Object Detection for Autonomous Lawn Mower based on YOLO V3

Gabrielle Simms
Computer Engineering Dept.
Florida Polytechnic University
Lakeland, Florida, U.S.
gsimms@floridapoly.edu

Muhammad Abid
Computer Science Dept.
Florida Polytechnic University
Lakeland, Florida, U.S.
mabid@floridapoly.edu

***Abstract*— This paper presents an obstacle and object detection algorithm based on the YOLOv3 model. These programs were designed to provide computer vision capabilities to an autonomous lawn mower. The purpose of extending the obstacle detection program is to recognize objects typically found in an airport (planes, luggage, airport signers, etc.). Finally, a distance estimation technique is implemented to notify the system when the object is within a certain vicinity of the unit.**

***Index Terms*—Object Detection, convolutional neural networks, YOLO, distance estimation**

I. Introduction

Object recognition is a term to describe a collection of related computer vision tasks that involve identifying objects in a digital space. Traditionally, object detection is implemented using machine learning where computer vision techniques are used to look at various features of an image (color histogram, edges, etc.) to identify groups of pixels that may belong to an object [9]. These features are then fed into a regression model that predicts the location of the object. However, in the last few years, the rapid advances of deep learning have greatly accelerated the momentum of object detection. Deep learning is an evolution of machine learning where a program can determine if its prediction is accurate without human intervention also known as unsupervised learning. With the advancement of the computing power of graphics processing units combined with deep learning techniques, the performance of object detectors and trackers has greatly improved, achieving significant breakthroughs in object detection.

The tasks associated with object detection include image classification and object localization. Image classification is the process of categorizing and labeling groups of pixels or vectors within an image based on predefined characteristics. The most popular architecture used in image classification based on deep-learning models is convolutional neural networks. Convolutional neural networks utilize a multilayer approach to create a hierarchical feature decomposition. For example, the first layer of a CNN may extract low-level features such as lines. The layers that evaluate this output may extract features that are combinations of low-level features such as shapes. This process continues through multiple (deep) layers until the features that are being extracted include faces, animals, and other objects [10]. Object localization refers to identifying the location of one or more objects in an image and drawing a bounding box around their extent [11]. A successful object detection algorithm executes both tasks with accuracy and speed.

One of the most popular applications of object detection is autonomous driving. Some applications may simply opt for obstacle detection which is the identification of an object without classification. This notifies the system of an obstacle such that the system can take precautionary measures to avoid the collision. More advanced applications use object classification to provide more information such as sign recognition, the object's trajectory, etc. This paper presents an implementation of an object detection algorithm based on the YOLOv3 model. These programs were designed to provide computer vision capabilities to an autonomous lawn mower. The purpose of extending the obstacle detection program is to recognize objects typically found in an airport (planes, luggage, airport signers, etc.). Finally, a distance estimation technique is implemented to notify the system when the object is within a certain vicinity of the unit follow.

II. Literature Review

There is extensive research in the field of object detection for autonomous systems. The work presented in [3] aimed to build weather-resistant autonomous lawn mowers which can operate in diverse environments and light conditions. Thus, this work evaluated the use of a low-cost camera for obstacle detection. The implementation uses the mean shift algorithm to aid in image segmentation by tracking the mean hue, and mean saturation. Using these values, the segmentation parameters are continuously updated. Similarly, the work in [5], uses image segmentation however, the work uses several preprocessing methods for object detection. However, these methods depend heavily on the color of the objects as well as the grass, in that the object must not be a similar color to that of the grass.

Other obstacle detection implementations utilize stereoscopic vision, the simulation of human binocular vision, therefore extending a camera's ability to perceive depth. For example, the work in [2] presents a two-part algorithm. The first part uses the disparity map from the stereo camera to detect salient obstacles, objects that stand out from the background. Using the disparity map, the algorithm considers the disparity values where values that gradually decrease from nearer points to farther points are ground points however, salient obstacles regions usually have the same or approximate disparity value. The second part of the algorithm uses an improved SVR (support vector regression) method to detect small obstacles. Stereoscopic vision can also be used to detect the distance to the object. However, this implementation depends on the use of a stereo camera which is not low-cost.

In contrast to these implementations, this paper presents a low-cost robust solution. Using artificial intelligence, the system acquires more information about the object which can aid in the development of additional avoidance maneuvers. The program is also equipped with a distance estimation technique such that the unit will stop once the object is within a certain threshold. While using machine learning can require higher processing power, this implementation was tested on the Jetson Nano 3GB development kit which is approximately $59.

III. Object Detection

The YOLO (You Only Look Once) algorithm is an object recognition algorithm that takes a regression approach by using CNN to predict class probabilities and draw bounding boxes simultaneously. YOLO is known for its speed, high accuracy, and learning capabilities. The YOLO algorithm begins by dividing the image into grids, where each grid is responsible for the detection and localization of the object it contains via CNN. YOLO uses a bounding box regression to predict the height, width, center, and class of objects.

To determine which bounding box accurately locates the object, intersection over union (IoU) is used in conjunction with non-maximum suppression. IoU provides the ratio of the overlap to the total area of the two boxes. This provides a sufficient approximation of how close the bounding box is to the original prediction. Non-maximal suppression suppresses all bounding boxes that have lower probability values. Using the two methods combined, YOLO takes the largest probability scores associated with each decision and suppresses the bounding boxes that have the largest IoU value with it. This step is repeated until each object is localized.

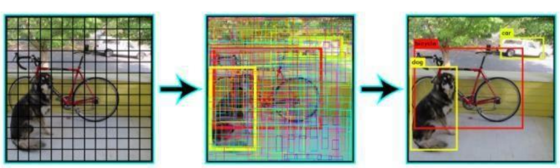

Figure 1: YOLO progressive output [11]

IV. Distance Estimation

To estimate the distance from the object, a virtual mesh was determined through experimental measurements. To facilitate a region of interest of 5ft, the camera was tilted such that the top of the image was the 5ft marker. Using a measuring tape, the Y-value that corresponds to each foot was recorded, which is displayed in table 1. The table can be interpreted such that objects with a y-coordinate of more than 0 are within five feet from the unit. Likewise, objects with a y-coordinate above 260 are less than one foot away from the object. The color of the pixels within these rows was changed for user reference.

Y-value	Distance (feet)
0	5
20	4
52	3
115	2
260	1

Table 1

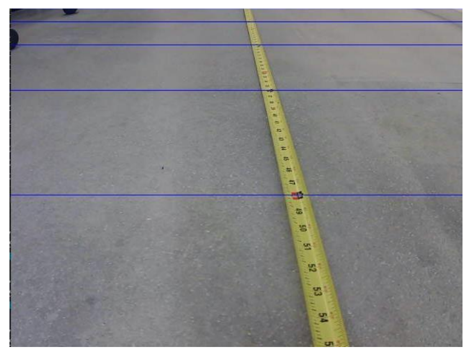

Figure 3: Virtual Mesh Output

Notice that the number of pixels between each foot gradually gets smaller thus, another formula is needed to closer approximate the object's distance. If an object is between two values (i.e. 1ft and 2ft), a closer approximation to the distance can be made using the formula.

$$D = F1(ft) - \frac{F1 - y}{F1 - F2}$$

Where F1 is the lower bound, F2 is the upper bound and y is the pixel location in the y-direction. For the best results, the Y-value should be the lowest point of the object. For example, if the lowest point of the detected object is at (150, 70). Using table 1, it can be determined that the object's distance is between 2ft and 3ft. The equations become

$$D = 2 + \frac{115 - 70}{115 - 52} = 2.71\,ft$$

V. RESULTS

A successful program will recognize the object as well as return its distance. Figure 2 displays the output of the YOLOv3 program. The program returns the name of the object as well as the confidence interval. Using this information, a control algorithm was designed and tested.

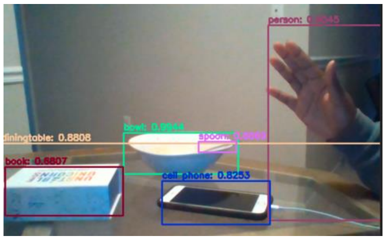

Figure 2: YOLOv3 output

The unit was tested on the Florida Polytechnic University campus. The following image displays a student walking in the path of the lawn mower. Using this information, the unit stops as well as turns off its blade for safety purposes. If the obstacle has not moved within 15 seconds, the unit will attempt to find an alternate route.

Figure 3: Lawn Mower obstacle detection

The program can be said to have the following success rates with obstacles typically found on the campus.

Scenario	Recognition	Distance
People	91.2%	87.65%
Trees	74.98%	84.34%
Animals	80.87%	76.23%
Rock	61.42%	90.12%

Table 2: Success Rate

While the program performed well in most scenarios, recognition errors can be attributed to a partial view of the object as well as the object size. The height of the grass in some scenarios also prevented the camera from obtaining exactly where on the ground the object was located, thus, yielding some distance estimation and recognition errors.

VI. FUTURE WORK

The future of this project includes training the YOLO model to recognize objects typically found in the airport environment. To do so, a large dataset of these objects is needed to conduct training, validation, and testing. Other advancement strategies include obtaining more information about the object such as its trajectory and to develop additional avoidance maneuvers. A method can also be designed such that the dataset can be modified to any environment. Furthermore, the distance estimation algorithm can also be more efficient if the y-values did not need to be obtained experimentally. This process can be automated with machine learning or through other mathematical computations to convert pixels to distance. By doing so, this method can be applied to various hardware tools.

VII. CONCLUSION

This paper presented an object detection implementation for an autonomous lawnmower. The purpose of extending the obstacle detection program is to recognize objects typically found in an airport (planes, luggage, airport signers, etc.). Future work includes training the YOLO model to recognize objects typically found in the airport environment. These programs were designed to provide computer vision capabilities to an autonomous lawn mower. Finally, a distance estimation technique is implemented to notify the system when the object is within a certain vicinity of the unit.

VIII. REFERENCES

[1] C. -J. Lee, T. -H. Tseng, B. -J. Huang, Jun-Weihsieh and C. -M. Tsai, "Obstacle detection and avoidance via cascade classifier for wheeled mobile robot," 2015 International Conference on Machine Learning and Cybernetics (ICMLC), 2015, pp. 403-407, DOI: 10.1109/ICMLC.2015.7340955.

[2] Y. Zhang, X. Xu, H. Lu and Y. Dai, "Two-Stage Obstacle Detection Based on Stereo Vision in Unstructured Environment," 2014 Sixth International Conference on Intelligent Human-Machine Systems and Cybernetics, 2014, pp. 168-172, DOI: 10.1109/IHMSC.2014.49.

[3] M. Franzius, M. Dunn, N. Einecke and R. Dirnberger, "Embedded Robust Visual Obstacle Detection on Autonomous Lawn Mowers," 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2017, pp. 361-369, DOI: 10.1109/CVPRW.2017.50.

[4] Youchun Xu et al., "A method of stereo obstacle detection based on image symmetrical move," 2009 IEEE Intelligent Vehicles Symposium, 2009, pp. 36-41, DOI: 10.1109/IVS.2009.5164249.

[5] H. Liu, Z. Yuan and Z. Su, "Design and realization of visual wireless autonomous lawn mower based on machine vision," 2014 11th International Computer Conference on Wavelet Actiev Media Technology and Information Processing(ICCWAMTIP), 2014, pp. 353-358, DOI: 10.1109/ICCWAMTIP.2014.7073425.

[6] T. Yao, S. Dai, P. Wang and Y. He, "Image based obstacle detection for automatic train supervision," 2012 5th International Congress on Image and Signal Processing, 2012, pp. 1267-1270, DOI: 10.1109/CISP.2012.6469703.

[7] S. T. Blue and M. Brindha, "Edge detection based boundary box construction algorithm for improving the precision of object detection in YOLOv3," 2019 10th International Conference on Computing, Communication and Networking

Technologies (ICCCNT), 2019, pp. 1- 5, DOI: 10.1109/ICCCNT45670.2019.8944852.

[8] B. Epshtein and S. Uliman, "Feature hierarchies for object classification," Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1, 2005, pp. 220-227 Vol. 1, DOI: 10.1109/ICCV.2005.98.

[9] Object detection guide. Fritz. (n.d.). Retrieved February 25, 2022, from https://www.fritz.ai/object-detection/

[10] Brownlee, J. (2020, April 16). How do convolutional layers work in Deep Learning Neural Networks? Machine Learning Mastery. Retrieved December 15, 2021, from https://machinelearningmastery.com/convolutional-layers-for-deep- learning-neural-networks/

[11] V7labs.com. 2022. YOLO: Real-Time Object Detection Explained. [online] Available at: https://www.v7labs.com/blog/yolo-object-detection [Accessed 25 February 2022]

[12] Brownlee, J. (2021, January 26). A gentle introduction to object recognition with deep learning. Machine Learning Mastery. Retrieved February 25, 2022, from https://machinelearningmastery.com/object-recognition-with-deep-learning/

[13] (2020, January 13). Image recognition with deep neural networks and its use cases. AltexSoft. Retrieved December 15, 2021