollard for autonomous mooring.

mooring stations in real world scenario. For this purpose, experiment setup is devised from the basic constituents that would make up vessel for autonomous mooring (shown in Fig. 3). In terms of sensor to acquire raw data from the surrounding camera and lidar is used. To acquire high resolution 2D image, a camera (Lucid Triton 5.0 MP model) is used. It has Sony IMX264 CMOS image sensor with 5 MP resolution and frame rate of 24 fps [1]. The light emitting from the surrounding is focused on to the camera's sensor using Fujincon HF6X-5M. It has a focal length of 6 mm, operation of focus range being $\infty - 10$ mm, and F-number of f1.9 - f16 [2]. Next, Ouster Lidar OS1 is used to acquire 3D point cloud of the surrounding [3]. Both the sensors are mounted on ABB robot to mimic the industrial robot placed on the ship/vessel. Further, the 3D model of bollard is designed (shown in Fig. 4) and then manufactured in-house at University of Agder (shown in Fig. 3).

C. Algorithm and Results

The over-all problem formulation for autonomous mooring deduces to the detection, classification and determining a single point co-ordinate to reference the bollard. The following procedures are taken to train and then infer the model.

[1]https://thinklucid.com/product/triton-5-mp-imx264/
[2]https://thinklucid.com/product/fujinon-c-mount-2-3-6mm-f-1-9/
[3]https://ouster.com/products/os1-lidar-sensor/

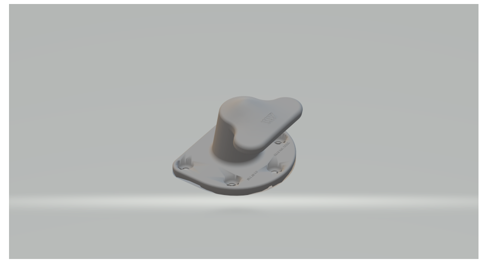

Fig. 4. 3D model of bollard used in the experiment.

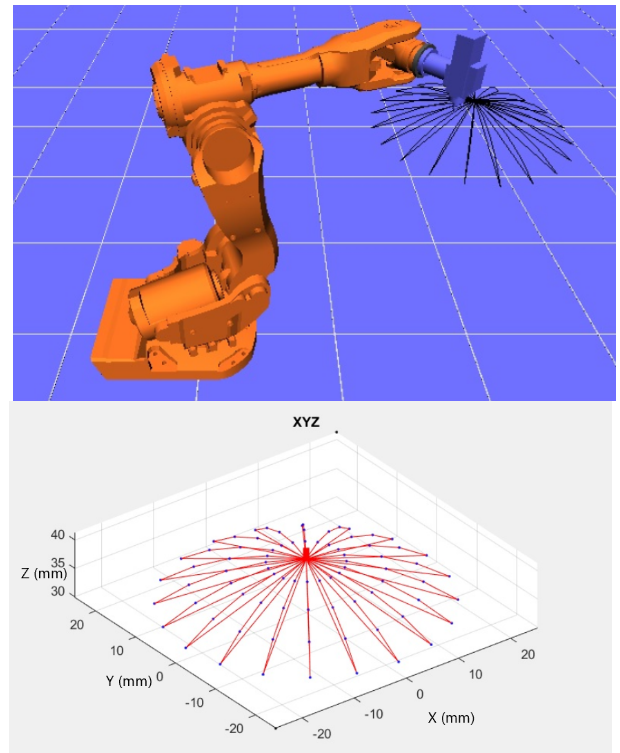

Fig. 5. Trajectory of camera placed on ABB robot to acquire images for training the ML model. (up) simulation environment (down) image capturing trajectory.

a) Sample prepration: Considering that the bollard needs to be detected at different angle and distance from the ship/vessel to be moored, the robotic arm is moved as the trajectory defined by Fig. 5. A total of 105 images are captured. A fraction is reserved for inference and rest is used to train the model.

b) Training - Mask R-CNN: Instance segmentation based framework - Mask R-CNN [11] is employed. To reduce the training time and lower the general errors, transfer learning is employed here. The model weights from pre-trained model obtained from COCO dataset [13] is chosen as the starting point and then use the learned feature to train the Mask R-CNN model. The appropriate deviations are made as compared to original model is made to meet the requirement. The number of classes is set to 2 (bollard and background), minimum confidence of detection to 0.9 (to reduce the false positive).

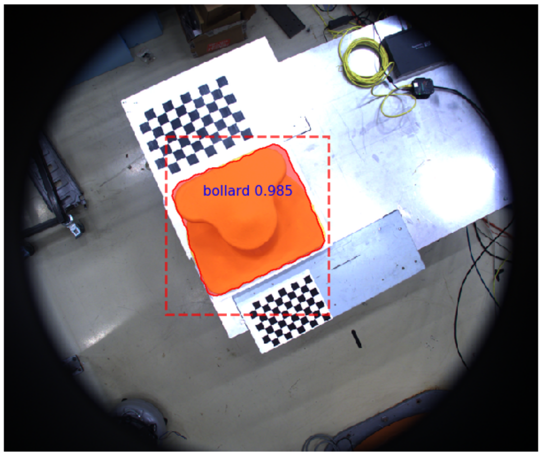

Fig. 6. Inference result. The model detects and classifies the bollard.

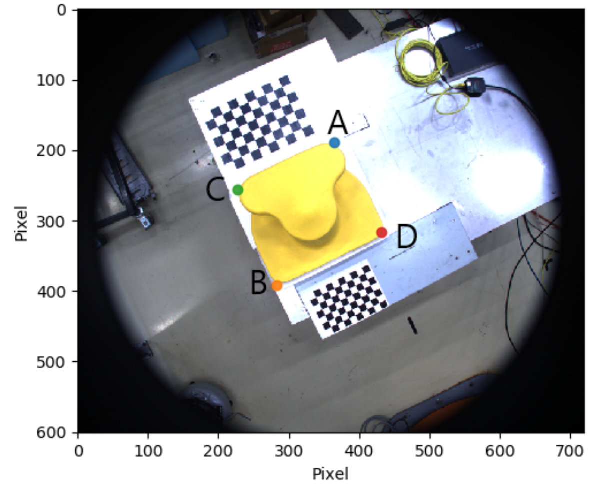

Fig. 7. Edges of bollard (A,B,C,D) detected using the proposed algorithm explained in Fig. 1.

c) Inference: The above trained model is used to infer the bollard (reserved for inference) to verify if the model is able to detect and classify the bollard based on estimation of the boundig box, class probability and the segmented mask. It was observed that the model works well and able to detect and classify the bollard with class probability of 0.99 shown in Fig. 6.

d) Post processing and proposed algorithm: From the Mask R-CNN, we have the mask of the bollard. Mask is a 2D array of elements of type boolean. It contains the co-ordinates of the pixel (x and y) corresponding to bollard as true and false otherwise. Given this information, post processing algorithm is developed and implemented to (a) precisely find the edge of bollard and (b) the single co-ordinate point that can be used to reference the bollard. The post processing algorithm to find the edges of the bollard from the segmented mask is explained in Algorithm 1 with the results in Fig. 7.

Algorithm 1 Localize edges of Bollard from Mask
Input: Mask obtained from Mask-RCNN, M
Output: Co-ordinates of edges of Bollard, E

Find first row from M where the value is true, rf
Select co-ordinates of last element in rf, A
Select co-ordinates of first element in rf, C
Find last row from M where the value is true, rl
Select co-ordinates of first element in ra, A
return P

Once the edges are determined, the point of intersection of the lines joining the opposite edges (AB and CD) gives the single reference point referring to the bollard (shown in Fig. 8). Similar procedure is applied on another random image reserved for inference. The model and the proposed algorithm both work well to determine the single reference point of the bollard (shown in Fig. 9).

In addition to the vision based approach, an algorithm to process the 3D point cloud captured from Lidar based on selecting related points based on distance, intensity, edge and

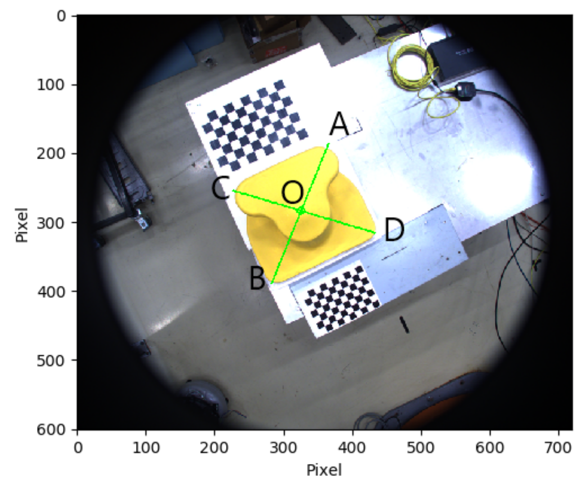

Fig. 8. The four co-ordinate points (A,B,C,D) obtained from Fig. 1 and Fig. 7 deduced to single reference point (O) referring to bollard.

field of view was also developed [14]. The results is shown in Fig. 10.

III. Conclusion and Discussion

We presented the choice of sensor and architecture to successfully detect and classify the bollard for autonomous mooring towards autonomous maritime navigation and related off-shore operations. We presented an algorithm that takes the segmented mask from Mask R-CNN as input and extracts the edge of bollard represented by four co-ordinate points. Further, from those four points, a single reference point is obtained to localize the bollard, thus translating the complexity to a single point, which in turn will help to reduce the localization error. The proposed algorithm is tested and verified by experimental data obtained in laboratory conditions. Even-though the image of bollard taken at different position and orientation of camera affected the results, however, the reference point was well

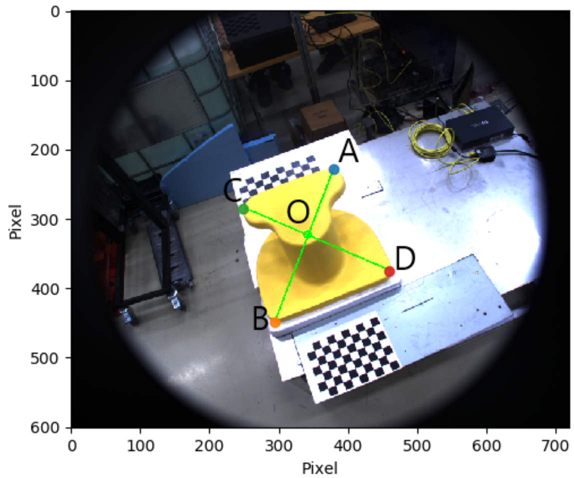

Fig. 9. Demonstrating repeatability of proposed algorithm. The proposed algorithm gives similar results on random image of bollard.

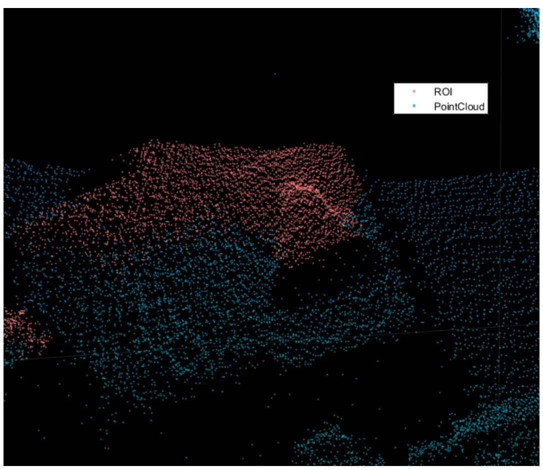

Fig. 10. Point cloud processing to select the points corresponding to bollard.

with in the accepted region (with in top surface of bollard). Future work includes testing the proposed algorithm against the bollard from real environment (one obtained from sea-shore).

While different sensors and sensor technologies are utilized for specific purposes, nevertheless, they fail individually to meet the requirements of growing decision capabilities with high accuracy. Therefore, it is important to perform an adequate sensor fusion so that the machine learning algorithms can exploit jointly the different complementary sources of data before the decision-making takes place. As a part of future work, the development of sensor fusion platform is in-progress [15]. It incudes data fusion from camera, co-ordinate of bollard obtained with the proposed algorithm with that of lidar point cloud to get a complete 3D profile of the surrounding and the location of bollard (thus optimizing the point cloud processing and getting better results as compared to Fig. 10) and implement planning and control algorithm. This in turn will help to localize the bollard with high precision and accuracy, thus moving towards our goal - by engineering solutions, the sea can be made more accessible, safe, reliable and in turn contributing to sustainable development.

Acknowledgment

The work was funded by the Norwegian Research Council, project number $261647/O20$, under the BIA Program.

References

[1] K. P. Divakarla, A. Emadi, and S. Razavi, "A cognitive advanced driver assistance systems architecture for autonomous-capable electrified vehicles," *IEEE Transactions on Transportation Electrification*, vol. 5, pp. 48–58, March 2019.

[2] M. BARBIER, E. BENSANA, and X. PUCEL, "A generic and modular architecture for maritime autonomous vehicles," in *2018 IEEE/OES Autonomous Underwater Vehicle Workshop (AUV)*, pp. 1–6, Nov 2018.

[3] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 580–587, June 2014.

[4] R. Girshick, "Fast r-cnn," in *2015 IEEE International Conference on Computer Vision (ICCV)*, pp. 1440–1448, Dec 2015.

[5] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, pp. 1137–1149, June 2017.

[6] T. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 936–944, July 2017.

[7] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779–788, June 2016.

[8] J. Redmon and A. Farhadi, "Yolo9000: Better, faster, stronger," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 6517–6525, July 2017.

[9] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "Ssd: Single shot multibox detector," in *Computer Vision – ECCV 2016* (B. Leibe, J. Matas, N. Sebe, and M. Welling, eds.), (Cham), pp. 21–37, Springer International Publishing, 2016.

[10] T. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *IEEE International Conference on Computer Vision (ICCV)*, pp. 2999–3007, Oct 2017.

[11] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in *2017 IEEE International Conference on Computer Vision (ICCV)*, pp. 2980–2988, Oct 2017.

[12] C. B. S. T. Molina, J. R. d. Almeida, L. F. Vismari, R. I. R. González, J. K. Naufal, and J. Camargo, "Assuring fully autonomous vehicles safety by design: The autonomous vehicle control (avc) module strategy," in *2017 47th Annual IEEE/IFIP International Conference on Dependable Systems and Networks Workshops (DSN-W)*, pp. 16–21, June 2017.

[13] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollar, and C. L. Zitnick, "Common objects in context," in *IEEE International Conference on Computer Vision (ICCV)*, 2014.

[14] R. Zhao, M. Pang, C. Liu, and Y. Zhang, "Robust normal estimation for 3d lidar point clouds in urban environments," *Sensors*, vol. 19, p. 1248, Mar 2019.

[15] D. Subedi, A. Jha, I. Tyapin, and G. Hovland, "Camera-lidar data fusion for autonomous mooring operation," in *2020 15th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, p. to appear, IEEE, 2020.