SHM PUBLISHER
SURYA HIJAU MANFAAT

Journal of Information System Exploration and Research

https://shmpublisher.com/index.php/joiser

p-ISSN 2964-1160 | e-ISSN 2963-6361

Identifying Coconut Maturity Levels Using CNN and YOLOv8 Deep Learning Algorithms

Alfaiz Alafi Luthfie[1*], Alamsyah[2]

[1,2]Department of Computer Science, Faculty of Mathematics and Natural Science, Universitas Negeri Semarang, Indonesia

DOI: https://doi.org/10.52465/joiser.v3i2.595
Received 08 July 2025; Accepted 22 August 2025; Available online 22 August 2025

Article Info

Keywords:
Coconut Maturity;
Object Detection;
YOLOv8;
CNN

Abstract

To improve the efficiency and accuracy of determining coconut maturity levels in the processing industry, this study proposes an automated detection system employing Convolutional Neural Networks (CNN) and the You Only Look Once version 8 (YOLOv8) algorithm to classify maturity levels from image data. This study introduces an automated detection system using Convolutional Neural Networks (CNN) and the You Only Look Once version 8 (YOLOv8) algorithm to identify coconut maturity levels from image data. A dataset of 230 coconut images was utilized, classified into two categories: Young Coconut and Mature Coconut. The YOLOv8 model was trained and evaluated using standard object detection metrics including mean Average Precision (mAP), precision, recall, and F1-score. The proposed model achieved a mAP of 90.5%, precision of 99.3%, recall of 94.2%, and F1-score of 96.6%, demonstrating high accuracy in detecting coconut maturity. This approach offers a practical and efficient alternative to manual assessment, contributing to improved accuracy and operational efficiency in agricultural practices.

1. Introduction

The agricultural industry plays a crucial role in meeting the food needs of people around the world [1]. As the human population continues to grow, increasing productivity and efficiency in food production is essential [2]. One of the key challenges faced by the agricultural sector today is the demand to produce high-quality food in a sustainable and cost-effective manner. Coconuts can be processed into various products, foods, and beverages, making them one of the important agricultural commodities [3]. These products range from coconut oil, coconut water, desiccated coconut, to cosmetics and health supplements, contributing significantly to both local economies and global trade. Coconuts also offer numerous benefits in both economic and health sectors [4], [5]. They are rich in nutrients, have antioxidant properties, and provide essential fatty acids, making them highly valuable in the health and wellness industry.

[*] ***Corresponding Author:***

Alfaiz Alafi Luthfie,
Computer Science Department, Faculty of Mathematics and Natural Sciences,
Universitas Negeri Semarang,
Semarang, Indonesia.
Email: faizalafiluthfie@students.unnes.ac.id

However, the harvesting process of coconuts presents many challenges, particularly in determining the maturity level accurately and efficiently, which directly impacts the quality of the fruit [6]. Incorrect maturity estimation can result in coconuts that are either too young, lacking in flavor and oil content, or too mature, making them unsuitable for certain applications. Harvesting coconuts either too early or too late can reduce their taste, texture, and nutritional value [7]. Most harvesting is still carried out traditionally through visual assessment, such as evaluating skin color and fruit size. However, this method has many drawbacks and is prone to errors due to varying human perception and subjectivity [8]. The assessment also varies depending on environmental lighting and the experience of the harvester, further affecting consistency. Additionally, it is time-consuming, inefficient, and requires high labor costs [9]. In regions where labor shortages exist, these traditional methods become even less sustainable.

Technological advancements have had a positive impact on agriculture [1], [10]. These developments include automation, remote sensing, Internet of Things (IoT), and Artificial Intelligence (AI) [11], all of which have revolutionized the way agricultural activities are carried out. One of the outcomes of technological development is Artificial Intelligence (AI) [12] which offers many advantages, including improving the quality and productivity of crop yields [13]. AI enables machines to perform tasks such as recognition, classification [14], and decision-making, which are critical in precision agriculture. Convolutional Neural Networks (CNN), for instance, have been widely used for visual recognition tasks in agriculture and beyond due to their strong performance in image-based classification [15]. Developing an AI-based system using specific algorithms to accurately identify the maturity level of coconuts is a potential solution to this problem [16]. Such systems can help reduce dependency on manual labor, minimize errors, and ensure more consistent and timely harvesting.

One of the most capable algorithms is YOLOv8 [17]. YOLO was introduced by Joseph Redmon in 2016 through his paper titled “You Only Look Once: Unified, Real-Time Object Detection,” which is used for object detection tasks [18]. The strength of the YOLO architecture lies in its ability to perform real-time detection with high accuracy. YOLOv8 is the latest version of the YOLO model developed by Ultralytics, designed for instance segmentation, image classification, and object detection [19], [20]. It brings improvements in terms of model speed, accuracy, and flexibility, making it suitable for real-world agricultural applications. Based on the background presented, this study focuses on implementing the YOLOv8 algorithm to identify the maturity level of coconuts as a step toward modernizing and optimizing harvesting practices.

2. Literature Review

Several previous studies have explored the use of machine learning and deep learning algorithms in determining fruit maturity levels. Caladcad et al. [21] employed algorithms such as Artificial Neural Network (ANN), Random Forest (RF), and Support Vector Machine (SVM), achieving an accuracy of 80%. However, their approach required tapping tools and involved limited data, making the method less practical. In contrast, the current research adopts visual image processing without any physical contact, offering a more efficient and non-invasive alternative.

Parvathi and Tamil Selvi [22] implemented Faster R-CNN combined with ResNet-50, which achieved a higher accuracy of 89%. Despite its accuracy, the model faced challenges in detection speed and computational complexity due to the heavy architecture. The innovation in this study lies in the use of YOLOv8, a model known for its faster processing and lightweight architecture, while still maintaining high accuracy.

Further, Mandava *et al.* [8] compared several object detection models, including YOLOv5, YOLOv4-Tiny, and MobileNet SSD, and reached an average accuracy of 87%. However, they had not yet utilized the latest version of YOLO, and MobileNet was found to be less optimal in performance. This study addresses those limitations by applying YOLOv8, which provides enhanced detection capability and performance compared to YOLOv5s.

The proposed research also integrates Convolutional Neural Networks (CNN) with YOLOv8, achieving a superior accuracy of 90.5%. While the detection time remains a challenge, the novelty of this study lies in combining CNN's feature extraction strength with YOLOv8's real-time detection capabilities to accurately identify coconut maturity levels. This integration demonstrates improved performance compared to previous works and presents a significant advancement in automated coconut maturity classification.

3. Method

The research method is an essential component when conducting a study. In this research, the YOLOv8 algorithm method is used, and the workflow is presented in the flowchart shown in Figure 1. A more detailed explanation will be provided in the following sections.

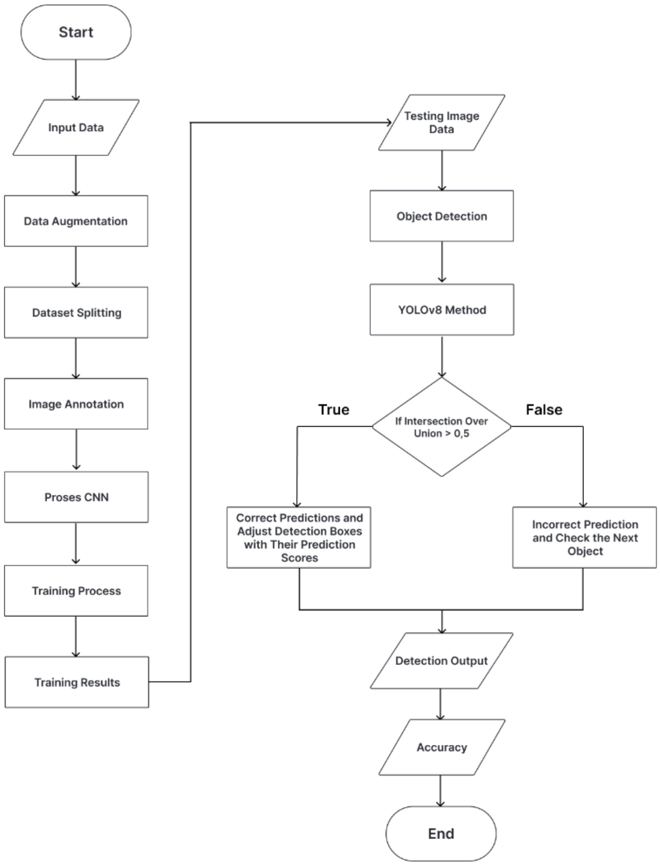


Figure 1. Research Procedure Flowchart

3.1 Input Data

The first stage involves inputting the dataset to be used for the training process. This dataset was obtained from the internet and consists of 202 images of coconuts still on the tree, representing various maturity levels, including young coconuts and mature coconuts.

3.2 Data Augmentation

The augmentation process was used to increase the quantity, quality, and variability of the training data from the dataset [23]. Augmentation techniques such as horizontal and vertical flipping, image scaling, and color transformations (saturation, hue, and brightness) were applied prior to collecting and processing the image data. The goal was to ensure effective model generalization for identifying the ripeness level of coconuts, enhance the accuracy of the model architecture, and reduce the risk of overfitting.

3.3 Dataset Splitting

The dataset was divided into three parts: test data, validation data, and training data. This division is an essential part of the machine learning training process. Its purpose is to ensure that the developed model can generalize well and perform effectively on new, unseen data. The dataset, consisting of 202 images with varying levels of coconut ripeness, was split into 19 images for testing, 39 for validation, and 144 for training.

3.4 Image Annotation

Information was added to the images, such as labels, bounding boxes, segmentation, metadata, and landmarks, according to their respective classes. The purpose of image annotation is to create a training dataset to be used in machine learning. The Roboflow web application was used to assist in the annotation process. The annotation process began by uploading the image dataset to the Roboflow web application, which simplifies and speeds up annotation. Each image is displayed in the interface, and two ripeness classes are created: Mature Coconut (label 0) and Young Coconut (label 1). A bounding box is drawn around each coconut object according to its ripeness class. Roboflow records the annotation data and saves it in .txt format. The process repeats automatically for each image until all images are annotated.

3.5 Training Process

The training process consisted of optimization, parameter tuning, data augmentation, backpropagation, and model evaluation. This process was conducted to ensure that the target object could be accurately detected. The result of the training is a weight file, which can be utilized by the system for object detection tasks.

3.6 CNN CSPNet

CSPNet (Cross Stage Partial Network) is a network architecture designed to reduce the computational complexity of deep learning while maintaining accuracy. It works by splitting the feature map into two parts: one is processed through convolutional layers, while the other bypasses them. This approach reduces gradient duplication, improves learning efficiency, and lowers computational load. In object detection architectures, CSPNet is commonly integrated into the backbone to enhance feature extraction, inference speed, detection accuracy, and resource efficiency. The main components of CSPNet include the input feature map, split layer, convolutional block, transition layer, and concatenation. The structure is illustrated in the Figure 2.

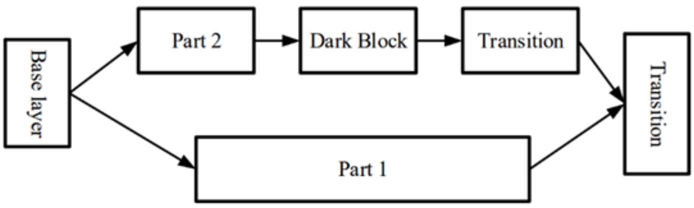


Figure 2. CSPNet Structure

3.7 YOLOv8 Method

The dataset that has undergone the image annotation process will be trained using YOLOv8. The YOLOv8 architecture consists of a backbone and a neck. In the backbone, the architecture is similar to YOLOv5, but the C3 module used in YOLOv5 is replaced with the C2f module based on CSP (Cross-Stage Partial) [24]. The C2f module is built using the C3 and ELAN (Efficient Layer Aggregation Network) modules from YOLOv7, which enhances YOLOv8's ability to capture gradient information more effectively and results in lighter training weights. At the end of the backbone, the SPPF (Spatial Pyramid Pooling Fast) module is used, which combines three 5x5 Maxpool layers to improve accuracy and reduce model weight.

In the neck, PAN-FPN (Path Aggregation Network - Feature Pyramid Network) is used to enhance feature information across different scales. It combines confidence boxes and regression outputs to improve accuracy by adding several upsampling methods and the C2f module. The YOLOv8 architecture can be seen in Figure 3.

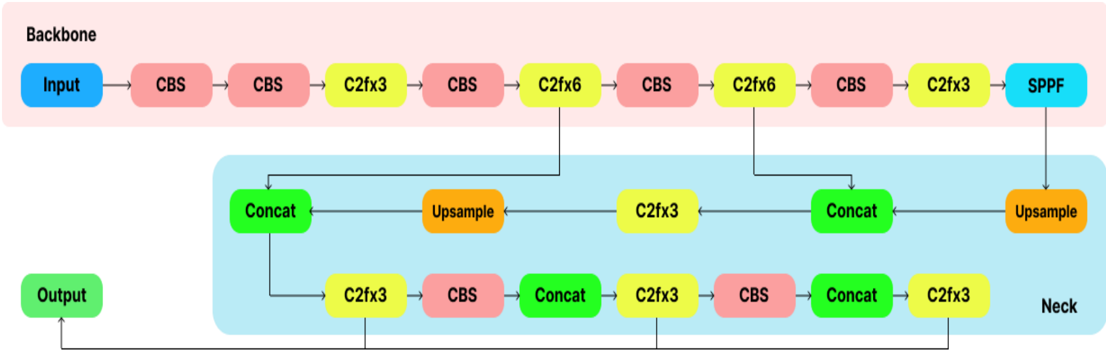


Figure 3. Yolov8 Architecture

CBS is a component of the YOLOv8 backbone that employs a cross-stage hierarchy by integrating and separating feature maps to optimize gradient flow. This design simplifies the computation process and improves accuracy, thereby enhancing the efficiency of feature extraction. The structure of CBS is shown in Figure 4.

Figure 4. CBS Structure

The SPPF module is positioned at the end of the backbone to expand the network's receptive field and enhance object detection accuracy for objects of different sizes and complexities. It employs pooling operations to extract multi-scale features, enabling the model to effectively capture contextual information [25]. The architecture of SPPF is illustrated in Figure 5.

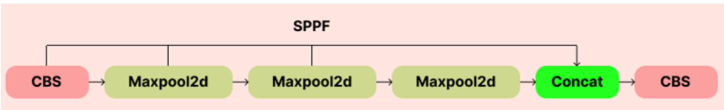


Figure 5. SPPF Structure

3.8 Accuracy Evaluation

The accuracy of YOLOv8 is determined using several key metrics, including precision, recall, mean Average Precision (mAP), and F1-score. The calculations for each of these metrics are as follows:

$$Recall = \frac{TP}{(TP+FN)} \tag{1}$$

$$Precision = \frac{TP}{(TP+FP)} \tag{2}$$

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i \tag{3}$$

$$F1 - Score = 2\ x \frac{(Precision\ x\ Recall)}{(Precision + Recall)} \tag{4}$$

The t-test is used to evaluate whether the independent variable has an individual or partial effect on the dependent variable. If the significance value is less than 0.05 and the t-count value is higher than the t-table, then if the significance value is smaller than the confidence level and the t-test value is greater, the alternative hypothesis is accepted, which indicates that the independent variable has a partial effect on the dependent variable [26].

4. Results and Discussion

Respondents reported their demographic details, including age, gender, province of origin, and education currently being pursued. ChatGPT is often used for what from a total of 52 respondents.

48.1% are male, 51.9% are female. Of the respondents, 98.1% are aged 17-22 years, while 1.9% are aged 29-34 years. Respondents mostly come from Central Java Province, with 86.5% while those from East Java, Papua, Southwest Papua, South Kalimantan, Banten, DKI Jakarta, and West Java are both 1.9%. Participants are mostly undergraduate students, with a total of 96.2% and around 1.9% are postgraduate students, then high school / vocational school / MA students, as much as 1.9%. Regarding the most frequent use of ChatGPT for what, it covers around 71.2% of respondents to search for information, and helps in conducting research and references. While for the preparation of academic assignments, as many as 23.1% and the least respondents use it to improve language skills, around 5.8%.

Results

a. Feature Visualization Using CSPNet

Each feature channel in the coconut dataset reveals specific patterns such as surface texture, skin color variations, and fruit edges. The feature visualization is displayed as 2D images with color bars representing intensity levels. Figure 6 shows various CSPNet feature channels, extracted through convolutional layers. The color bar on the right indicates the intensity values, where brighter colors represent more relevant areas, and darker tones indicate less relevant or weakly patterned regions. For example, in Feature Channel 1, values range from 1.5 to -2.0, with high positive values highlighting key areas identified by the model.

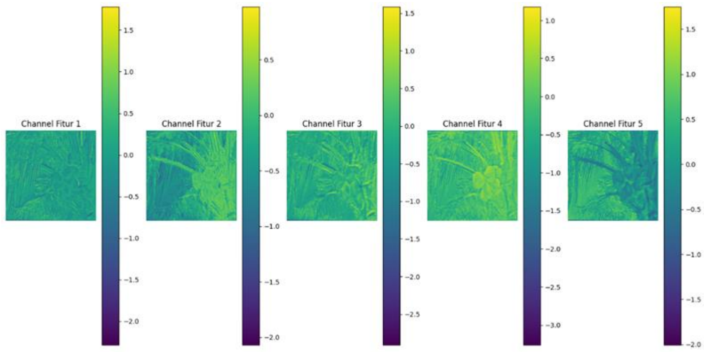


Figure 6. Feature Extraction Results

b. Test Results

The training process was conducted over 150 epochs, during which the model continuously updated its weights based on the calculated loss. The training performance is visualized in the graph shown in Figure 7, which illustrates the model's learning progress. The box_loss curve shows a consistent decline, indicating improved object detection accuracy. Similarly, the cls_loss and dfl_loss curves decrease steadily, reflecting better classification and bounding box regression. The precision, recall, and mAP metrics show increasing and high values, confirming the model’s strong performance in both detection and classification tasks.

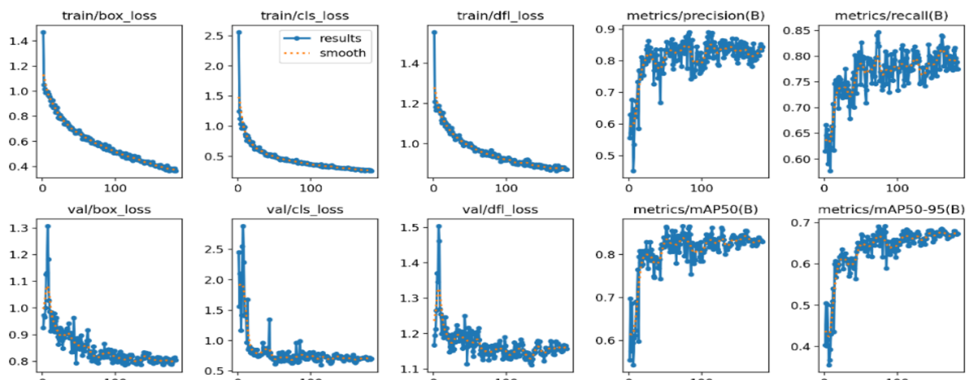


Figure 7. Training Performance Graph

The confusion matrix illustrates the performance of the model in classifying coconut ripeness into three categories: Kelapa Matang (Mature Coconut), Kelapa Muda (Young Coconut), and Background. The model correctly predicted 129 instances of Kelapa Matang and 160 instances of Kelapa Muda. However, there were some misclassifications: 19 Kelapa Matang were misclassified as Kelapa Muda, and 18 Kelapa Muda were misclassified as Kelapa Matang. Additionally, the model predicted 25 instances of Background as Kelapa Matang and 42 as Kelapa Muda, indicating some false positives related to the background class. Despite these errors, the model shows strong classification performance, particularly in distinguishing between the two ripeness classes. The confusion matrix can be seen in Figure 8.

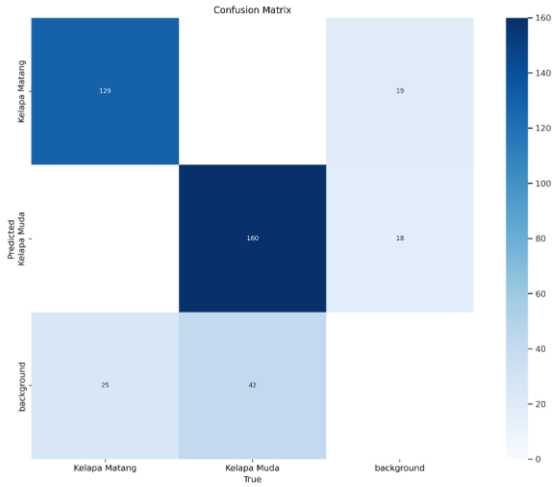
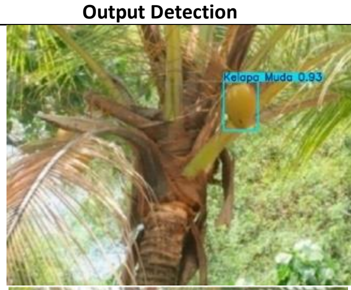
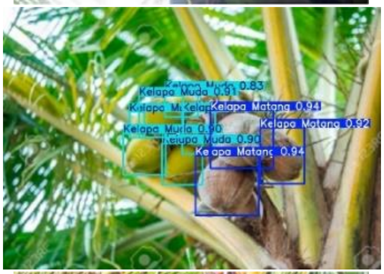
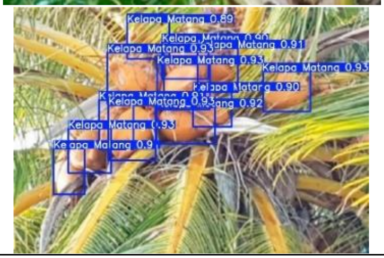


Figure 8. Confusion Matrix

The testing was conducted using the trained model files obtained from the training stage with both the CNN and YOLOv8 algorithms. The test results can be seen in Table 1.

Table 1. Test Result

Data Test	Output Detection
	
	
	

Data Test	Output Detection
	
	
	
	
	

The CNN YOLOv8 model demonstrated excellent performance in detecting the ripeness level of coconuts, achieving a mAP of 90.5%, precision of 99.3%, recall of 94.2%, and an F1-score of 96.6%. This indicates that the model is quite effective in accurately identifying coconut ripeness, with only a few prediction errors. A summary of the results can be seen in Table 2.

Table 2. Model Training Performance Evaluation Results

No	Classes	Precision	Recall	F1-Score	mAP
1	Ripe Coconut	0.991	0.959	0.975	0.904
2	Young Coconut	0.995	0.925	0.958	0.907
	Average	0.993	0.942	0.966	0.905

4.2 Discussion

In this study, the proposed method using a combination of Convolutional Neural Network (CNN) and YOLOv8 was compared with several previous research models in the field of coconut maturity classification. The comparison aims to evaluate the performance and improvements offered by the proposed model. Details of the model comparison are presented in Table 3.

Table 3. The Comparison with Previous Model

Author	Year	Model	Result
Caladcad et al [21]	2020	ANN, RF, SVM	mAP: 80%
Parvathi & Tamil Selvi [22]	2021	Faster R-CNN + ResNet-50	mAP: 89%
Mandava et al [8]	2023	YOLOv5 Mobilenet SSD	mAP: 87%
Proposed Method	**2025**	**YOLOv8 CNN**	**mAP: 90.5%**

The proposed method outperforms these previous studies by integrating CNN for feature extraction with YOLOv8, a newer and more efficient object detection algorithm. This combination achieved a mean Average Precision (mAP) of 90.5%, indicating improved accuracy and reliability in detecting coconut maturity. Therefore, the proposed model offers a more optimal solution with higher performance, and contributes to advancing automated systems in agricultural image analysis.

5. Conclusion

Coconut maturity level plays a crucial role in determining the quality and usability of coconut-based products. Manual methods that rely on visual observation are still commonly used but are prone to human error, subjectivity, and inefficiency. Therefore, an automated classification model is proposed in this study to identify coconut maturity levels based on image data. The effectiveness of the YOLOv8 model combined with CNN was successfully demonstrated in this research, as shown by high mAP, precision, recall, and F1-score values. The research results show that YOLOv8 is capable of accurately detecting the ripeness of coconuts, achieving a mAP of 90.5%, precision of 99.3%, recall of 94.2%, and an F1-score of 96.6%. The model demonstrates strong performance under various environmental conditions, although certain challenges remain in cases with low lighting and poor image quality. With its impressive speed and accuracy, YOLOv8 is well-suited for future real-time applications, such as video monitoring in plantations and quality control systems based on fruit ripeness detection.

References

[1] D. Uztürk and G. Büyüközkan, "Industry 4.0 technologies in Smart Agriculture: A review and a Technology Assessment Model proposition," *Technol Forecast Soc Change*, vol. 208, Nov. 2024, doi: 10.1016/j.techfore.2024.123640.

[2] F. Costa, S. Frecassetti, M. Rossini, and A. Portioli-Staudacher, "Industry 4.0 digital technologies enhancing sustainability: Applications and barriers from the agricultural industry in an emerging economy," *J Clean Prod*, vol. 408, Jul. 2023, doi: 10.1016/j.jclepro.2023.137208.

[3] A. Ateneo, F. M. Dayrit, and Q. Nguyen, "Improving the Value of the Coconut with Biotechnology Improving the Value of the Coconut with Biotechnology Part of the Organic Chemistry Commons, and the Other Chemistry Commons." [Online]. Available: https://archium.ateneo.edu/chemistry-faculty-pubs

[4] S. Bhoj, A. Manoj, and S. Bhaskar, "Usage potential and benefits of processed coconut shells in concrete as coarse aggregates," *Mater Today Proc*, 2023, doi: 10.1016/j.matpr.2023.03.529.

[5] F. Camargo Prado, J. De Dea Lindner, J. Inaba, V. Thomaz-Soccol, S. Kaur Brar, and C. R. Soccol, "Development and evaluation of a fermented coconut water beverage with potential health benefits," *J Funct Foods*, vol. 12, pp. 489–497, Jan. 2015, doi: 10.1016/j.jff.2014.12.020.

[6] R. P. M. Aba, M. B. Z. Luna, J. C. Villasis, and A. A. A. Ching, "Characterization of mature coconut (Cocos nucifera L.) water from different varieties," *Food and Humanity*, vol. 2, p. 100248, May 2024, doi: 10.1016/j.foohum.2024.100248.

[7] W. H. Coulibaly *et al.*, "Nutritional profile and functional properties of coconut water marketed in the streets of Abidjan (Côte d'Ivoire)," *Sci Afr*, vol. 20, Jul. 2023, doi: 10.1016/j.sciaf.2023.e01616.

[8] R. K. Mandava, H. Mittal, and N. Hemalatha, "Identifying the maturity level of coconuts using deep learning algorithms," *Mater Today Proc*, Sep. 2023, doi: 10.1016/j.matpr.2023.09.071.

[9] I. Kutyauripo, M. Rushambwa, and L. Chiwazi, "Artificial intelligence applications in the agrifood sectors," *J Agric Food Res*, vol. 11, Mar. 2023, doi: 10.1016/j.jafr.2023.100502.

[10] F. Maffezzoli, M. Ardolino, A. Bacchetti, M. Perona, and F. Renga, "Agriculture 4.0: A systematic literature review on the paradigm, technologies and benefits," *Futures*, vol. 142, Sep. 2022, doi: 10.1016/j.futures.2022.102998.

[11] A. Alamsyah, B. Prasetiyo, M. F. Al Hakim, and F. D. Pradana, “Prediction of COVID-19 Using Recurrent Neural Network Model,” *Scientific Journal of Informatics*, vol. 8, no. 1, pp. 98–103, May 2021, doi: 10.15294/sji.v8i1.30070.

[12] M. R. H. Polas, A. Afshar Jahanshahi, A. I. Kabir, A. S. M. Sohel-Uz-Zaman, A. R. Osman, and R. Karim, “Artificial Intelligence, Blockchain Technology, and Risk-Taking Behavior in the 4.0IR Metaverse Era: Evidence from Bangladesh-Based SMEs,” *Journal of Open Innovation: Technology, Market, and Complexity*, vol. 8, no. 3, Sep. 2022, doi: 10.3390/joitmc8030168.

[13] V. Sachithra and L. D. C. S. Subhashini, “How artificial intelligence uses to achieve the agriculture sustainability: Systematic review,” *Artificial Intelligence in Agriculture*, vol. 8, pp. 46–59, Jun. 2023, doi: 10.1016/j.aiia.2023.04.002.

[14] A. Alamsyah and D. A. Anggraeni, “Detection of Indonesian Sign Language System using Convolutional Neural Network (CNN) with Nadam Optimizer,” 2024, pp. 352–359. doi: 10.2991/978-94-6463-589-8_32.

[15] A. Alamsyah and A. Izzulhaq, “Implementation of the Convolutional Neural Network Algorithm Using the ResNet50V2 Architecture for Pneumonia Detection,” 2024. [Online]. Available: https://journal.unnes.ac.id/journals/JM/index

[16] J. Chen *et al.*, “Detecting ripe fruits under natural occlusion and illumination conditions,” *Comput Electron Agric*, vol. 190, Nov. 2021, doi: 10.1016/j.compag.2021.106450.

[17] Y. Zhang, “YOLO Series Target Detection Technology and Application,” 2023.

[18] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” 2016. [Online]. Available: http://pjreddie.com/yolo/

[19] R. Sapkota, D. Ahmed, and M. Karkee, “Comparing YOLOv8 and Mask R-CNN for instance segmentation in complex orchard environments,” *Artificial Intelligence in Agriculture*, vol. 13, pp. 84–99, Sep. 2024, doi: 10.1016/j.aiia.2024.07.001.

[20] M. Sohan, T. Sai Ram, and Ch. V. Rami Reddy, “A Review on YOLOv8 and Its Advancements,” 2024, pp. 529–545. doi: 10.1007/978-981-99-7962-2_39.

[21] J. A. Caladcad *et al.*, “Determining Philippine coconut maturity level using machine learning algorithms based on acoustic signal,” *Comput Electron Agric*, vol. 172, May 2020, doi: 10.1016/j.compag.2020.105327.

[22] S. Parvathi and S. Tamil Selvi, “Detection of maturity stages of coconuts in complex background using Faster R-CNN model,” *Biosyst Eng*, vol. 202, pp. 119–132, Feb. 2021, doi: 10.1016/j.biosystemseng.2020.12.002.

[23] G. Lin, J. Z. Jiang, J. Bai, Y. W. Su, Z. H. Su, and H. S. Liu, “Frontiers and developments of data augmentation for image: From unlearnable to learnable,” Feb. 01, 2025, *Elsevier B.V.* doi: 10.1016/j.inffus.2024.102660.

[24] E. Casas, L. Ramos, C. Romero, and F. Rivas-Echeverría, “A comparative study of YOLOv5 and YOLOv8 for corrosion segmentation tasks in metal surfaces,” *Array*, vol. 22, Jul. 2024, doi: 10.1016/j.array.2024.100351.

[25] H. Lou *et al.*, “DC-YOLOv8: Small-Size Object Detection Algorithm Based on Camera Sensor,” *Electronics (Switzerland)*, vol. 12, no. 10, May 2023, doi: 10.3390/electronics12102323.

[26] A. Melda, Y. E. Wahyuningsih, and S. R. Sani, “The Effect of Extreme Economic Assistance on Community Income in Kaway XVI District, West Aceh Regency (Case Study of 6 Villages in Kaway XVI District),” *Jurnal Akuntansi, Manajemen dan Ilmu Ekonomi (Jasmien)*, vol. 4, no. 04, pp. 145–155, 2024.