# Deep Learning Model to Predict the Ripeness of Oil Palm Fruit

Keywords: Fruit Ripeness Classification, Oil Palm, YOLO.

Abstract:

This study explores the application of deep learning, specifically the YOLOv8 model, for predicting the ripeness of oil palm fruit bunch through digital images. Recognizing the economic importance of oil palm cultivation, precise maturity assessment is crucial for optimizing harvesting decisions and overall productivity. Traditional methods relying on visual inspections and manual sampling are labor-intensive and subjective. Leveraging deep learning techniques, the study aims to automate maturity classification, addressing limitations of prior methodologies. The YOLOv8 model exhibits promising metrics, achieving high precision and recall values. Practical applications include deployment in production areas and real-time field scenarios, enhancing overall production processes. Despite excellent metric results, the model shows potential for further improvement with additional training data. The research highlights the effectiveness of YOLOv8 in automating the ripeness classification oil palm fruit bunches, contributing to sustainable cultivation practices in diverse agricultural settings.

## 1 INTRODUCTION

Colombia stands as the largest producer of oil palm (Elaeis guineensis) in the Americas and ranks fourth globally, annually yielding millions of tons (Fedepalmas, 2019). The economic and social significance of this cultivation has spurred interest in sustainable development models. Within the global vegetable oil market, oil palm holds a pivotal role, serving as a primary source for industries such as food, cosmetics, and biofuels. The economic importance has led to a heightened focus on optimizing cultivation practices to meet the growing demand (Corley & Tinker, 2015).

A critical aspect of the cultivation process is the precise assessment of oil palm fruit bunch maturity, influencing harvesting decisions and overall productivity. The quality of palm oil is deeply intertwined with the ripeness of the fruit. It's ideal for the fruit to reach an exact level of ripeness, steering clear of both being excessively green or overly ripe

extremes. When the fruit is immature, it lacks sufficient oil content, and when overly ripe, it detaches too easily, leading to significant oil loss as the fruits separate from the bunches. The classification of the fruits is closely tied to how easily they detach from the bunch and a certain color change, ultimately dependent on individual experience and visual assessment.

Traditional methods relying on visual inspections, manual sampling, and expert judgment are labortime-consuming, intensive. and subjective. potentially introducing inaccuracies. Classifying the fruits is closely associated with how easily they detach from the bunch and a certain change in color, which ultimately relies on individual experience and visual assessment. Many studies have been conducted, some related to computer vision attempts based on color. Others have used sensors, which also did not provide great outcomes, partly because there are different palm varieties that alter the shape of bunch and fruits (Lai et al., 2023).

alp https://orcid.org/0000-0002-3031-2334

b https://orcid.org/0000-0002-7135-2092

https://orcid.org/0000-0003-1651-3831

The advent of deep learning, particularly convolutional neural networks (CNN), offers a transformative avenue to automate complex visual tasks, including image classification. The application of deep learning techniques to agricultural processes, such as maturity assessment, has shown promising results in enhancing accuracy and efficiency (Mohanty et al., 2016).

One of the major challenges in conducting predictive work regarding the ripeness of oil palm fruit bunches lies in acquiring appropriate images for fruit maturity detection. Typically, fruits are segregated, and images are captured either when they are on the ground or while still on the tree (Suharjito et al., 2023).

There are two critical moments requiring fruit maturity classification: 1) while the fruit is still on the tree to determine the optimal harvesting time and 2) when it's within the production area before oil extraction. However, it's uncommon to find work or images of fruits during this latter stage, despite it being arguably the most crucial. Large companies usually have fruit suppliers, and accurately classifying incoming fruit is essential. Additionally, for a final evaluation of one's own fruits, determining their maturity is crucial.

In fruit unloading zones, there are often inclined ramps or reception platforms where fruits are transported from trucks to the oil extraction area. During transit on these ramps, fruit bunches are typically not well-separated and may stack on top of one another. Our aim is precisely to develop a model capable of classifying fruit at this stage of the process. Hence, this work's primary objective is to establish a database using images captured specifically on these loading ramps.

In response to the limitations of traditional methods and building upon promising prior deep learning research, this study aims to harness deep learning for oil palm fruit bunch maturity classification. Primary objectives include developing a robust deep learning model capable of accurately distinguishing between different maturity stages, utilizing images to capture dynamic changes in fruit bunches over time. To achieve these goals, images at various maturity states will be annotated, and YOLOv8 will be employed for maturity detection. This study seeks to provide technological advancement, enhancing maturity assessment accuracy, and contributing to sustainable practices in oil palm cultivation.

#### 2 RELATED WORKS

The use of video data for crop monitoring has emerged as a valuable tool in precision agriculture. Video-based approaches provide a dynamic understanding of crop growth and maturation processes over time. Successfully applied in various crops such as grapes (Kangune et al., 2019; Zhao et al., 2023) and wheat (Virlet et al., 2016), this methodology showcases its potential to capture temporal changes in oil palm fruit bunches.

Recent research has made significant strides in the maturity classification of oil palm fruit, leveraging advanced technologies. Many studies rely on noninvasive methods, predominantly visual-based, avoiding direct contact with the fruit. Some authors employ computer vision and machine learning systems, extracting color features or other image characteristics using methods like support vector machine (SVM) (Septiarini et al., 2019) and artificial neural networks (ANN). For example, Septiarini A. et al. (2021) use different machine learning algorithms as Naïve Bayes, SVM and ANN. Others utilize Raman spectroscopy, as demonstrated by Raj T. et al. (2021) employing Raman signal features as input for KNN. Considering the importance of segmentation in traditional machine learning and/or computer vision methods, some authors have focused on this aspect (Septiarini et al., 2020).

The integration of deep learning techniques into agriculture has gained ground, offering innovative solutions to various challenges, including crop monitoring, disease detection, and yield prediction. Deep learning models, particularly Convolutional Neural Networks (CNN), have shown remarkable success in image-based tasks, providing a foundation for their application in maturity classification. Recent works, including the use of convolutional neural networks capable of classifying oil palm fruit through knowledge transfer, for example, Suharjito et al., (2021), compare various CNN models, such as MobileNetV1, MobileNetV2, NASNet Mobile, and EfficientNetB0, with transfer learning (Suhariito et al., 2021). On the other hand, models such as YOLO show promising results when it comes to classifying multiple fruits in a single image with internal segmentation. Authors using the YOLO model have employed various versions, ranging from YOLOv3 (Mohd Basir Selvam et al., 2021) to YOLOv5 (Mansour et al., 2022). Some authors have even compared YOLO with other CNN models (Junior & Suharjito, 2023; Mansour et al., 2022).

However, effective classification models depend on a robust database, emphasizing the fundamental role of correct data labeling. Selecting the number of maturity grades to classify and building a high-quality database are key to model performance. Some authors have dedicated efforts to label databases for this purpose. Divergence exists in the number of maturity classes used by different authors, ranging from 2 (ripe, unripe)(Saleh & Liansitim, 2020) to 6 (unripe, under ripe, ripe, over-ripe, empty, abnormal) (Suharjito et al., 2023). Up to 7 classes have even been used, attempting to differentiate all possible options (Herman et al., 2020). But if having few classes can be detrimental, as it may not cover all maturity options or attempt to group many different types of maturity into one class, having too many can also be problematic. It is challenging to have images for all maturity styles or types of fruits because some use the abnormal class, which is for when the fruit has issues. The truth is that distinguishing different patterns would be ideal whenever we have a sufficient amount of data for each class.

Another important aspect when classifying oil palm fruits is the timing of classification. As mentioned earlier, there are two critical moments: when they are on the tree ready to be harvested and when they are in the production stage to determine how they were collected. Both stages are significant—the first for efficient harvesting. The second is crucial for quality control of the harvest. While this stage may not seem directly related to harvesting, it does ensure control over the quality of the process and allows evaluation of other suppliers a company might have. Depending on the production area, this can become quite complex; classifying fruit by fruit is impractical due to inclined surfaces in production areas, causing the fruits to be closely packed rather than completely separated.

Most articles that have developed models to predict the ripeness of oil palm fruit have done so with individual bunches or completely separated from each other, for example, on the ground (Junior & Suharjito, 2023; Mansour et al., 2022; Saleh & Liansitim, 2020; Suharjito et al., 2023). This isn't typical in a production area as it's challenging to separate cluster by cluster to classify them all. This work aims to classify the fruits in the final production stage where they pass through an inclined ramp before oil extraction.

These collective efforts highlight diverse approaches and methodologies to enhance the accuracy of oil palm fruit maturity classification, forming a basis for understanding the challenges in maturity assessment. The demonstrated potential of deep learning in recent works contributes to the context of the proposed research on oil palm fruit

bunch maturity classification. In our case, we propose three classes to avoid noise between closely related classes, considering their significance for workers in the oil palm industry. The database will be built using images of fruits from a video taken on an inclined ramp in the production area. Furthermore, given the capabilities demonstrated by previous versions of the YOLO model in other works, we will use the latest version of this model.

## 3 DATA AND METHODS

# 3.1 Methodology

The methodology followed is shown in Figure 1 below. First, data collection is performed from videos by capturing frames. The images are labeled, selecting the bunch based on their ripeness. Afterward, the data is preprocessed, meaning it is resized to the same dimensions and augmented. Following all the data preprocessing, the model training takes place, involving tasks such as splitting the data into training, testing, and validation sets, training the model over several epochs, and validating it. The data set was divided into percentages: 85% training, 10% testing, and 5% validation.

Next, the steps of this methodology are described in more detail.

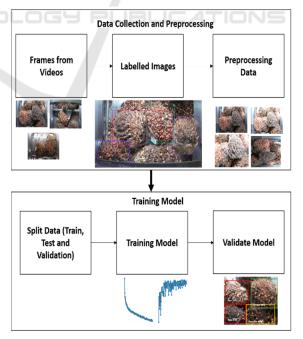


Figure 1: Methodology.

# 3.2 Data Collection and Preprocessing

data were extracted from videos and meticulously labeled by experts. To label the oil palm fruit bunches, different ripening stages were considered. Based on the available data and aiming for practicality in maturity prediction, only 3 maturity stages were utilized from the 4 classic types shown in Figure 2. In this figure, a) represents the unripe stage, crucial as the fruit isn't yet ready and might not be useful for oil production due to low oil content. The mature stage, depicted in b), is considered ideal as it allows for the extraction of the highest oil yield without losses. Stage c) indicates overripeness, leading to losses as the fruits easily detach from the bunch and might remain unused. Lastly, stage d) represents the fruit being beyond overripe, continuing its deterioration and entering a state considered rotten or spoiled. Despite its significance, due to limited available data, this latter stage was not individually considered; instead, it was merged with the previous stage. Therefore, stages c) and d) were labeled as overripe. Finally, the fruit bunches were labeled as unripe, ripe, and overripe.

Two types of palms were considered: Elaeis guineensis and hybrids OxG (E. oleifera x E. guineensis). Roboflow software facilitated image labeling, yielding 65 images from the videos (see Figure 3 for one example of the images). Within these images, fruit bunch were annotated, resulting in 390 labeled bunch: 65 unripe, 203 ripe, and 122 overripe.



Figure 2: Example of ripeness level of oil palm fruit bunch.

One of the first things to consider is the difficulty of the images. As can be seen, for example, in the figure 3, there are many fruits grouped together, which can be challenging to identify. This is precisely one of the issues with the YOLO model, identifying objects that may be overlapped. The photos were taken on a loading ramp in the production area, so the bunch may shift, and sometimes they can be on top of each other or too close. Even due to the detachment of fruits from the bunch, there may be some bunch of loose fruits that resemble bunches.



Figure 3: Example of image taken from the video.

Before training, image pre-processing was conducted. Initially, the image size was adjusted to 640x640 pixels. Various data augmentation processes were then applied, including flipping, rotating, cropping, saturating, adjusting brightness, and exposure alterations. The applied data augmentation processes are as follows:

- Flip: Horizontal, Vertical
- 90° Rotate: Clockwise, Counter-Clockwise, Upside Down
- Crop: 0% Minimum Zoom, 20% Maximum Zoom
- Rotation: Between -15° and +15°
- Shear: ±15° Horizontal, ±15° Vertical
- Saturation: Between -15% and +15%
- Brightness: Between -15% and +15%
- Exposure: Between -15% and +15%

The data augmentation process was performed during training, obtaining 165 images to train the model, significantly increasing the data per classes: 156 immature, 642 mature, and 309 overripe.

#### 3.3 YOLOv8 Model

In this study, we employed the YOLOv8 model for object detection, the latest version to date in the You Only Look Once (YOLO) series. Developed by

Ultralytics, this model is renowned for its real-time capabilities (Jocher et al., 2023).

The architecture is based on a CNN. It utilizes a simple CNN to predict bounding boxes and class probabilities in a single pass. YOLOv8 is a multiscale model, employing three scale-detection layers. This model is at the forefront of real-time object detection, providing a balance between accuracy and speed, making it a valuable tool for various applications.

Our utilization of YOLOv8 involved fine-tuning on labelled datasets, including pre-processing steps such as resizing images to 640x640 pixels.

#### 3.4 Metrics

Object detection involves not only the classification of an object but also the classification of several objects. In each case, it is necessary to evaluate whether the detection position is correct.

To assess the performance of the YOLOv8 model for object detection, various metrics are employed to evaluate accuracy and efficiency. These metrics provide information about the model's ability to accurately identify objects within an image.

# 3.4.1 Intersection over Union (IoU)

Intersection over Union (IoU) is a metric used to evaluate the overlap between the predicted bounding box and the ground truth bounding box. It is calculated by dividing the area of overlap between the two boxes by the area of their union. IoU provides a measure of how well the predicted box aligns with the actual object location.

$$IOU = \frac{Area \ of \ Intersection}{Area \ of \ Union} \tag{1}$$

#### 3.4.2 Precision and Recall

Precision and recall are fundamental metrics quantifying the model's accuracy in correctly identifying positive instances (precision) and capturing all relevant instances (recall).

# 3.4.3 Mean Average Precision (mAP)

One key metric is the mean Average Precision (AP), measuring the average accuracy across different object classes. Where (AP) is calculated as the area under the precision-recall curve. This metric is crucial for understanding the overall effectiveness of the YOLOv8 model in various scenarios.

This metric is often calculated under an IoU threshold. For example, mAP50 calculates the mean

of AP at an IoU threshold of 0.5, considering only predictions with IoU greater than or equal to 0.5. This metric is useful when flexibility in bounding box matching is needed. On the other hand, mAP50-95 is the mean of AP at different IoU thresholds from 0.5 to 0.95, calculated at 0.05 intervals. These two metrics provide information about the model's ability to accurately detect objects at different levels of overlap between predicted bounding boxes and ground truth.

#### 4 RESULTS AND DISCUSION

In this study, the YOLOv8 model was employed to train the previously described dataset. The advantage of using a model like this is the ability to detect fruit bunches in images without the need for prior segmentation. This contributes to a faster model. The model was trained using transfer learning for 300 epochs with a batch size of 16.

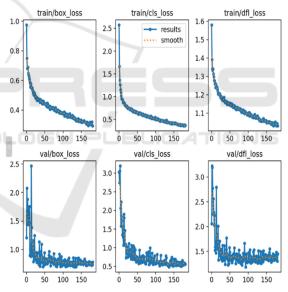


Figure 4: Training loss graphs of the model.

Figure 4 illustrates the training loss graphs. In the figure, graphs related to three losses that play a significant role in the performance of a YOLO model can be observed: loss related to bounding box regression (box\_loss), loss associated with classification accuracy (cls\_loss), and distribution focal loss (dfl\_loss). The box\_loss measures the accuracy of predicted bounding boxes around objects, indicating the alignment between predicted and actual object boundaries. Meanwhile, cls\_loss evaluates the precision of object classification, reflecting the model's ability to correctly identify object classes.

Lastly, dfl\_loss is a variant that aids in mitigating class imbalance and challenging examples during model training, thereby enhancing its capability to handle varied classes and complex instances. The first row contains three graphs related to losses, showing a clear trend of decreasing loss, indicating that more training epochs could have been performed. The subsequent three graphs are about the validation losses. The limited number of validation images results in somewhat unstable loss despite a decrease at the same level. This instability may be attributed to the scarcity of validation data.

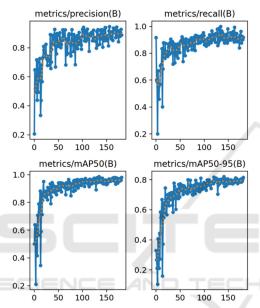


Figure 5: Training metrics graphs of the model.

Figure 5 shows the graphs that represent precision and recall, achieving values of 96.5% and 95%, respectively. Finally, mAP50 and mAP50-95 graphs, crucial for measuring model precision, reached significant values of 98% for mAP50 and 80% for mAP50-95. The validation resulted in an mAP of 94.3%.

Table 1: Average precision by class.

| Class    | Precision |
|----------|-----------|
| Unripe   | 92        |
| Ripe     | 85        |
| Overripe | 93        |

Table 1 displays the precision results for each class in the validation set. It can be observed that the intermediate class, 'mature,' has the lowest precision, while 'green' and 'overripe' exhibit the highest precision.

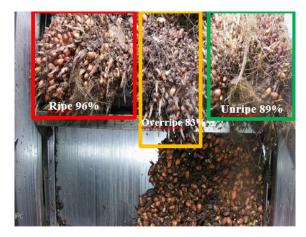


Figure 6: Image classify by the model.

As depicted in Figure 6, the model adeptly identifies fruits in all three classes. It can be seen in the figure that the model also provides a percentage of the classification of the identified object, which can also be a factor in the prediction, selecting only objects with a specific threshold. Such a model holds great utility for camera deployment in production areas, where photos are taken on a conveyor where fruits pass before being taken to tanks for oil extraction. Although this marks the final production stage, it is crucial for evaluating the overall production process. Additionally, the model could be applied in the field, using a smartphone camera to classify fruits in real-time, assisting fruit pickers in harvesting at the right moment of ripeness.

Despite achieving excellent metric results, the model shows promise for even better performance with additional training data.

Consideration could be given to adding another class for empty or rotten bunch. A detailed analysis revealed challenges, especially in classifying overripe bunch, indicating the need for a nuanced approach. Although overripe bunch tend to lose many fruits, they are not necessarily empty or semi-empty, introducing another level of complexity, but the quantity of such instances was insufficient to establish a separate class in this study.

Figure 7 shows how the model makes a mistake in classifying the bunch at the bottom right as overripe when it is actually ripe. It can be seen that this specific bunch has lost fruits, but despite that, it is not overripe. Perhaps the loss of fruits was due to the transportation process in production and not the maturity state. These can become classic errors and are challenging to detect even by some experts. However, increasing the quantity of images in the data is believed to significantly aid in improving distinctions like this.



Figure 7: Image classify by the model, with wrong identification.

## 5 CONCLUSIONS

The research presented in the article leverages the YOLOv8 model for identifying the maturity of fruit bunches through digital images, demonstrating its application in production scenarios and considering it a promising tool for fruit harvesting. Several key observations can be extracted:

- The study focuses on classifying fruits as ripe or overripe, highlighting the model's ability to discern different stages of fruit maturity, being more accurate in the unripe and ripe stages.
- Deep learning, particularly YOLO variants, proves effective in various fruit detection scenarios, capable of identifying objects in images with multiple items without the need for prior segmentation.
- The analyzed models exhibit real-time capabilities, with applications in complex orchard scenarios, contributing to timely fruit classification and harvest decisions.
- Experimental results, especially with the YOLO v8 model, emphasize its robustness in addressing variations in lighting and unstructured grape growth environments.
- In conclusion, the research underscores the versatility and effectiveness of YOLOv8 and related models in the detection, classification, and identification of the maturity of oil palm fruit bunches in diverse agricultural settings.
- For future work, it is recommended to split the overripe class into two, adding the empty or rotten bunch class to learn different patterns more effectively and increase the quantity of images.

 Another recommendation for future work is to utilize different types of photos, capturing fruit bunches not only on the loading ramp but also while on the tree. This approach would provide broader coverage for the final application's usability.

#### **ACKNOWLEDGEMENTS**

This piece of research is a part of the project which has received funds from The Royal Academy of Engineering under the Award Distinguished International Associates (DIA).

Furthermore, we want to thank Unipalma S.A.S for letting us access their facilities and for all support provided to carry out our project.

#### REFERENCES

Corley, H., & Tinker, P. (2015). *The oil palm*. Wiley Online Library. https://doi.org/10.1002/9781118953297

Fedepalmas. (2019, December). The Oil Palm Agribusiness in Colombia.

Herman, H., Susanto, A., Cenggoro, T. W., Suharjito, S., & Pardamean, B. (2020). Oil Palm Fruit Image Ripeness Classification with Computer Vision using Deep Learning and Visual Attention. In *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)* (Vol. 12, Issue 2, pp. 21–27). https://jtec.utem.edu.my/jtec/article/view/5543

Jocher, G., Chaurasia, A., & Qiu, J. (2023). *Ultralytics YOLOv8*. https://github.com/ultralytics/ultralytics

Junior, F. A., & Suharjito. (2023). Video based oil palm ripeness detection model using deep learning. *Heliyon*, 9(1), e13036. https://doi.org/10.1016/J.HELIYON. 2023.E13036

Kangune, K., Kulkarni, V., & Kosamkar, P. (2019). Grapes
 Ripeness Estimation using Convolutional Neural network and Support Vector Machine. 2019 Global Conference for Advancement in Technology (GCAT), 1–5. https://doi.org/10.1109/GCAT47503.2019.
 8978341

Lai, J. W., Ramli, H. R., Ismail, L. I., & Wan Hasan, W. Z. (2023). Oil Palm Fresh Fruit Bunch Ripeness Detection Methods: A Systematic Review. *Agriculture 2023, Vol. 13, Page 156, 13*(1), 156. https://doi.org/10.3390/AGRICULTURE13010156

Mansour, M. Y. M. A., Dambul, K. D., & Choo, K. Y. (2022). Object Detection Algorithms for Ripeness Classification of Oil Palm Fresh Fruit Bunch. *International Journal of Technology*, 13(6), 1326– 1335. https://doi.org/10.14716/IJTECH.V1316.5932

Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Using Deep Learning for Image-Based Plant Disease

- Detection. Frontiers in Plant Science, 7, 1419. https://doi.org/10.3389/fpls.2016.01419
- Mohd Basir Selvam, N. A., Ahmad, Z., & Mohtar, I. A. (2021). Real Time Ripe Palm Oil Bunch Detection using YOLO V3 Algorithm. 19th IEEE Student Conference on Research and Development: Sustainable Engineering and Technology towards Industry Revolution, SCOReD 2021, 323–328. https://doi.org/10.1109/SCORED53546.2021.9652752
- Raj, T., Hashim, F. H., Huddin, A. B., Hussain, A., Ibrahim, M. F., & Abdul, P. M. (2021). Classification of oil palm fresh fruit maturity based on carotene content from Raman spectra. *Scientific Reports* 2021 11:1, 11(1), 1–11. https://doi.org/10.1038/s41598-021-97857-5
- Saleh, A., & Liansitim, E. (2020). Palm oil classification using deep learning. *Science in Information Technology Letters*, *1*(1), 1–8. https://doi.org/10.31763/sitech.v1i1.1
- Septiarini, A., Hamdani, H., Hatta, H. R., & Anwar, K. (2020). Automatic image segmentation of oil palm fruits by applying the contour-based approach. *Scientia Horticulturae*, 261, 108939. https://doi.org/10.1016/J. SCIENTA.2019.108939
- Septiarini, A., Hamdani, H., Hatta, H. R., & Kasim, A. A. (2019). Image-based processing for ripeness classification of oil palm fruit. 2019 5th International Conference on Science in Information Technology (ICSITech), 23–26. https://doi.org/10.1109/ICSITech46713.2019.8987575
- Septiarini, A., Sunyoto, A., Hamdani, H., Kasim, A. A., Utaminingrum, F., & Hatta, H. R. (2021). Machine vision for the maturity classification of oil palm fresh fruit bunches based on color and texture features. *Scientia Horticulturae*, 286, 110245. https://doi.org/10. 1016/J.SCIENTA.2021.110245
- Suharjito, Elwirehardja, G. N., & Prayoga, J. S. (2021). Oil palm fresh fruit bunch ripeness classification on mobile devices using deep learning approaches. *Computers and Electronics in Agriculture*, *188*, 106359. https://doi.org/10.1016/J.COMPAG.2021.106359
- Suharjito, Junior, F. A., Koeswandy, Y. P., Debi, Nurhayati, P. W., Asrol, M., & Marimin. (2023). Annotated Datasets of Oil Palm Fruit Bunch Piles for Ripeness Grading Using Deep Learning. Scientific Data, 10(1), 72. https://doi.org/10.1038/s41597-023-01958-x
- Virlet, N., Sabermanesh, K., Sadeghi-Tehran, P., & Hawkesford, M. J. (2016). Field Scanalyzer: An automated robotic field phenotyping platform for detailed crop monitoring. *Functional Plant Biology*: FPB, 44(1), 143–153. https://doi.org/10.1071/FP16163
- Zhao, J., Hu, Q., Li, B., Xie, Y., Lu, H., & Xu, S. (2023).
  Research on an Improved Non-Destructive Detection
  Method for the Soluble Solids Content in Bunch-Harvested Grapes Based on Deep Learning and
  Hyperspectral Imaging. Applied Sciences, 13(11).
  https://doi.org/10.3390/app13116776.