

Fast and Effective Identification of Window and Door Openings for UAVs' Indoor Navigation

Onder ALPARSLAN

Hezârfen Aeronautics and Space Technologies Institute
National Defense University
Istanbul, Turkey
oalparslan@hho.edu.tr

Dr.Omer CETIN

Hezârfen Aeronautics and Space Technologies Institute
National Defense University
Istanbul, Turkey
ocetin@hho.edu.tr

Abstract— As one of today's popular research field, autonomous unmanned systems are widely used in entertainment, search and rescue, health, military, agriculture and many other fields with the advantages of technological developments. Object detection is one of the methods used for autonomous vehicles to gather and report information about its environment during mission. With the ability to detect and classify objects, an unmanned vehicle can determine the type and number of objects around it and use this information in its autonomous motion and path planning or reporting the objects with the desired features. In an indoor environment, the data coming from an optic sensor provides enough capability to detect crossing points for an intelligent unmanned aerial vehicle such as doors, windows and stairs. However, sliding objects like windows and doors need to be classified as open or closed if this information will be used for path planning. Moreover, the crossing point's width needs to be calculated to ensure it is safe to go through it. So as to provide a robust navigation plan, LIDAR sensor data can be used to ensure object classification. In this study, a method to fuse LIDAR and optic sensors data has been developed and by the way a stronger object detection method is created. A crossing point is classified as an open/closed and determined whether to have enough space or not to pass through. This method was operated in a developer board mountable in a flying quadrotor which makes it possible to use it in a real-time scenario. The success of the approach has been shown with different scenarios.

Keywords— Autonomous unmanned systems, path planning, data fusion, deep learning, object classification.

I. INTRODUCTION

An autonomous unmanned air vehicle (UAV)'s travel can be provided by the help of guidance devices which allow the robots to go through pre-defined trajectory or the UAV itself has the capability to understand its environment and plan a route to the target by moving around the obstacles. This route can be updated in every step by controlling the vehicle's position and surrounding objects. Creating a robust unmanned vehicle depends on a reliable and effective path planning strategy [1]. While it is very possible to locate a mobile robot's position in outdoor with external reference systems like Global Positioning System (GPS), it is not possible in indoor environment like inside the buildings, caves, etc. An autonomous mobile vehicle in a building needs to obtain its position, map the environment, update it simultaneously to ensure loop closures and abstain from wasting time. In an unknown environment, the problem of

building up a map and localization simultaneously with the help of sensors is known as Simultaneous Localization and Mapping (SLAM) [2,3]. The problem of understanding the surrounding environment and robot's current position is figured out with SLAM which is a method that builds up a consistent map and locates the robot on the created map [2,3]. While SLAM or the other known methods are used effectively for robot navigation, these methods are not directly interested in what are the surrounding objects and obstacles. Methods like Kimera [4] can provide dynamic scene graph of the environment with metric-semantic 3D reconstruction. On the other hand, an object classified as an obstacle or a space determined as an aisle both may be used in the path planning. Moreover, if an unmanned vehicle is to use some specific crossing points or some objects as a temporary target, there is a need for another mapping algorithm including the capability of defining objects. This competence can be provided by object detection algorithms with the help of computer vision technologies.

The big computation cost and detection accuracy were two of the biggest challenges of object detection for using it real time in UAVs. By means of Convolutional Neural Networks (CNN) and Graphical Process Unit (GPU) based computing power, object detection can now be implemented to small mobile vehicles using developer boards. Convolutional neural networks produce very accurate results for detecting and classifying objects and it has become reliable. On the other side, parallel graphic processors work together to run neural network faster, so it is possible to use applications in real world even on developing boards. Besides, researchers have started to seek new mechanisms which speed up the CNN to work with less memory and fewer computational resources, such as compression and quantization of the networks [5].

An autonomous vehicle which requires to determine its route in an unknown environment by using the particular objects nearby, needs to have an object detection integrated SLAM method. This has not been used in any research according to the recent inquires. But it would be very useful for certain tasks such as creating a path using doors, windows, or stairs. However, the classification of crossing points such as doors and windows are not solely enough to plan a successful path planning. It is definitely needed to have

extra information about if it is open or closed. It can be succeeded by two different approaches. The first one is using depth images and semantic segmentation to classify the object as closed or open like in [6]. While the success of these approaches is limited, they are designed for one type of object. The second approach is using another sensor to validate object's position.

In the context of this study, to find crossing points of indoor environment for a UAV's navigation, the most known objects are classified through CNN and the pose of the object is validated by 2D LIDAR (Laser Imaging Detection and Ranging) data by which platform can decide if it needs to avoid or include it in path planning. Considering this kind of UAVs have already camera and LIDAR sensor for various purposes, a small autonomous vehicle can classify possible crossing objects as open or closed without using extra sensor and weight.

In our study, considering navigation of an UAV in a disaster indoor zone, doors, windows and stairs are taken into account as possible gateways and they all need to be classified as open or closed. To provide this validation an approach to match camera vision with another sensor data to ensure object's pose has been developed. Moreover, the width of the detected object is computed to control if it is larger than the width of the UAV. By the way, a safe crossing space can be marked as a possible gateway to another room or space and used in path planning of the UAV. In the subsequent section, current studies for classifying indoor objects, open and close classification and fusion methods are reviewed, in section three the methodology used is explained and in section four the experimental study and results are demonstrated. It is explained what has been learnt, acquired and what is to be done in the future works in the last section.

II. LITERATURE REVIEW

Many models have been proposed with the use of CNN for object recognition and classification until recently. Mask R-CNN is a convolutional network model by reinforcing R-CNN [7]. There are other studies that increase the number of frames per second (fps) that can be processed, such as the Faster R-CNN to increase the performance of R-CNN [8]. However, although these two-level models have a good detection success, the detection speed is relatively low and the memory consumption is high. This case makes two-level detection algorithms almost impossible to use in mobile robots considering the existing processor technologies especially on small platforms that can carry very limited load.

To have a lower computation cost and faster results, one level models have been proposed. Compared to the two-level models, You Only Look Once (YOLO) [9-11], Single Shot Detector (SSD) [12], Deconvolutional Single Shot Detector (DSSD) [13] and RetinaNet [14] algorithms appear to be able to detect objects much faster. This relatively high speed is coming from the simpler and shorter algorithms. YOLO

ignores the wide pipeline which optimizes individual components. It focuses on directly to being faster without giving up from detection accuracy and it succeeds to be fast by its structure [9]. The simple network provides 45 frames per seconds, while fast version can reach to more than 150 fps. YOLO takes the input image as a whole both in training and test and thereby it figures out contextual information about categories. As it is said highly generalizable [9], when applied to new domains or unexpected inputs, it is still successful.

Looking to the usage of computer vision for mobile robots' navigation, in numerous works, visionary data has been used for path planning. Moghadam et al. [15], combined the 2-dimensional laser data with stereo camera data. In this way, they succeeded to perceive 3-dimensional buildings. Sabe et al. [16], benefited from the stereo camera to avoid obstacles and move around on different surfaces. However, to be successful, it must have sufficient texture knowledge. The biggest disadvantage of using the camera as a sensor is there may be deviations in distance data on homogeneous environments such as a flat wall and in low light environments. In such cases, texture validation and surface validation techniques are applied [17]. Pomerleau handled path tracking as a classification problem and succeeded in detecting the deviation of the vehicle from the road line with the system he trained with artificial neural networks [18]. Ran et al. trained the images using CNN which are gathered with a spherical camera and they ensured that the robot turned to the correct route by determining how many degrees it was traveled from the desired direction [19]. Hadsell and his friends labeled the environment by dividing it into 5 classes in order to provide a pathway on off roads [20]. Surfaces such as trees and buildings are classified as obstacle or super obstacle, surfaces such as soil and asphalt as ground or super-ground, and navigation is planned on the land where the robot can go between obstacles. Many other methods also successfully classified the environment as obstacles, targets or corridors, yet they didn't look for what exactly is the surrounding objects. Understanding what the crossing point is and moving to that point in accordance with this information may be helpful for the robot's navigation.

In an indoor space the most probable gateways to another area are windows, doors and stairs. In our sense, so as to route the UAV to explore the area, these objects need to be classified. Many researchers have preferred CNN [21] and 3D information [22,23] to detect doors. Kwak et al. presented a method to detect the door and the handle in an office to open and go into the room [24]. For the detection of door, they used STAR Detector which is based on matching the features of image with a reference one.

In [21], a method is introduced to detect door and cabinets and door handles with 3D Kinect camera. The system was trained with 510 images of doors and 420 images of cabinets using YOLO. They also used the depth data of 3D sensor to create point cloud. In [25], they obtained the open doors by using point clouds and detecting gaps in the wall. An improvement of this study has been introduced for detecting

semi open doors in [26]. It is implied in the study that bigger dataset provides better results. The study [6] was carried out to perform door detection in low powered computers. They used RGB-D sensors and a good classification accuracy. However, the speed of the method (1 FPS) can be a drawback for the algorithm.

Other than depth and RGB sensors, using LIDAR can be beneficial for route planning. In [27], they benefited from 360 degrees view of the LIDAR sensor for obstacle avoidance and navigating safely. In [28], authors presented a semantic situation awareness system to obtain robot's position by the help of LIDAR, CNN and indirect-EKF. Many other researchers also used LIDAR sensor to obtain unmanned vehicle's position or path planning.

As seen in previous studies, detection of crossing waypoints indoor and classification as open or closed is possible and LIDAR information can be used for obstacle detection and navigation. Inspired by these researches, it is quite possible to compute the distance and width of detected objects and use this information for autonomous unmanned aerial vehicle's indoor navigation. It provides precise guidance to cross openings and safely navigate by means of reconsolidation of two different sensor data.

III. METHODOLOGY

In this study, a method has been developed to conceive if an object is open or closed by fusing the LIDAR data and classified object data provided by optical sensor. For this purpose, one LIDAR and one optical camera used as sensors. A CNN algorithm to detect indoor objects used and a fusion method developed within the study. The general flowchart of methodology can be seen in Fig. 1.

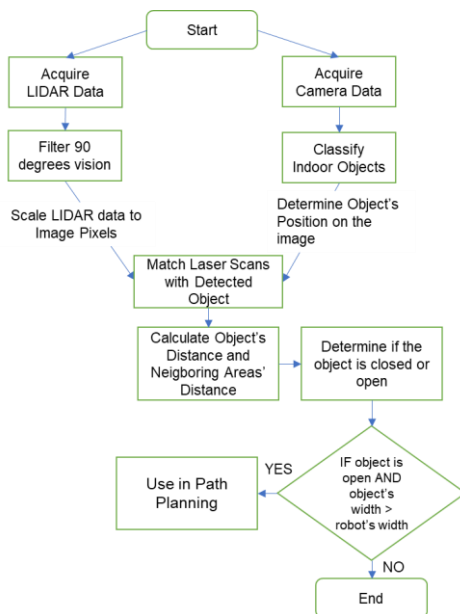


Figure 1. Flowchart of the fusion of optic and LIDAR sensors data to ensure crossing points for indoor navigation of autonomous UAVs

A. Sensor Features

The main purpose of the study is to match the different types of data from different sensors without delay that can

perceive different characteristics of the objects and environment in the field of view. To provide this, the fusion method should be applicable without knowing the physical properties of the sensors and should be independent of sensor architectures, manufacturers and models. Considering one of the biggest challenges of small platforms, which is payload, lighter equipment required to preferred as much as possible. An indoor flying unmanned vehicle also should be smaller to pass easily through limited space. Focusing on this point, an ordinary optical camera and 2D LIDAR device have been chosen as sensors. The camera, Logitech Brio [29] provides 90 degrees extended view in 4K. It produces at least 30 fps and it is only 63 grams. RP Lidar-A2 has been used as 2D LIDAR sensor. It works from 0.15 m to 12 m distance and scans the area 360 degree. One of the most important features sample duration is 0.25 millisecond which is compatible with camera producing 30 images in a second. Since it is lightweight and easy to implement, it is aimed to see the efficiency of 2D LIDAR on a UAV which is able to move vertically in 3D space.

The study is based on a UAV task in a disaster zone without any environmental help. When choosing sensors, better or less qualified sensors could be preferred. However, considering the UAV already has these sensors, the same sensors can also be used for other purposes like performing SLAM method by benefiting from LIDAR or looking for an object with the help of the camera.

B. Receiving Optic Data and Detecting Objects

An unmanned aerial vehicle equipped with camera sensor displays the environment lively. When the vehicle is navigating, it approaches to the objects from different angles and sometimes under various lightings. For an autonomous robot, classification algorithm needs to work in real-time with accurate results. This is helpful for object detection if the method misses or falsely detected the object in the first frame. In recent years it seems one stage models like YOLO, SSD and RetinaNet have made the detection and classification possible in unmanned vehicles.

One of the disadvantages of recent fast algorithms, they have difficulties in detecting small-sized objects. Whereas it is quite advantageous making the neural network simpler and smaller, it is getting difficulty to cope with small sized objects in the images. It is recommended to use a slower method or increase the input image size manually if sacrificing from detection speed is pleasant. Having a good detection rate using only one convolutional network, YOLO appears to be the fastest algorithm among the current detection algorithms. As stated by Bochkovski et al. [30], when YOLOv4 is compared to its peers, it does not fall short in detection success, but shows a clear better performance in object detection speed.



Figure 2. Detection results

There are numerous datasets to classify doors, windows and stairs. In this study, an open source dataset [31] has been combined with images taken from internet and trained with YOLOv4. By using 4000 images with 15000 iterations, the accuracy is better than using directly OpenImages [32] dataset. As seen in Fig. 2, doors can be detected with this trained data successfully. The detected object's pixel position can be saved as seen in (1) where x is the starting pixel value of the object and w is the width and h is the height of an object.

$$O_i = \{x_i, w_i, y_i, h_i\} \quad (1)$$

C. Finding Optical Image in 2D LIDAR vision

One of the most important points of fusion is locating the sensors. For sensor alignment, LIDAR and camera sensors are positioned in the same horizontal direction and adhere to each other so as to scan the same region. The camera, LIDAR and computer located one under the another as seen in Fig. 3. In the horizontal plane, sensors' location is the same. The sensors are at different heights, however, it doesn't affect the fusion algorithm due to the two-dimensional nature of LIDAR.



Figure 3. Placing of sensors

The camera has 90 degrees view. Thus, the first thing to do is limiting the LIDAR data with 90 degrees where is the same area in camera's scope. LIDAR's 360 degrees vision produce data for every scan in 0.25 millisecond. Although it is not possible to hinder the LIDAR to turn around itself and scan 360 degrees, unneeded scan data can be eliminated. The desired scans in the area where the camera is able to see are subtracted from LIDAR's whole scan and only the front-view data is taken into the method such as in (2) where r_α is the

distance of the sensor on α position. In our alignment the scans between the 45 and 135 degrees meet with camera range which is why limited the desired scans are limited in (2). The list consists N number of distance data coming from LIDAR for the frontside of the UAV. It could be more than 90 values as the values are taken fractional.

$$R = \{r_\alpha : 45 < \alpha < 135\} \quad (2)$$

Object's bounding box data saved in the list β_i where x is the pixel value of the bounding box's left border and w is width of the bounding box. Camera's object related pixel values are divided into 1920/90 to scale to match with LIDAR sensor's 90 degrees vision in that position as in (3). Thus, any area in the LIDAR view can be matched horizontally with the camera view and the detected object area is marked on the laser scans.

$$\beta_i = \frac{x_i}{1920/90}, \frac{x_i+1}{1920/90} \dots \frac{x_i+w_i}{1920/90} \quad (3)$$

To find the distance values of related pixels, β_i 's distance values are taken into list R_i as in (4).

$$R_i = [r_{\beta_1}, r_{\beta_2}, \dots, r_{\beta_N}] \quad (4)$$

By using LIDAR measures, any area's distance to sensor can be taken from this list. This provides to obtain any detected object's distance. Like seen in Fig. 4, a chair's distance to sensor is measured accurately. Moreover, LIDAR data is illustrated with grayscale values as seen in Fig. 4 to present a better understanding of the place.

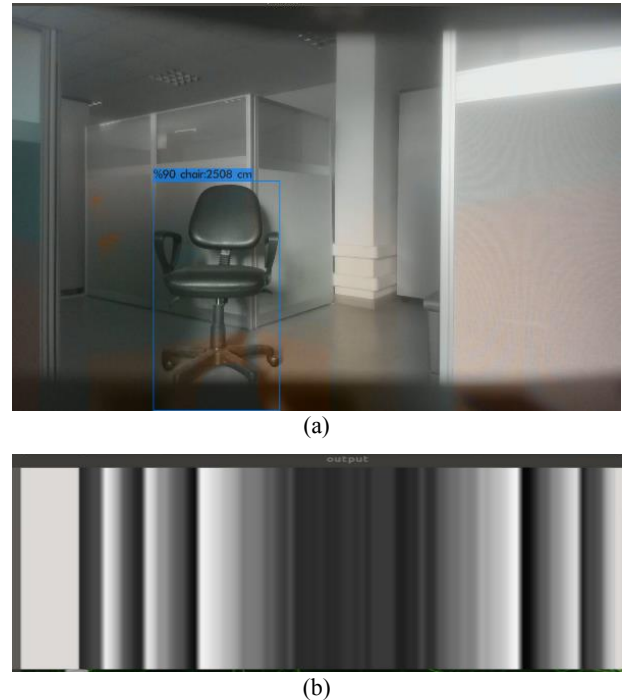


Figure 4 (a). Computing object's distance with LIDAR (b) Normalized grayscale demonstration of 2D LIDAR data

D. Comparison of Distance Information with Detected Objects

Matching the LIDAR points with image provides both calculating the distance with the classified object and understanding if the object is open or close. To succeed this, the center and width of the detected object is taken from the detection algorithm as pixel values and converted the viable values for LIDAR. LIDAR seeks the desired objects from 0 to 90 degrees. It is known that an open door differentiates from neighbor elements by distance. Using this difference helps to classify detected object as open or closed. To achieve this distinction, a method to compare distance values of object area and neighboring area has been implemented.

The left and right neighboring areas are found by going to left and right sides of the objects as much as half of the object's width as seen in (5) and (6).

$$\beta_{il} = \frac{x_i - w_i/2}{1920/90}, \frac{x_i - w_i/2 + 1}{1920/90}, \dots, \frac{x_i}{1920/90} \quad (5)$$

$$\beta_{ir} = \frac{x_i + w_i}{1920/90}, \frac{x_i + w_i + 1}{1920/90}, \dots, \frac{x_i + 3/2(w_i)}{1920/90} \quad (6)$$

The average distance of object area, left neighboring area and right neighboring area are compared. Left and right neighboring areas' mean average is computed like in (7) and compared as in (8).

$$D_i = AVG(R_i), D_{il} = AVG(R_{il}), D_{ir} = AVG(R_{ir}) \quad (7)$$

$$STATUS = \begin{cases} OPEN, & |D_i - D_{il}| > 0.4 \\ OPEN, & |D_i - D_{ir}| > 0.4 \\ CLOSED, & OTHERWISE \end{cases} \quad (8)$$

As seen on Fig. 2, the detected objects neighbors' average needs to be lower distance values if it is open. Yet, the difference needs to be higher than a threshold. Otherwise, a door sill, for example, can mislead the algorithm.

IV. EXPERIMENTAL STUDY

To illustrate the method, a mechanism to travel indoor and find the crossing points has been emulated. The mechanism consists of a computer, a camera and a LIDAR sensor which is enough to test the proposed fusion method. The mechanism is not implemented on a UAV, while it is considered to mount easily. Even if a flying robot navigates with angular movements in 3D space, detection of these movements and modifying the effects on sensors have already figured out in the literature with numerous studies. Thereby, although the effects of angular movements on sensors are ignored in this experimental working environment, in practice, the methods necessary to correct these effects will have to be applied to the sensor data. To provide both computing and lightweight requirements, Jetson AGX Xavier developer board is used to benefit from general-purpose computing on graphics processing units. This board and the sensors can easily be mounted on a small unmanned aerial vehicle.

A. Sensors Data Fusion Application

The test software is developed on developer board with C++ and Python programming languages. The image from the camera has 1920x1080 pixels. While the width of image is 1920, the data from LIDAR sensor is 90 degrees width

which means a normalization is needed to match two visions. To overlap sensors' data, LIDAR data is magnified to 1920 pixels scale. Thus, the detected object area is marked on the laser scans. As the camera and the LIDAR sensors are in the same position in horizontal axis, the normalization provides to take measures from the same places in the environment.

B. Application of Method

The distance of the object to the vehicle is determined by averaging the laser scans' measurement for the desired object. Moreover, the distance of neighboring sides of object also measured with the same method. A closed door or windows do not show quite a difference from the neighboring walls. The door or window sills (difference from the wall level) can create some variation, however. To decide the exact state, a threshold value is applied which is determined experimentally. If the difference value between the neighboring walls and detected object is higher than 40 cm, it means it is an open crossing point. Otherwise, it is a closed object. This threshold value is determined by considering the maximum wall diameter in present-day buildings.

While flying platform is wandering around, it detects crossing points in real-time and asks to LIDAR if it is open or closed. The returning data is taken into detection algorithm and label of the object is detailed as open or closed as shown in Fig. 5. Although it was aimed to determine the doors and windows and the transition areas inside the building within the scope of this study, the main difference between the door and window features was their height from the ground. The proposed method was tested with doors and windows where it is observed that glass surfaces do not hinder the system working and can be easily adapted for the detection of other transitions.

The crossing point is classified successfully in both examples. Also, the distance and width of the object is taken. This provides good consequences to use the information for vehicle's navigation.

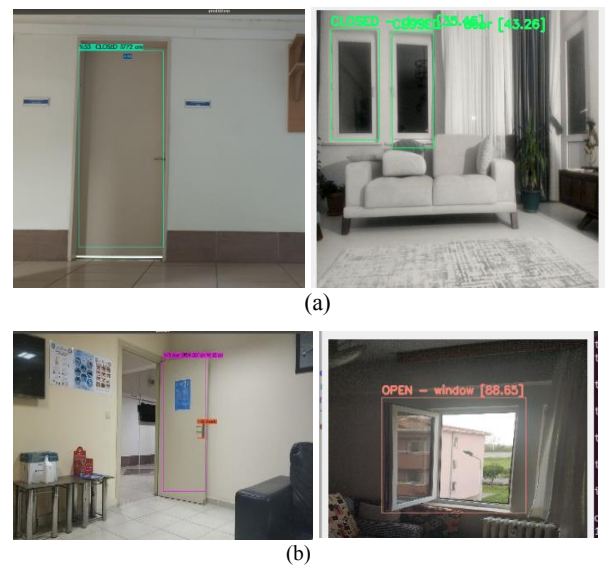


Figure 5 (a). Classifying of closed objects (b) Classifying of open objects

C. Discussion of Application Results

In the application, crossing waypoints in an indoor environment has been detected by the help of YOLOv4 algorithm on live video. To comprehend object's position as closed or open, LIDAR data has been matched with camera image and distance of the object and neighboring areas measured. To be able to have a clearer vision of the environment, illustration of distance data has been implemented by visualizing the LIDAR data on an image window as such in Fig. 6.

Looking to results, using the distance data, object's position is decided successfully in the test. The threshold value between the detected object and neighboring walls provided to classify object accurately even if the door's position on the wall like in Fig. 5 (b). In this case, the object's position can not be used for path planning. Yet, it is still true that there is a door or window and an opening which can guide the platform for navigation. Beyond this point, LIDAR data can be used to plan UAV's trajectory as it is known that there is an opening in the neighboring area of the detected object.

The test was run on Jetson AGX Xavier which is a small developer kit, to rehearse a real life problem. The developer kit performs the classification algorithm in 17 FPS average which is satisfactory for a mobile platform. The fusion method did not slow down this speed and proved its success for real time scenarios.

The width of the object is calculated to understand if the crossing way is larger than UAV's width. One drawback of the study is measuring the width of the detected object. Even though the calculation method works well, the bounding box data from detection algorithm is not completely correct. Because of this, the method computes the width of bounding box inaccurately. In Fig. 5, incorrect bounding box results in finding the door's width smaller than the real value.

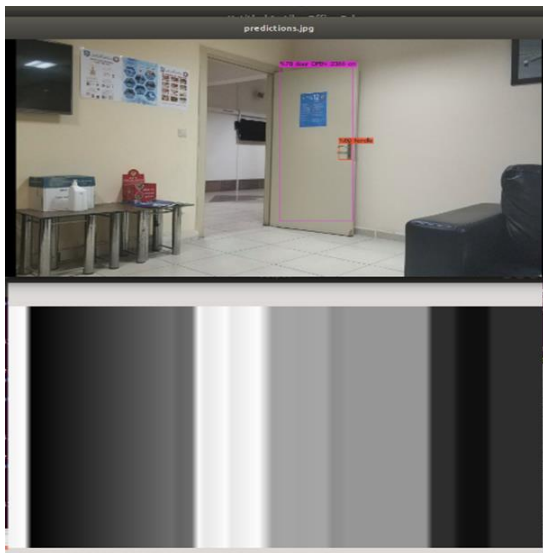


Figure 6. Illustration of distance data

V. CONCLUSIONS

In this study, a sensor fusion approach was introduced between LIDAR and camera sensor data to be able to use this

information in navigation of UAVs. Using an object classification algorithm, indoor environment objects such as doors, stairs and windows were detected. To be able to find these objects' distance to the sensors, LIDAR data was matched with image pixels and distance matrix of the related pixels was used to find object's distance. Putting in the same horizontal axis, LIDAR and camera sensor were performed together and distance of any classified object was calculated. Moreover, the position of the object as closed or open was obtained by comparing the distance values of object with object's neighboring area. Also, the width of the object was found by using camera's distance data to the object, camera's view angle and object's occupation in pixels. With this information, a UAV in an indoor environment like a disaster zone can find the crossing points to other rooms, decide if it is an opening and whether is usable for UAV's itself. The classification information could be helpful for finding a desired place in the zone or evacuating the area promptly, so it would be valuable to use it in the navigation plan. The method was tested on a platform in real time. Considering real time scenarios, a developer board with a small weight and limited computing power was used. The fusion method did not slow down the classification algorithm and it is fast enough to use in indoor UAVs.

The proposed methods can be used as an input for a UAV's path planning algorithm. Taking into account an open door or window's position is to the side of the closed position and not a crossing point, the detected object's position cannot be used in navigation. However, a method taking the classification information and using the neighboring area as a possible crossing point in the navigation algorithm can be developed. Another drawback is the accuracy of computation of the object's width. Many factors including detection algorithms, sensor quality and computers can affect this data. However, the size, weight and limited payload of UAV can constrain the use of better sensors and computers. To overcome this problem, an advanced algorithm or dataset to draw bounding boxes more precisely is needed, so that UAV can determine whether an opening has enough space for the UAV to pass through. A study is planned to test these points and usage of 3D LIDAR sensor. Thereby, the success of 3D and 2D LIDAR sensors can be compared to each other.

REFERENCES

- [1] Hassani, Imen, Imen Maalej, and Chokri Rekik. "Robot path planning with avoiding obstacles in known environment using free segments and turning points algorithm." *Mathematical Problems in Engineering* 2018 (2018).
- [2] J. Leonard and H. Durrant-Whyte, "Simultaneous map building and localization for an autonomous mobile robot," *Proceedings IROS '91:IEEE/RSJ International Workshop on Intelligent Robots and Systems '91*, no. 91, pp. 1442–1447, 1991.
- [3] C. Cadena, L. Carlone, H. Carrillo, Y. Latif, D. Scaramuzza, J. Neira, I. Reid, and J. Leonard, "Past, Present, and Future of Simultaneous Localization and Mapping: Toward the Robust-Perception Age," *IEEE Transactions on Robotics*, vol. 32, no. 6, pp. 1309–1332, 2016.
- [4] Cruz, Nicolás, Kenzo Lobos-Tsunekawa, and Javier Ruiz-del-Solar. "Using convolutional neural networks in robots with limited computational resources: detecting NAO robots while playing soccer." *Robot World Cup*. Springer, Cham, 2017.
- [5] Rosinol, A., Violette, A., Abate, M., Hughes, N., Chang, Y., Shi, J., ... & Carlone, L. (2021). Kimera: from slam to spatial perception with 3d dynamic scene graphs. *arXiv:2101.06894*.

- [6] J. G. Ramôa, L. A. Alexandre and S. Mogo, "Real-Time 3D Door Detection and Classification on a Low-Power Device," 2020 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC), Ponta Delgada, Portugal, 2020, pp. 96-101.
- [7] K. He, G. Gkioxari, P. Dollár, R. Girshick, "Mask R-CNN", 2017 IEEE International Conference on Computer Vision (ICCV), pp. 2980-2988, 2017.
- [8] S. Ren, K. He, R. Girshick, J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137-1149, June 2017.
- [9] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, "You Only Look Once: Unified Real-Time Object Detection", 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779-788, 2016.
- [10] J. Redmon, A. Farhadi, "YOLO9000: Better Faster Stronger", 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 6517-6525, 2017.
- [11] J. Redmon, A. Farhadi, YOLOv3: An Incremental Improvement, 2018.
- [12] W. Liu, D. Anguelov, D. Erhan et al., SSD: Single Shot MultiBox Detector, pp. 21-37, 2015.
- [13] C. Y. Fu, W. Liu, A. Ranga et al., DSSD: Deconvolutional Single Shot Detector, 2017.
- [14] T. Y. Lin, P. Goyal, R. Girshick et al., "Focal Loss for Dense Object Detection", IEEE Transactions on Pattern Analysis & Machine Intelligence, no. 99, pp. 2999-3007, 2017.
- [15] Moghadam, Peyman, Wijerupage Sardha Wijesoma, and Dong Jun Feng. "Improving path planning and mapping based on stereo vision and lidar." 2008 10th International Conference on Control, Automation, Robotics and Vision. IEEE, 2008.
- [16] Sabe, Kohtaro, et al. "Obstacle avoidance and path planning for humanoid robots using stereo vision." IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA'04. 2004. Vol. 1. IEEE, 2004.
- [17] PTGrey, 2004. Point grey research inc. <http://www.ptgrey.com>
- [18] Pomerleau, Dean A. "Efficient training of artificial neural networks for autonomous navigation." Neural computation 3.1 (1991): 88-97.
- [19] Ran, Lingyan, et al. "Convolutional neural network-based robot navigation using uncalibrated spherical images." Sensors 17.6 (2017): 1341.
- [20] Hadsell, R.; Sermanet, P.; Ben, J.; Erkan, A.; Scoffier, M.; Kavukcuoglu, K.; Muller, U.; LeCun, Y. Learning long-range vision for autonomous off-road driving. J. Field Robot. 2009, 26, 120–144.
- [21] A. Llopart, O. Ravn and N. A. Andersen, "Door and cabinet recognition using convolutional neural nets and real-time method fusion for handle detection and grasping", 2017 3rd International Conference on Control Automation and Robotics (ICCAR), pp. 144-149, April 2017.
- [22] T. H. Yuan, F. H. Hashim, W. M. D. W. Zaki and A. B. Huddin, "An automated 3d scanning algorithm using depth cameras for door detection", 2015 International Electronics Symposium (IES), pp. 58-61, Sep. 2015.
- [23] S. Meyer Zu Borgsen, M. Schöpfer, L. Ziegler and S. Wachsmuth, "Automated door detection with a 3d-sensor", 2014 Canadian Conference on Computer and Robot Vision, pp. 276-282, May 2014.
- [24] N. Kwak, H. Arisumi and K. Yokoi, "Visual recognition of a door and its knob for a humanoid robot", 2011 IEEE International Conference on Robotics and Automation, pp. 2079-2084, May 2011.
- [25] B. Quintana, S. A. Prieto, A. Adán and F. Bosché, "Door detection in 3d colored laser scans for autonomous indoor navigation", 2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN), pp. 1-8, Oct 2016.
- [26] B. Quintana Galera, S. Prieto, A. Adan and F. Bosché, "Door detection in 3d coloured point clouds of indoor environments", Automation in Construction, vol. 85, pp. 146-166, 2018
- [27] A. Moffatt, E. Platt, B. Mondragon, A. Kwok, D. Uryeu and S. Bhandari, "Obstacle Detection and Avoidance System for Small UAVs using a LiDAR," 2020 International Conference on Unmanned Aircraft Systems (ICUAS), Athens, Greece, 2020, pp. 633-640,
- [28] J. L. Sanchez-Lopez, C. Sampedro, D. Cazzato and H. Voos, "Deep learning based semantic situation awareness system for multirotor aerial robots using LIDAR," 2019 International Conference on Unmanned Aircraft Systems (ICUAS), Atlanta, GA, USA, 2019, pp. 899-908.
- [29] Logitech Brio Specs, <https://www.logitech.com/tr-tr/products/webcams/brio-4k-hdr-webcam.960-001106.html>.
- [30] Bochkovskiy, Alexey, Chien-Yao Wang, and Hong-Yuan Mark Liao. "YOLOv4: Optimal Speed and Accuracy of Object Detection." arXiv preprint arXiv:2004.10934 (2020).
- [31] Arduengo, Miguel, Carme Torras, and Luis Sentis. "Robust and adaptive door operation with a mobile manipulator robot." arXiv preprint arXiv:1902.09051 (2019).
- [32] Kuznetsova, A., Rom, H., Alldrin, N., Uijlings, J., Krasin, I., PontTuset, J., et al. (2018). The open images dataset v4: Unified image classification, object detection, and visual relationship detection at scale.