Robotic detection of leaf angle and grasp point with low-cost sensors

Junzhe Wu, Wangqi Ma, Girish Chowdhary

***Abstract*— Leaf angle is a significant architectural parameter of plant canopy due to its influence on solar light absorption and photosynthetic efficiency, and hence also on the plant growth and productivity. Traditional way for leaf angle measurement is manual, and researchers need to go to the field and use a protractor to collect leaf angle data. However, manual leaf measurement is labor intensive, not ergonomic, and often not consistent between different people. In this paper, TerraSentia, a field robotic system equipped with low-cost sensors is utilized to automatically detect leaf angle and grasp point. The proposed method can produce a consistent leaf rolling angle estimate quantitatively and qualitatively on multiple corn leaves, especially on leaves with multiple different angles.**

Keywords - Neural network, Sensor Fusion, Leaf angle, Grasp point

I. Introduction

From the United Nations, the world population is increasing by around 1.13 percent per year and will grow from 7.4 billion in 2016 to 8.1 billion in 2025. By 2050, the world population is projected to reach 9.7 billion. Urbanization will continue to develop at a rapid pace with the numerous income increases, and urban populations will make up about 70% of the world's population, compared to 49% today. To feed these increased populations, especially the richer population in the urban area, food production must increase by 70% [1]. In order to solve this problem, the efficiency of food production should be improved in technical ways such as plant breeding. Maize, the world's most productive crop, is important as human food and animal feed, therefore plant breeders in the world aim to cultivate high-yield and stress-tolerant maize varieties [2]. However, yield is one of the most difficult traits to inherit in plant breeding [3], and therefore leaf angle is selected as the main trait for breeding. Leaf angle is a significant architectural parameter of plant canopy due to its influence on solar light absorption and photosynthetic efficiency, and hence also on the plant growth and productivity [4]. A mass of leaf angle data from different plants in the field is required for the breeding of maize leaf angle. Traditional way for leaf angle measurement is manual which means researchers need to go to the field at fixed times or at particular phenological stages and use a protractor to collect leaf angle data [5]. The temperature in summer is hot, and some crops like corn are difficult to walk through. Therefore, the manual way of leaf angle measurement is slow and costly, and there is an urgent need for an automated method which can produce a consistent leaf angle estimate quantitatively and qualitatively on multiple corn leaves. Most automated methods of leaf angle measurement aim at reconstructing a three-dimensional point cloud of leaves and stems [6]. However, for large crops like maize, reconstruction from a single camera is not sufficient [6], thus sensor fusion technology is utilized in this paper.

J. Wu, W. Ma and G. Chowdhary are with the Agricultural & Biological Engineering Department, University of Illinois Urbana-Champaign, Champaign, IL 61820, USA.

Sensor fusion is widely adopted in the autonomous driving field because the types of road scenarios in real urban environments are diverse and can change rapidly where only one kind of sensor is not capable of getting all the significant information from the environment. The same sensor fusion technology can also be applied for recognition of fruits in the automated harvesting, and the exact 3D coordinates of detected objects can be obtained from point cloud data which benefits the future steps in automated harvesting. [7] developed an automatic apple recognition method which applied sensor fusion before the classification. Depth camera was used in this study to produce fused colored point cloud data, and an RGB-based segmentation algorithm was used to get colored point cloud data of apple as a dataset. Then an improved 3D descriptor was utilized to extract color and 3D features from the dataset, and these features were fed to a SVM-based classifier. This classifier is capable of predicting apples in 3D bounding boxes with 92.3% accuracy. [8] proposed another method which applied sensor fusion after the classification to recognize the green pepper and estimated the 3D pose of the stem for the next cutting step. Machine vision technology was firstly used to segment the green pepper from the background of leaves. After that, coherent point drift algorithm was applied to project the points from LiDAR to the image plane, and the point cloud inside the recognition area was filtered to calculate the 3D coordinates of detected green pepper's stem. The detection result from this method is not accurate enough where some leaves are mistakenly identified as green peppers. Hence, deep learning methods should be used for the recognition to obtain a more accurate detection result from the image.

The main contributions of this paper include: (1) A neural network was trained to detect the leaves with high horizontal level, (2) Point cloud data from depth camera and vision data from camera were combined via sensor fusion to get the leaf angle and grasp point. The rest of the paper is organized as follows. In Section II, the field robotic system is introduced. In Section III, the sensor fusion framework for leaf rolling angle estimation and grasp point selection are detailed.In Section IV, the experimental results of leaf rolling angle are discussed. In Section V, a conclusion of this paper and suggested future work are provided.

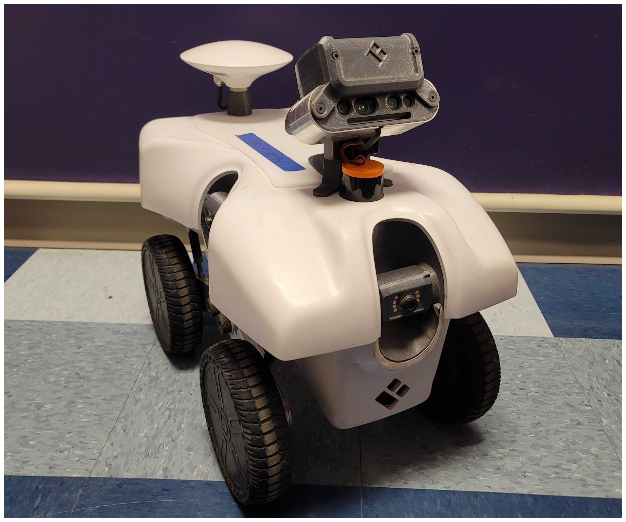

Fig. 1. The field robotic system

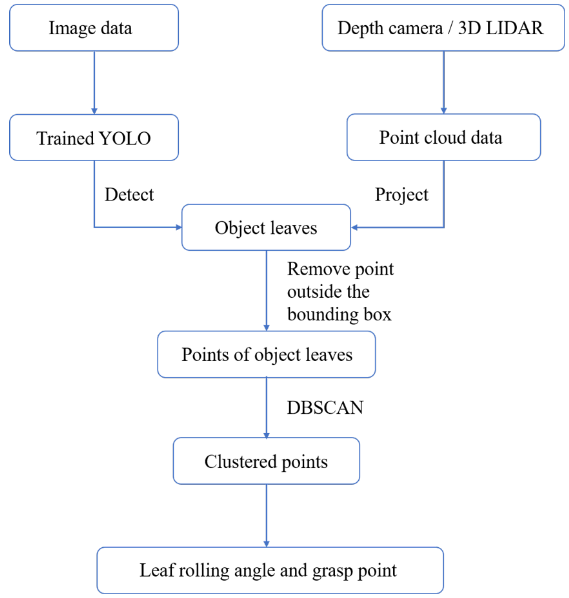

Fig. 2. Framework of the proposed sensor fusion method

II. Field Robotic System

A Field Robotic System, termed TerraSentia, was developed in the Distributed Autonomous Systems Laboratory [9]. This agricultural robot is designed to autonomously navigate between the rows of corn and collect information such as corn number, stem height and leaf angle for breeding and research. It is also equipped with a range of sensors as shown in Fig. 1. One real-time kinematic (RTK) GPS was mounted on the rear of the top of the robot for the navigation. The depth camera is installed on the top for the sensor fusion task. Two 2D LiDARs are installed on the top and trailing end of the robot respectively to acquire points from the horizontal plane and moving vertical plane. The resolution angle of the 2D LiDAR is 0.25°, the range of it is 270°, and it measures data at 40Hz. Images are recorded with a USB board camera mounted on the three sides of the robot to get information from the surrounding environment. The frame rate of the normal camera and depth camera is 30 fps when the resolution is 1920 x 1080 pixels.

III. Methodology

In this application, image data and 3D point cloud data are used for the sensor fusion. As illustrated in Fig. 2, image data is first fed into the YOLO network for object detection, then the corresponding 3D point cloud data is projected into the image plane. After that, only the points inside the bounding boxes remain and the DBSCAN algorithm is used to cluster these residual points. Lastly, each cluster of points can be used to estimate leaf rolling angle and grasp point.

A. Neural Network for Detection

There are diverse neural network architectures for image processing currently benefiting from the rapid development of deep learning technology in recent years. However, YOLOv3 [10] is utilized in this paper based on the following reasons: Firstly, YOLOv3 applies a single neural network to divide the full image into regions, then bounding boxes and probabilities are predicted for each region. Secondly, YOLOv3 is extremely fast and accurate.

B. Dataset and Training

In the detection process, an appropriate corn leaf should be selected at first. The corn leaves with high horizontal level are the perfect object since they are easier for future manipulation. A total of 1100 images are labeled, and 1000 images are randomly selected as the training dataset while the rest 100 images are the testing dataset. The data augmentation technology applied during training included horizontally flipping, translation, scaling, and brightness adjustment. The network was trained for 200 epochs with an initial learning rate of 0.001, a learning rate scheduler which scales the learning rate with 0.1 after each 10 steps, and a batch size of 16. The average precision measured at .5 IOU threshold of 200^{th} epoch is 71.10%.

C. Projection of Point Cloud data

As shown in Fig. 3, the projected points in the right image have the same shape as the leaves in the bounding boxes in the leaf image which demonstrates the accuracy of the extrinsic matrix. In the right image, the black part is the missing points in measurement from the depth camera, and for other parts inside the bounding boxes, the color is whiter when the distance is larger. The point cloud data from a depth camera is accurate when the distance is small, but the number of missed points is increasing when the distance becomes larger, thus the points whose distance is larger than the threshold is deprecated in the next DBSCAN [11] algorithm. In the clustering result for the residual point cloud inside the bounding boxes (fig. 4), one color represents one cluster of data, and the algorithm works well on most parts of the point cloud except the point cloud of overlapped leaves. After that, one cluster of data is selected based on a factor which is

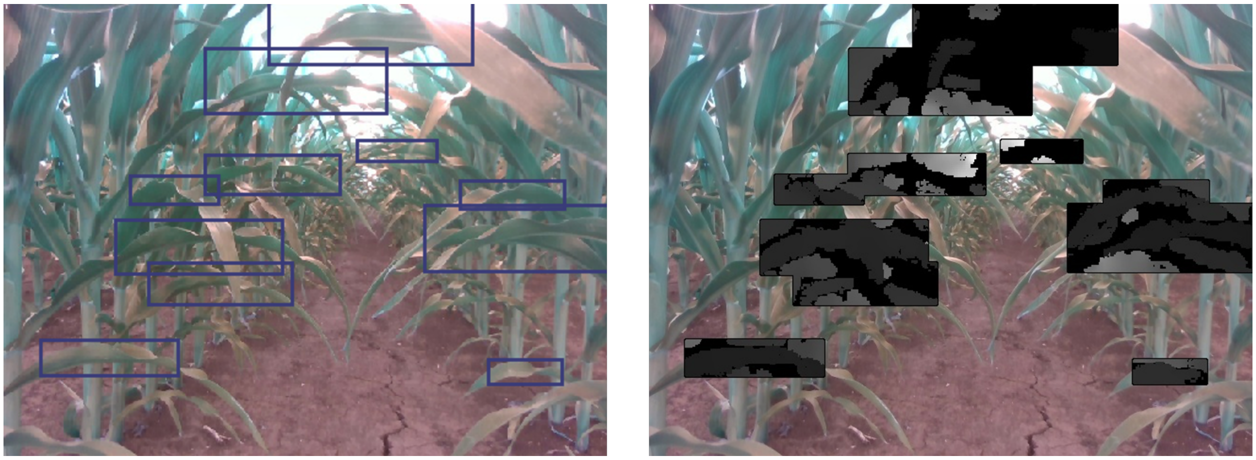

Fig. 3. Projection of point cloud from Depth camera in the corn field

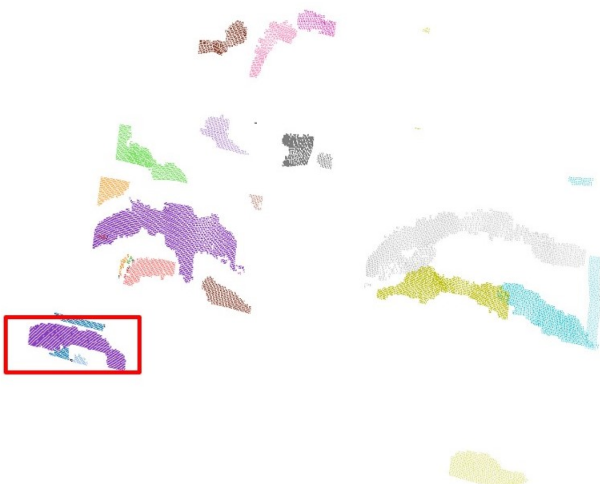

Fig. 4. The clustering result for the residual point cloud inside the bounding boxes

proportional to quantity of points and inversely proportional to distance. In this circumstance, the selected data is the purple cluster inside the red box.

D. Leaf Rolling Angle Detection and Grasp point Detection

As illustrated in Fig 5, the selected leaf is slightly rolling and there are three different rolling angles in this leaf which are obvious in the point cloud of this leaf, thus point cloud data is used to compute the average rolling angles for these three parts.

Computing the leaf rolling angle is equal to computing the normal vector of point cloud surface which is usually estimated directly from normal vectors of each point in point cloud.The normal vectors of a point can be approximated by calculating the normal vectors of the plane fitted according to the points in the neighborhood, thus the original problem is transformed into the least square plane fitting estimation

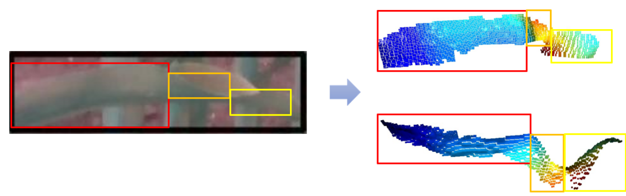

Fig. 5. The clustering result for the residual point cloud inside the bounding boxes

problem presented as follow equation:

$$\min_{A,B,C,D} \sum_{i=1}^{n} (Ax_i + By_i + Cz_i + D)^2 \quad \text{s.t.} \quad A^2 + B^2 + C^2 = 1 \tag{1}$$

Where x_i, y_i, z_i are coordinates of points in the neighborhood, and A, B, C, D are the coefficients of the 3D plane.

By taking the derivation, setting it equal to 0, and eliminating D in the equation set, the following linear system of equations can be obtained:

$$M \begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} \overline{x^2} - \bar{x}^2 & \overline{xy} - \bar{x}\bar{y} & \overline{xz} - \bar{x}\bar{z} \\ \overline{xy} - \bar{x}\bar{y} & \overline{y^2} - \bar{y}^2 & \overline{yz} - \bar{y}\bar{z} \\ \overline{xz} - \bar{x}\bar{z} & \overline{yz} - \bar{y}\bar{z} & \overline{z^2} - \bar{z}^2 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \end{bmatrix} = 0 \tag{2}$$

$$\text{s.t.} \quad A^2 + B^2 + C^2 = 1$$

Where M is covariance matrix, $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, $\overline{xy} = \frac{1}{n}\sum_{i=1}^{n} x_i y_i$, and the rest of the algebraic expressions are in a similar fashion.

In general, the covariance matrix is non-singular, so there is no exact solution to the above equation, but PCA [12] can be used to obtain the estimated solution which is the normalized eigenvector corresponding to the minimum eigenvalue of the covariance matrix M. Then the angles between each normal vector and angle vector [0,1,0] are computed to get the leaf rolling angle distribution (fig. 6). We can assume this distribution is the mixture of several Gaussian distributions and use the EM algorithm [13] to get the average and variance of these Gaussian distributions where k-

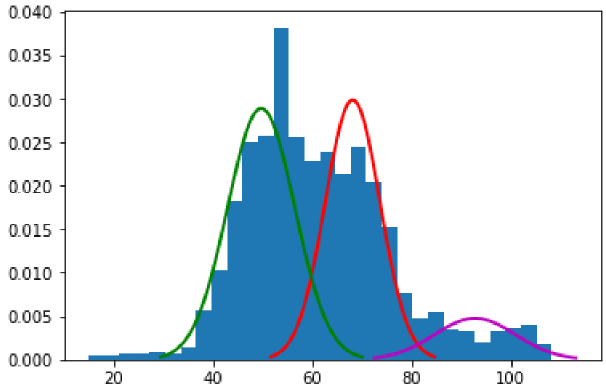

Fig. 6. The clustering result for the residual point cloud inside the bounding boxes

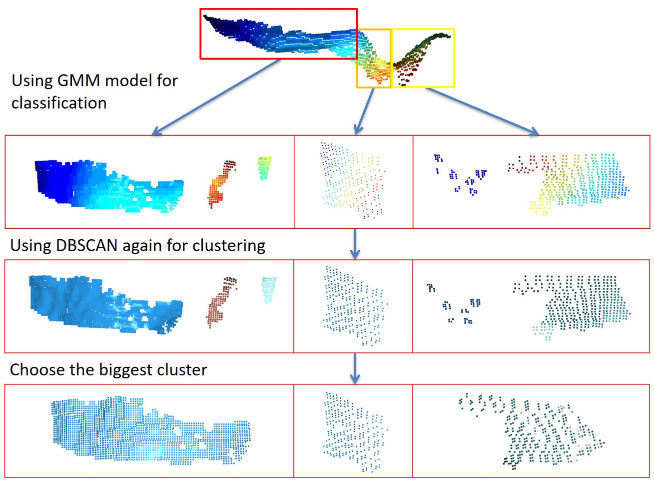

Fig. 7. The clustering result for the residual point cloud inside the bounding boxes

means algorithm [14] is used for the initialization. However, the right number k of clusters is not obvious, thus the G-means algorithm [15] is applied to automatically choose k. The obtained means are [68.16245498, 92.82219438, 49.73753005] which are the average leaf rolling angles of the selected leaf, and the obtained variances are [30.41397885, 66.67874526, 45.74359508]. The regular density function of these three Gaussian components is shown in Fig. 6. After that, the obtained means, variances, and weights are fed into the Gaussian Mixture Model for the classification. As shown in Fig. 7, the left and right clusters in the classification result from the GMM model have some misclassified points since there might be similar angles in different parts of leaves. This misclassification could be solved by using the DBSCAN algorithm again and then choosing the biggest cluster. Finally, the coordinates of grasp points are computed by averaging the coordinates of points in three chosen clusters.

IV. Experiments and Results

After the real corn withered, the plastic corn models in our lab are used for experimental leaf rolling angle measurement. In this experiment, the collected data were fed into the pipeline to get the detection result and computed leaf rolling angle. After that, the detected leaves in the real world were found according to the detection result in the image, then a protractor was used to measure the leaf rolling angle. As indicated in the experimental result, totally 46 angles from 24 leaves were measured, and the root mean square error (RMSE) is 6.53 which is acceptable considering the error in the manual measurement. The scatter diagram of measured angle and computed angle is shown in the Fig 8.

V. Conclusion and Future Work

In this paper, a YOLOv3 model was well trained for the detection of leaves with high horizontal levels. Moreover, an innovative pipeline using sensor fusion was developed to compute the leaf surface orientation and optimal punching position. In this pipeline, different sensors were calibrated to a unified coordinate system, then the point cloud data were projected to the image plane to match detected leaves. With these isolated leaf point cloud inside the bounding boxes, DBSCAN was utilized for clustering, and the normal vectors of each point in one cluster were calculated to get the leaf rolling angle distribution, then a Gaussian mixture model was applied to compute the multiple different rolling angles in one leaf. In the future research, the detector can be further optimized by finding a large and related dataset and using the transfer learning technology. The detector models are also updated in YOLOv4 [16] and YOLOv5 [17], thus our dataset can be fed into the new models to get better performance. Moreover, the point cloud of overlapped leaves cannot be separated by the DBSCAN algorithm which can cause problems in the calculation of leaf rolling angle and grasp point. In order to avoid these problems, the labeled overlapped leaves in the existing dataset should be deleted, and another model needs to be trained to no longer detect overlapped leaves.

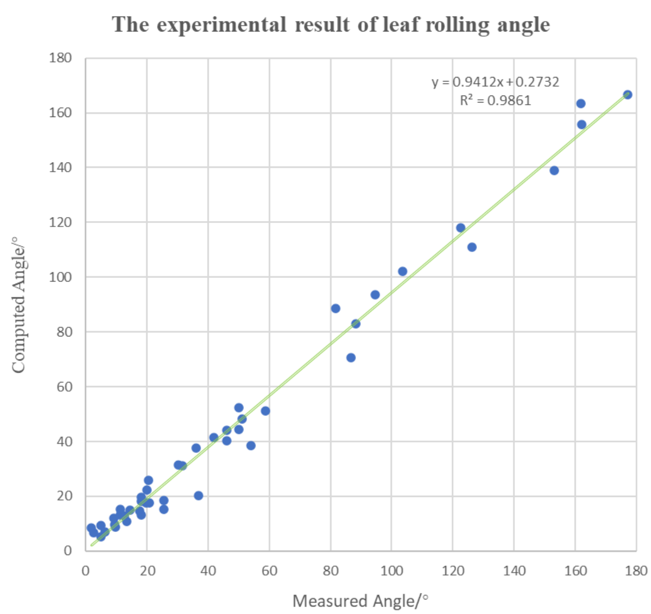

Fig. 8. The clustering result for the residual point cloud inside the bounding boxes

REFERENCES

[1] A. D. Tripathi, R. Mishra, K. K. Maurya, R. B. Singh, and D. W. Wilson, "Estimates for world population and global food availability for global health," in *The role of functional food security in global health*, pp. 3–24, Elsevier, 2019.

[2] Y. Che, Q. Wang, Z. Xie, L. Zhou, S. Li, F. Hui, X. Wang, B. Li, and Y. Ma, "Estimation of maize plant height and leaf area index dynamics using an unmanned aerial vehicle with oblique and nadir photography," *Annals of botany*, vol. 126, no. 4, pp. 765–773, 2020.

[3] R. A. Richards, G. J. Rebetzke, M. Watt, A. T. Condon, W. Spielmeyer, and R. Dolferus, "Breeding for improved water productivity in temperate cereals: phenotyping, quantitative trait loci, markers and the selection environment," *Functional Plant Biology*, vol. 37, no. 2, pp. 85–97, 2010.

[4] M. B. Mantilla-Perez and M. G. Salas Fernandez, "Differential manipulation of leaf angle throughout the canopy: current status and prospects," *Journal of Experimental Botany*, vol. 68, no. 21-22, pp. 5699–5717, 2017.

[5] R. T. Furbank and M. Tester, "Phenomics–technologies to relieve the phenotyping bottleneck," *Trends in plant science*, vol. 16, no. 12, pp. 635–644, 2011.

[6] M. Minervini, H. Scharr, and S. A. Tsaftaris, "Image analysis: the new bottleneck in plant phenotyping [applications corner]," *IEEE signal processing magazine*, vol. 32, no. 4, pp. 126–131, 2015.

[7] Y. Tao and J. Zhou, "Automatic apple recognition based on the fusion of color and 3d feature for robotic fruit picking," *Computers and electronics in agriculture*, vol. 142, pp. 388–396, 2017.

[8] P. Eizentals and K. Oka, "3d pose estimation of green pepper fruit for automated harvesting," *Computers and Electronics in Agriculture*, vol. 128, pp. 127–140, 2016.

[9] Z. Zhang, E. Kayacan, B. Thompson, and G. Chowdhary, "High precision control and deep learning-based corn stand counting algorithms for agricultural robot," *Autonomous Robots*, vol. 44, no. 7, pp. 1289–1302, 2020.

[10] A. Farhadi and J. Redmon, "Yolov3: An incremental improvement," in *Computer Vision and Pattern Recognition*, pp. 1804–02, Springer Berlin/Heidelberg, Germany, 2018.

[11] M. Ester, H.-P. Kriegel, J. Sander, X. Xu, *et al.*, "A density-based algorithm for discovering clusters in large spatial databases with noise.," in *kdd*, vol. 96, pp. 226–231, 1996.

[12] G. H. Dunteman, *Principal components analysis*. No. 69, Sage, 1989.

[13] G. Xuan, W. Zhang, and P. Chai, "Em algorithms of gaussian mixture model and hidden markov model," in *Proceedings 2001 International Conference on Image Processing (Cat. No. 01CH37205)*, vol. 1, pp. 145–148, IEEE, 2001.

[14] K. Wagstaff, C. Cardie, S. Rogers, S. Schrödl, *et al.*, "Constrained k-means clustering with background knowledge," in *Icml*, vol. 1, pp. 577–584, 2001.

[15] G. Hamerly and C. Elkan, "Learning the k in k-means," *Advances in neural information processing systems*, vol. 16, pp. 281–288, 2003.

[16] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "Yolov4: Optimal speed and accuracy of object detection," *arXiv preprint arXiv:2004.10934*, 2020.

[17] J. Glenn, *Yolov5*, 2020.