BOROS AI. The Computer Visualization Chemist

Yehyun Bin[1], Jihwan Oh[2] and Jihoon Park[3]

[1]Cornerstone Collagate Academy of Seoul, Seoul, South Korea
[2] Vivan and Stanley Gangnam International Scholars, Seoul, South Korea
[3]The Stony Brook School, New York, US

binkano@gmail.com, john.hwan0919@gmail.com, and theprun06@gmail.com

Abstract – Computer visualization is being utilized in various portions of society. Facial recognition, item identification and separation, and high-speed movement isolation, are some examples of how computer visualization is working as an eye of an automotive process. Quality control in a factory setting is one area where computer visualization is being aggressively used to maintain a consistent environment. However, industrial settings which deal with chemicals, non-transparent liquids, and various glassware have yet to utilize computer visualization to the degree of other industrial fields. Herein, we propose a computer visualization algorithm, BOROS AI, which is based on the YOLO11 computer visualization model. The parameters of the BOROS AI, have been deliberately tuned to function properly in visually unfavorably environments. Various tests have been conducted to assess the BOROS AI's potential when the target objects are at different

distances, have different material composition, and are placed in an obstructed and complex manner.

Keywords: Glassware, Computer Visualization, YOLO11, Computer Chemist, Laboratory Apparatus Identification

1. Introduction

While brilliant human chemists are at the forefront of advancements in chemical science, sometimes even they are forced to do mundane tasks which take up a significant amount of time. As such, an innocent yearning for a automated chemical assistant are automaton was sometimes uttered from the lips of

With the introduction of artificial intelligence technology in societies' everyday lives, the precision and accuracy with which it is utilized has come under the spotlight. Previous endeavors to incorporate vision technology was mostly proof of concept, and such endeavors focused primary on making sure the technology worked, with minimal emphasis on fine-tuning. However, the technology has advanced to such a degree that programmers are now focused on increasing the efficiency of the visualization program. Parameters to consider are draw distance, light consideration, and image catch time.

One prospective object detection model is YOLO11. YOLO, which stands for You Only Look Once, was introduced in 2015. Recently released on the September of 2024, YOLO11 builds upon its previous versions to provide a real-time object detection model that can identify its target in a small span of time [1]. Specifically, the training time is worth noting, as it seems that the algorithm does not require an intensive and copious amount of data to train it [2]. As such, the small sample size of the input dataset is worth noting.

2. Background

2.1 What is YOLO11

YOLO11 is the latest iteration of an object detection algorithm. The history of YOLO started in 2015, and until its 4^{th} iteration, its framework was darknet based. It had its humble beginnings in basic classification and object detection, and by the time of its 4^{th} iteration, it was capable of object tracking and multi-object detection.

When YOLO1 was first introduced, it was a new computer system that could recognize objects by making a bounding box and had class label identification. YOLO2 first introduced anchor boxes increased YOLO's image accuracy; Anchor boxes allowed YOLO to detect objects' shapes and sizes more accurately. On YOLO3, new multi scale predictions were added. This made allowed YOLO3 to detect small object a layered and more detailed approach [3,4].

The most significant upgrade was when YOLO 4 was released. YOLO4 started to aggressively use AI, which meant that the program itself has a deep learning system, which in turn leads to improvements in image recognition. Also they focused on multi-task learning combining many tasks in one network. This improvement greatly improved the system's efficiency [5,6].

From 2020, YOLO started to use the Pytorch framework, and from the 5^{th} version, YOLO started to utilized instance segmentation [2,7]. In a short time of less than 10 years, computer vision technology has seen astonishing improvements annually. Fundamentally speaking, each new version of the YOLO computer vision technology has improved adaptability, segmentation, estimation, and most importantly, object detection.

YOLO has the strength in that it can be used anywhere when people need recognition. For example, they can be used in agricultural fields when they need to monitor crop and livestock management. Also, it could be used in factory security, and wildfire detection. From these examples, people can notice that YOLO can be used in any different field when they needed, and it is strong that it could use when humans cannot be recognized every time. Also YOLO makes the image single pass, which means that when the program predicts and recognizes the object; it is efficient, which makes it faster and better result.

For this research, YOLO11 was chosen as the main algorithm. YOLO11 shows rapid image identification and the ability to identify an object in real time [8,9]. Additionally it could efficiently detect objects that are moving and it training protocol required a relatively smaller library for training.

2.2 Computer Visualization Mechanism.

Although the research was based on the YOLO11, it was unanimously decided that name of the algorithm itself did not properly convey the purpose of the team. The AI was renamed to BOROS, to properly reflect the fact most chemical glass apparatuses are made of borosilicate glass.

Figure 1 shows the basic operating procedure of the BOROS AI. A camera serves as the initial means of object identification, and then the BOROS AI proceeds to label the object in question in a short time span. The BOROS AI has been trained to properly identify 4 different types of chemical equipment: Erlenmeyer Flasks, Round Bottom Flasks, Graduated Cylinders, and Beakers.

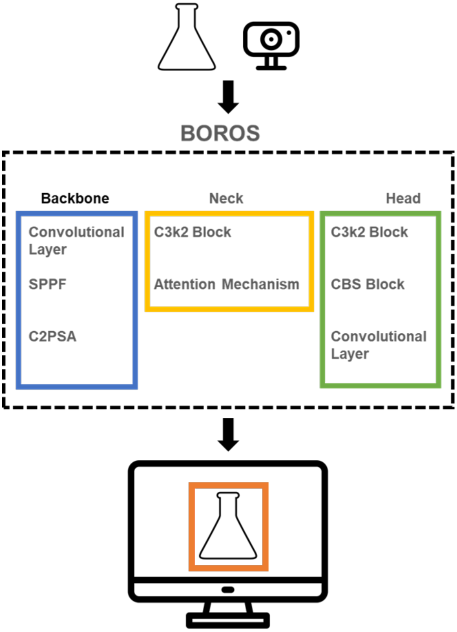

Figure 1. Schematic of the BOROS AI

3. Materials and Methods

3.1 BOROS AI Training

The first step regarding the training of the BOROS AI was obtaining a chemical apparatus library. This was done by taking real life photos of the previously mentioned chemical glassware. Additionally, stock images available online was also incorporated into the BOROS AI training.

A total of 1420 images total were used for the training of the BOROS AI (**Figure 2)**. Considering the fact that several thousand images are usually used for computer visualization training, the YOLO11 architecture within the BOROS AI stood true to its advertised strength; less training required compared to previous versions.

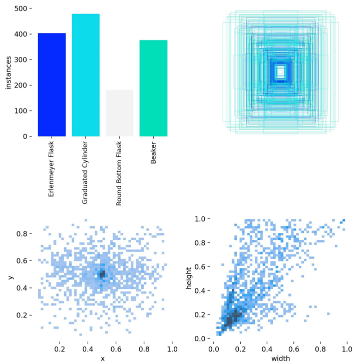

Figure 2: Image quantity for different glassware

3.2 BOROS AI Training Result Analysis

Although the analysis results showed a number of quantitative analysis results, only the four most important graphs will be shown.

Figure 3. displays multiple line plots that show the training progress and performance metrics of the model. Each plot represents a different metric or loss function tracked during the training process.

For all four images, the x-axis indicated the Epoch number, or how many times a mock trial occurred. More specifically, and epoch is a complete repetition of the training data set through the neutral network during training. The more epochs there are, the better the model is trained, as can be seen in the decrease in loss values as the epoch number increases. However, it is imperative to note that during the learning of fundamental patterns, if the epoch number was too high (in excess) the model becomes overly complicated and fails to identify objects when subject to real test conditions.

The y-axis represents the loss value. Loss can be an indicator or measure of how well the model is performing during training. Better model performance is shown by lower loss values, and worse performance is indicated by larger values. Reducing the loss value is the aim during training.

There are two types of loss: box_loss, also known as bounding box regression loss and cls_loss, or classification loss. box_loss calculates the difference between the ground truth and the expected bounding box dimensions and coordinates. The estimated bounding boxes are more accurate when the box_loss is smaller. The ratio of the intersection of the prediction's box and the union would be nearly one if it matched the ground truth exactly. On the other hand cls_loss determines the difference between the ground truth and the expected class probabilities for each object in the picture. A lower cls_loss indicates that the model is more successful in predicting the objects' class.

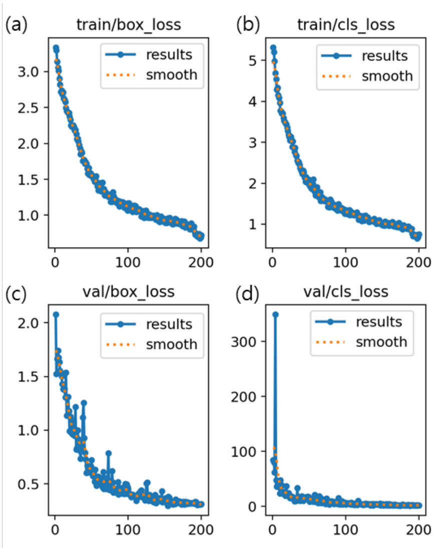

Figure 3. Loss graph at various Epoch values: (a) train/box_loss, (b) train/cls_loss, (c) val/box_loss and (d) val/cls_loss

For this research **Figure 3 (a)** evaluates how well the model predicts the bounding box (bbox) coordinates and the area around objects that are recognized during training. A lower value indicates that the anticipated bounding boxes are nearer the ground truth bounding boxes. As the model improves at locating the objects in the image, this loss diminishes. The degree to which these expected bounding box coordinates coincide with the excellent truth boxes is measured by the box loss.

Figure 3 (b) classification loss on the training data. This parameter tracks how accurately the model is classifying objects into the correct categories. A lower train/cls_loss means the model is making fewer mistakes when classifying objects.

Figure 3 (c) shows the bounding box loss on the validation set, which is obtained using the validation (activity of verifying the accuracy) data rather than the training data. It is calculated similarly to the train/box_loss and thus, the loss should follow a similar trend to the train/bos_loss.

Figure 3 (d) specifies the classification loss on the validation data, and it shows how well the model can classify objects in images it hasn't seen during training ower **val/cls_loss** indicates that the model generalizes well to new data. Arguably, **Figure 3 (d)** is the most important result of the training procedure.

3.2 Test Subject Identification Parameters

A number of parameters were tested in this particular experiment. Previous research had already proven that previous generations of the YOLO computer visualization model was capable of identifying glassware in a laboratory environment. Such laboratory environments are characterized by a black experimental table and a LUX reading of over 1,000 units. To test the limits of the BOROS AI, the image identification was conducted under a low light intensity environment which averaged 380 – 400 LUX units. Additionally, the background and table upon which the glassware placed was white to confuse the BOROS AI to the highest degree. Independent variable included length from camera, whether the chemical apparatus in questions was glass or plastic, and different volumes.

4. Experimental Results

4.1 Individual Object Identification

The first test of the BOROS AI focused on its ability to identify individual glassware types.

4.1.1 Glass Erlenmeyer flask Identification

An Erlenmeyer flask is typically characterized by a triangular shape. It has a round bottom which stays in contact with the surface upon which it stands. Additionally, it has a top that gradually narrows as it travels upward. When seen sideways, it is a triangle.

For the Glass Erlenmeyer flask identification test (**Figure 4)**, BOROS AI successfully identified at four distances: 30 cm, 60 cm, 90 cm, and 120 cm. Although the result at 30 cm showed a value of 0.64, the other three cases had values higher than 0.86.

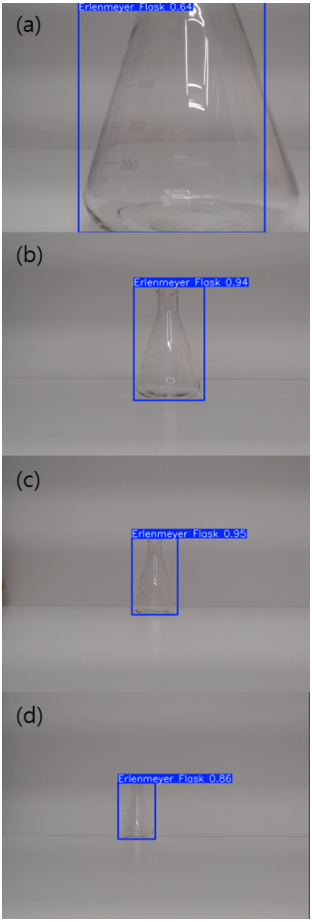

Figure 4: Glass Erlenmeyer flask at (a) 30 cm, (b) 60 cm, (c) 90 cm and (d) 120 cm.

4.1.2 Glass Graduated Cylinder Identification

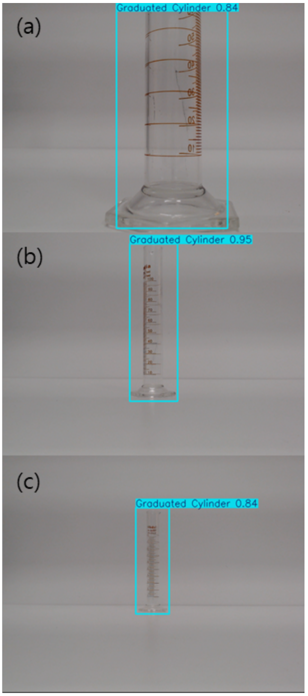

Figure 5: Glass graduated cylinder at (a) 30 cm, (b) 60 cm, and (c) 90 cm

A graduated cylinder is characterized as a narrow column that retains its circumference throughout the entirety of its shape. Some graduated cylinders have a larger base to help stabilize its footing.

For the graduated cylinder identification test (**Figure 5)**, BOROS AI successfully identified at three distances: 30 cm, 60 cm, 90 cm. All three distances showed acceptable values of 8.4 or above. However, the graduated cylinder was unable to be identified at a distance of 120 cm.

4.1.3 Round Flask Identification

A round flask, like is namesake, has a spherical bottom. Due to its area of contact with a flat surface, scotch tape was utilized to fasten the round flask to the table.

For the round flask identification test (**Figure 6)**, BOROS AI successfully identified at four distances: 30 cm, 60 cm, 90 cm, and 120 cm. All four distance yield impressive values with the lowest being 0.89

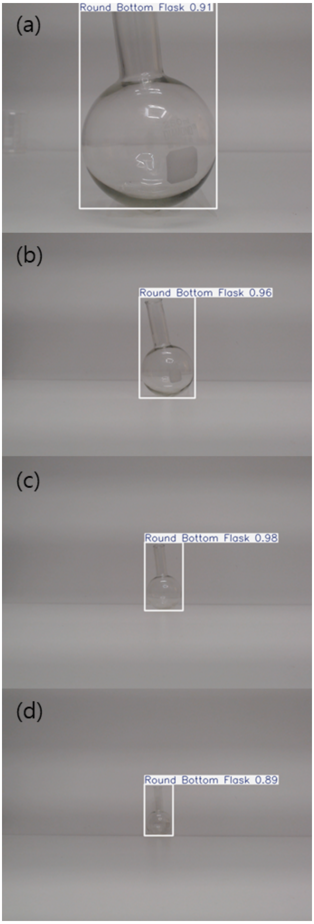

Figure 6: Glass round bottom flask at (a) 30 cm, (b) 60 cm, (c) 90 cm, and (d) 120 cm

One aspect to note is the relatively lower training volume for the round bottom flask. The uniqueness of its shape contributed to the ease with which BOROS AI was able to successfully and efficiently identify the round bottom flask.

4.1.4 Glass Beaker Identification

The beaker is arguably the most widely used laboratory apparatus during the early academic years. It possesses a circumference much larger than that of a graduated cylinder, and is consistent throughout its length. Additionally, it is capable of containing a greater amount of liquid compared to the graduated cylinder. As for the fonts and the sizes of the headings, this manuscript in itself constitutes a good example. The paper should be written in A4 (210mm

For the beaker identification test (**Figure 7**), BOROS AI successfully identified at three distances: 30 cm, 60 cm, 90 cm. All three distances showed values of 7.5 or above. However, the beaker was unable to be identified at a distance of 120 cm.

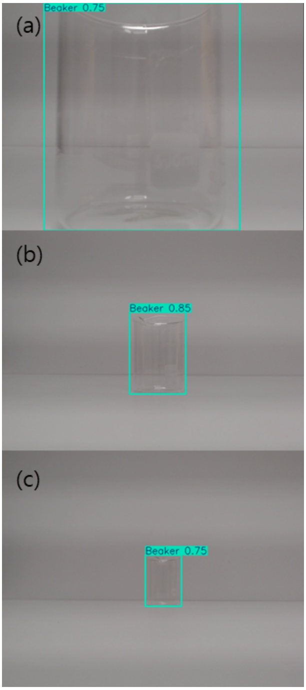

Figure 7: Glass beaker at (a) 30 cm, (b) 60 cm, and (c) 90 cm

4.2 Material and Size Diversification and Image Identification

The initial preliminary testing of the BOROS AI showed that it was capable of individual glassware identification. Distances of 60 cm and 90 cm showed on average, the highest values. Next, a more complicated experiment was conducted where the multiple apparatuses of different material composition and size were placed side-by-side.

Three Erlenmeyer flasks are set side by side at different distances (**Figure 6 (a) and (b))**. From left to right, the volume of each flask is 100 ml, 250 ml, and 500 ml respectively. The 100 ml and 250 ml are made of plastic, while the 500 ml is made of glass. All three flasks are successfully identified.

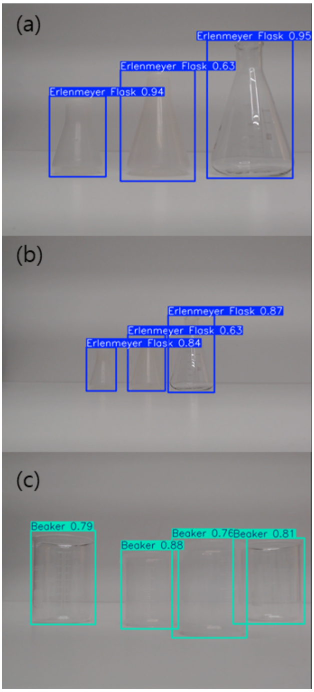

Figure 8: Different material and size combinations. (a) is two Erlenmeyer flasks followed by a larger glass Erlenmeyer flask at 60 cm and the same combination and (b) 90 cm. (c) is four beakers of different size and materials (glass, plastic, plastic, glass sequence) identified at 60 cm.

Next, four beakers of different material composition and volume were placed side-by-side. From left to right, the volume of each beaker was 500 ml, 250 ml, 500 ml, and 250 ml. The composition was glass, plastic, plastic, and glass. Once again the BOROS AI successfully identified each beaker within its scope.

4.3 Real Environment Simulation

The next step in testing the limits of the BOROS AI involved different object identification. In this portion, all four of the previous glassware convened on one position (**Figure 9)**.

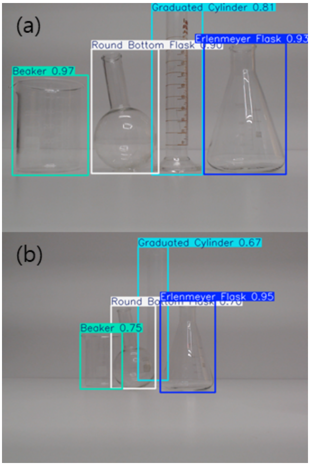

Figure 9. Different glass apparatuses (a) side-by-side at 60 cm and (b) obstructing each other at 90 cm.

When placed in an organized manner (**Figure 9 (a)**), BOROS AI was able to identify all four glassware with impressive results. The lowest value of 0.81 was given by the graduated cylinder while the other three glassware gave results of 9.0 and above.

When the four glassware were placed 30 cm further and in an obstructed manner, the values generally dropped compared to the values in the organized version. Nevertheless, BOROS AI was successful in its identification effort.

4.4 Disorganized Environment Simulation

The final experiment was conducted in a complex environment. Uneven draw distance, obstructed views, uneven glassware quantity distribution, and even vertical alignment were tested.

While all three situations took into account the previous parameters and then mixed them even more, the BOROS AI was successful in its attempts to identify each situation. Note the relatively low values of each situation.

Figure 10 (c) is especially of great interest. The chaos involved in this particular situation can be seen to mimic the disorganization that is characteristic of high school chemistry labs.

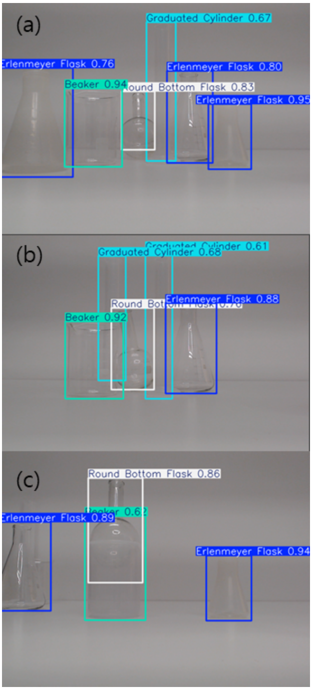

Figure 10: Complex environment simulation. (a) increases only one type of apparatus, (b) is a multi-obstruction environment, and (c) contains vertical placement.

That the BOROS AI was able identify such a situation illuminates its possible use as an advanced CCTV. Indeed, while CCTV can only record, the BOROS AI will be to properly isolate and identify what is happening inside a particular environment even in the absence of a human coordinator.

5. Conclusion

By utilizing the YOLO11 to his advertised strengths, the BOROS AI was born. At the moment, the BOROS AI is the first of its kind; it is the first YOLO11 used in to assess a laboratory environment. Previous researches have used YOLO to capture glassware before, but its operating capacity falls short compared to BOROS AI. The distance at which BOROS can identify objects is further than the most contemporary research comparison [13]. Additionally, no other research has sought to identify objects in generally unfavorable conditions, especially dark environments.

Due to the nature by which BOROS AI encounters data, its output is dependent on many external factors. Light reflection, light refraction, distance from the camera, and even the camera itself contribute to BOROS AI's output values. The effects of light are especially apparent when the glassware in question is not specified by any unique physical characteristic.

This was very apparent during the initial testing of the BOROS AI. When the object in question had unique shape, such as the triangle Erlenmeyer flask or the spherical round bottom flask, BOROS AI had an easier time than when it was assessing object without such physical characteristics. For instance, the round bottom flask had a shorter training session compared to its other counterparts **(Figure 2)**. Hoever, it triangler counterpart had much higher training volume. The difference in training volume may have arisen due to the different light reflection and refraction that occurs in the Erlenmeyer flask. The spherical character of the round bottom flask allows for a more uniform distribution of light rays at various incidence angles.

While the BOROS AI exceeded expectation, it has many ways to improve. It has yet to be tested against an moving object. Object identification during either sudden movement or steady movement should be the next logical step in improving the model.

References

[1] Ultralytics. (2024, November 7). *Yolo11* Ultralytics YOLO Docs. https://docs.ultralytics.com/models/yolo11/

[2] Khanam, R., & Hussain, M. (2024). YOLOv11: An Overview of the Key Architectural Enhancements.

[3] Dupont, M. (2024, August 21). *History of Yolo: From Yolov1 to yolov10*. Labelvisor. https://www.labelvisor.com/history-of-yolo-from-yolov1-to-yolov10/#:~:text=YOLO%2C%20or%20You%20Only%20Look,identification%20within%20a%20single%20network

[4] Buhl, N. (2024, November 5). *Yolo Object Detection explained: Evolution, algorithm, and applications*. Encord. https://encord.com/blog/yolo-object-detection-guide/#:~:text=YOLO%20requires%20a%20single%20forward,it%20faster%20and%20more%20efficien.

[5] AllegroAdmin1. (2022, October 29). *The battle of speed vs. accuracy: Single-shot vs two-shot detection meta-architecture*. ClearML. https://clear.ml/blog/the-battle-of-speed-accuracy-single-shot-vs-two-shot-detection

[6] Sara , G. (2024, August 20*). Yolo model for real-time object detection: A full guide: Viam*. RSS. https://www.viam.com/post/guide-yolo-model-real-time-object-detection-with-examples.

[7] Qin, Y., Pan, Z., & Shao, C. (2023). Defect detection method of phosphor in glass based on improved yolo5 algorithm. *Electronics*, 12(18), 3917. https://doi.org/10.3390/electronics12183917.

[8] Sapkota, R., Meng, Z., Churuvija, M., Du, X., Ma, Z., & Karkee, M. (2024). *Comprehensive Performance Evaluation of Yolo11, Yolov10, Yolov9 and Yolov8 on Detecting and Counting Fruitlet in Complex Orchard Environments*. https://doi.org/10.36227/techrxiv.172954111.18265256/v1.

[9] Sapkota, R., & Karkee, M. (2024). Yolo11 and Vision Transformers Based 3D Pose Estimation of Immature Green Fruits in Commercial Apple Orchards for Robotic Thinning. https://doi.org/10.36227/techrxiv.173014437.72236643/v1

[9] Sapkota, R., & Karkee, M. (2024). Yolo11 and Vision Transformers Based 3D Pose Estimation of Immature Green Fruits in Commercial Apple Orchards for Robotic Thinning. https://doi.org/10.36227/techrxiv.173014437.72236643/v1

[10] Ibrahim, M. (2024, May 17*). Yolov5 object detection tutorial*. W&B. https://wandb.ai/mostafaibrahim17/ml-articles/reports/YOLOv5-object-detection-tutorial--Vmlldzo3OTk1NTQ4.

[11] Mamunur, R. (2022, May 24). *Yolo V8 Result Grap*. Stack Overflow. https://stackoverflow.com/questions/75809415/yolo-v8-result-grap.

[12] Egele, R., Mohr, F., Viering, T., & Balaprakash, P. (2024). The unreasonable effectiveness of early discarding after one epoch in neural network hyperparameter optimization. *Neurocomputing*, 597, 127964. https://doi.org/10.1016/j.neucom.2024.127964

[13] Cheng, X., Zhu, S., Wang, Z., Wang, C., Chen, X., Zhu, Q., & Xie, L. (2023). Intelligent vision for the detection of chemistry glassware toward AI robotic chemists. *Artificial Intelligence Chemistry*, *1*(2), 100016. https://doi.org/10.1016/j.aichem.2023.100016

Yehyun Bin
She is currently a junior at 1Cornerstone Collagate Academy of Seoul, South Korea

Jihwan Oh
He is currently a junior at Vivan and Stanley Gangnam International Scholars, Seoul, South Korea

Jihoon Park
He is currently a junior at The Stony Brook School, New York, US