Implementation of Convolutional Neural Network Algorithms to Detect the Ripeness of Palm Fruits based on Image Colors

Aditya Guna Dharma[1*], Iwan Purnama[2], Budianto Bangun[3]

[1,2,3] Faculty of Science and Technology, Universitas Labuhanbatu, Sumatera Utara Indonesia.
[*]Corresponding Author:
Email: adityagunadharma59@gmail.com

Abstract.

The ripeness of oil palm fruit greatly affects the quality of the oil produced. Usually, humans manually detect the maturity of oil palm fruit, a process that can be inconsistent and time-consuming. Therefore, there is a need for automated methods that can improve accuracy and efficiency in determining the maturity of oil palm fruit. The study proposes using the Convolutional Neural Network algorithm to detect oil palm fruit maturity based on digital color images. We expect a convolutional neural network, known for its effectiveness in image analysis, to deliver accurate results in classifying fruit maturity. The aim of this study is to develop and test a model of a convolutional neural network that can classify oil palm fruit into three maturity categories: raw, ripe, and rotten. This study employs a maturity-categorized data set of oil palm images. We built and trained a convolutional neural network model using this data set. We evaluate the model performance using four main metrics: accuracy, precision, recall, and f1-score. We perform an in-depth analysis to assess model performance in each maturity category. The confusion matrix yielded an accuracy of 74.6%. The convolutional neural network model developed showed the highest precision results for the ripe fruit (86.00%), followed by rotten fruit (72.30%), and raw fruit (67.30%). We obtained the highest recalls (85.40%) for the raw fruit category, followed by ripe fruits (75.50%) and rotten fruits (72.30%.

Keywords: *CNN, Confusion Matrix, Deep Learning, Palm Fruit, Ripeness.*

1. INTRODUCTION

Palm oil provides employment, income, and foreign exchange to the country and is one of the plantation crops that plays an important role in the national economy. Palm oil is currently developing very quickly. As the number of people's needs increases, so does the production and area. In 2022, Indonesia will have 15.34 million hectares of oil palm plantations. Oil palm plantations also provide employment and income for around 5.30 million farming families, as well as contributing foreign exchange of 20.2 billion dollars [1]. Oil palm plants produce palm oil and palm kernel, serving as the primary raw materials for the production of crude palm oil (CPO) in Indonesia. CPO possesses immense potential as a raw material for diverse oil products, catering to both food and non-food needs. Indonesia's palm oil production has increased every year since 1980. Since then, this amount has continued to increase by an average of 11.48% per year [2].

Many variables influence the quality of palm oil; these include water content, dirt content, free fatty acid content, and, most importantly, the maturity level of the palm fruit [3]. However, determining the maturity level of palm fruit in palm oil mills presents numerous challenges, including the need for sorting to produce high-quality palm fruit at the correct maturity level. Palm oil processing companies and farmers manually select palm fruit. The identification process has several drawbacks, including the lengthy time required, the tendency for humans to become worn out and bored with monotonous tasks, the perception of fruit as unique, and the production of varying product yields due to limited visual abilities. The observer's mental state also greatly influences this process. This can also cause inconsistencies during the selection process. Due to its lengthy manual execution, this process necessitates the use of machines on a large industrial scale. Efficient palm oil production is highly dependent on the maturity of the palm fruit at harvest. Improper fruit ripening can result in low productivity and quality of the oil produced [4]. Therefore, we need a technology that can identify the maturity level of oil palm fruit automatically and accurately. This system must be able to identify the maturity level of oil palm fruit in three categories: unripe, ripe, and rotten. Producers can then process the mature palm oil fruit into crude oil for consumption or daily needs. We should be able to use images of oil palm fruit automatically to differentiate between these categories. We expect this system to automate the fruit ripeness inspection process, increase efficiency, and produce more objective measurements than manual methods [5].

Previous studies have utilized computer vision and machine learning technology to detect fruit ripeness in various types of agricultural commodities. Previous studies have used image processing techniques to detect the ripeness of bananas [6] and apples [7]. In addition, researchers have applied machine learning-based classification algorithms like Decision Tree, Random Forest, and Support Vector Machine to identify the ripeness level of citrus and tomato fruits [8], [9]. Meanwhile, several previous studies have focused on detecting the ripeness of oil palm fruit using computer vision and machine learning techniques [4]. Researchers have developed methods such as classifying maturity levels using the blue-red fluorescent ratio, detecting maturity levels using active multi-band optical sensors, and applying machine learning-based segmentation and classification methods. The main challenge, however, is the variation in visual characteristics of ripe or immature fruit. In other research, applying the K-Means Clustering method based on RGB and HSV colors showed that the model was able to distinguish between unripe, quite ripe, and ripe palm fruit with an accuracy level of 64.58% [10]. One potential weakness is the mismatch between accuracy for test data (79.16%) and training data (50%). This difference suggests that the model may overfit to the training data. Another study employed an optical probe and found a correlation between the ripeness level of oil palm fruit and its hardness [11]. A weakness of this study is that it does not mention external factors that could influence the measurements, such as environmental light conditions or fruit surface variations (e.g., dirt, humidity), which could potentially influence photodiode readings.

Researchers have proven the effectiveness of the Convolutional Neural Network (CNN) algorithm in classifying digital images, which includes recognizing the type and quality of fruit [12]. The CNN algorithm has proven to be effective in automatically extracting visual features and classifying fruit ripeness with high accuracy [7]. We can apply this approach to identify the maturity level of oil palm fruit based on digital color images. Because the CNN method attempts to imitate the image recognition system in the human visual cortex, it has the ability to process images in the same way as humans [13]. Convolutional neural networks are able to automatically extract relevant features from input images and perform classification with high accuracy [14]. CNN is a powerful deep learning architecture that has demonstrated extraordinary effectiveness in digital image processing applications, including object detection and classification [15]. CNN models can automatically extract relevant visual features from input images and then use a series of convolution and pooling layers to build a powerful feature representation. This capability makes CNNs particularly suitable for tasks such as automatic detection and classification of maturity levels of oil palm fruit, which often show subtle visual variations between ripe and immature samples.

This research aims to apply the CNN algorithm to detect the maturity of oil palm fruit based on color images. We need a fruit image data set with valid maturity labels to apply CNN to the problem of oil palm fruit maturity detection. Following the image preprocessing stage, we can train the CNN model to understand the relationship between the fruit's visual features and its ripeness level. We evaluate the model by measuring metrics like classification accuracy, precision, recall, and f1-score on independent test data sets. Automating the recognition of palm fruit maturity stages will help the palm oil industry optimize production processes and improve the quality of the oil produced. Using the CNN model's powerful feature extraction and classification tools, this study makes it possible to evaluate fruit fertility more accurately and objectively than with manual methods. This leads to higher productivity and sustainability in the palm oil supply chain.

II. METHODS

This research uses the CNN algorithm, a well-known deep learning technique capable of accurately classifying image data, to identify images of palm fruit. The CNN algorithm undergoes processing with the assistance of the Python and Tensorflow programming packages in the Google Colab text editor. It works by recognizing objects or images as input, and the output is the level of object recognition accuracy. Figure 1 shows the stages of the research carried out.

Fig. 1. Research Stages

Data collection was the first step. The dataset consists of images of palm fruit from three categories: unripe, ripe, and rotten palm. We collected the dataset from various internet sources, in addition to directly taking signature photos with a cellphone camera based on the specified class. The dataset comprises 1530 images, with 80% serving as training data and the remaining 20% as testing data. We will use the training data to build a model, while we will use the testing data to evaluate the newly created model. The second stage involves resizing each image of oil palm fruit to 128x128 pixels with a channel size of 3, an input shape of 128x128x3, a batch size of 16, and an epoch of 50, which will subsequently serve as a dataset. The next step involves augmenting the image before it enters the network. This design employs two convolution layers for the number of convolution processes involved. Each convolution has a different number of filters and kernel size. The dropout process begins with the flattening process, which transforms the feature map from the pooling layer into vector form. We usually refer to this process as the fully connected layer stage. CNN carried out testing to evaluate the model it produced. We conducted the model testing at the training stage. This study used 307 images as the test data. After obtaining the test results, the next process is to evaluate the model by measuring metrics such as classification accuracy, precision, recall, and f1-score.

III. RESULT AND DISCUSSION

At this stage, the first step involves data collection. We carry out the data collection process using two methods: downloading photos from the internet based on predetermined requirements and directly taking photos using a mobile phone camera. Once we collect the data, we use the Photoshop application to process the image, removing the background and changing it to white. This is done to increase the program's ability to read image pixels because if the image has a default background whose color changes, it will affect the program's capacity to read image pixels. Figure 2 shows the images before and after background removal.

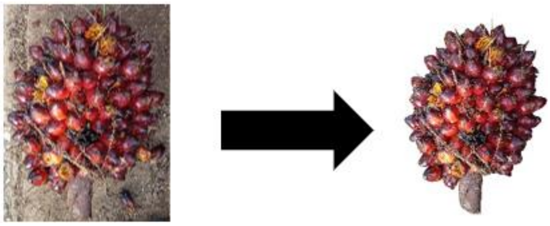

Fig. 2. Delete Image Background

The train_test_split technique from the Sklearn module, which divides the data into 80% and 20%, results in the division of the amount of data used in this research into two parts: training data, which totals 1223 images, and testing data, which totals 307 images.

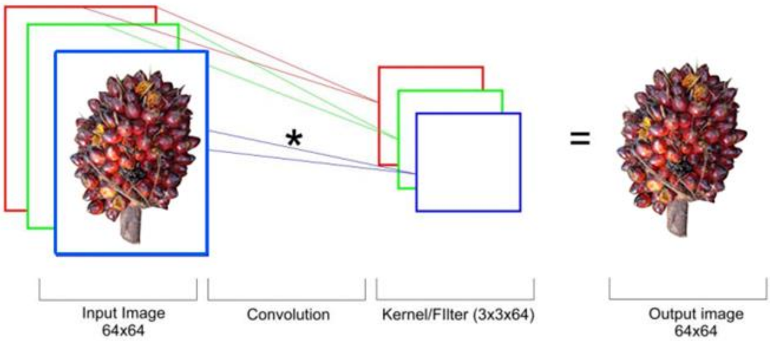

Fig. 3. Convolution Process

Figure 3 shows a diagram of the convolution process in CNN. The convolution process is a basic step in the CNN architecture, which aims to extract features from the input image. This process involves applying filters or kernels to the image to produce a feature map. This feature map is a more abstract representation of an image that contains important information that is useful for classification. The input image is a picture of oil palm fruit with a resolution of 64x64 pixels. This image will go through a convolution process to extract important features. The convolution process involves applying filters or kernels to the input image. This process employs a 3x3 kernel. This kernel moves across the input image, performing a dot product operation between the kernel and the part of the image it passes through. This operation places a value at the appropriate location on the feature map. We apply the kernel or filter, a small matrix (in this case 3x3), to the input image. A CNN employs numerous kernels (64 in this figure), each of which extracts distinct features from the input image, including edges, textures, and color patterns. The output image is the result of the convolution process, and it has the same dimensions as the input image (64x64), but with extracted features. Often referred to as a feature map, this output image holds crucial information for subsequent classification. By using multiple kernels, CNNs can extract different types of features from the same image. The convolution process is a critical step in CNN architecture to extract features from input images. By using multiple kernels, CNNs can capture various types of features that are useful for classification. This research uses the convolution process to detect the maturity of oil palm fruit based on digital color images. The fully connected layer then uses the extracted features from the input image to make a final prediction about the fruit maturity class (unripe, ripe, rotten). This process allows the CNN model to effectively classify images based on the visual features extracted from the input image.

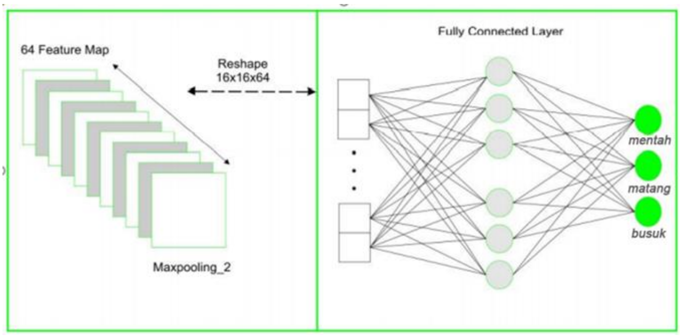

Fig. 4. Fully Connected Process

Figure 4 shows the fully connected process, which is the final part of the CNN architecture. The fully connected layer converts the image's features into vectors after several layers of convolution and pooling to carry out the final classification. 64 A feature map is the result of several previous convolution and pooling layers. Each feature map represents the results of feature extraction from the input image. The MaxPooling process reveals the feature map's specific dimensions in this image. The MaxPooling technique reduces the dimensions of feature maps, thereby reducing the number of parameters and computations in the network. This process helps filter important features by taking the maximum value from each pooling region. The MaxPooling process transforms the feature map into a long vector, preparing it for insertion into the fully connected layer. The dimensions of 16x16x64 in this case indicate that the feature map undergoes reshaping into a 1D vector prior to its entry into the fully connected layer. Neurons in the fully connected layer fully connect to all neurons in the previous layer. Each neuron in the fully connected layer receives input from all the neurons in the previous layer, allowing the network to combine the extracted features and make classification decisions. The output classes are the last neurons in the fully connected layer, representing the different classes of palm fruit maturity: unripe, ripe, and rotten. Each output neuron provides a score or probability that the input image falls into one of these classes. The class with the highest score is the model's final prediction.

For the final classification, the fully connected process in the CNN architecture is critical. The fully connected layer consolidates and processes the extracted image features to make predictions after undergoing several layers of convolution and pooling. Understanding this process allows us to see how the CNN model converts input images into classification decisions that are useful in detecting palm fruit ripeness. This CNN model uses visual features from digital color images to differentiate between raw, ripe, and rotten palm oil fruit with fairly high accuracy, as shown in the previous confusion matrix. This fully connected process is a key step in achieving these predictions.

Table 1. Confusion Matrix

Matrix		**Class Prediction**		
		Raw	Ripe	Rotten
Real Class	Raw	76	8	17
	Ripe	15	80	11
	Rotten	22	5	73

Table 1 displays a confusion matrix derived from testing the palm oil image dataset using 307 test data points. We use a confusion matrix as a tool to evaluate the performance of classification models. This matrix shows the number of correct and incorrect predictions made by the model compared to the true value of the test data. This matrix helps to understand how well the model classifies each class. Real class indicates the original class of the test data. Class prediction shows the class predicted by the model. According to the confusion matrix, the results for the raw class were: 76 correct predictions as raw, 8 incorrect predictions as mature, and 17 incorrect predictions as rotten. For the Ripe class, there are 15 incorrect predictions as raw, 80 correct predictions as ripe, and 11 incorrect predictions as rotten. In contrast, the Rotten class yielded 22 incorrect raw predictions, 5 incorrect mature predictions, and 73 correct rotten predictions. The resulting True Positives (TP) are Raw: 76, Ripe: 80, and Rotten: 73. The resulting false positives (FP), Raw: 15 (from the Ripe class) + 22 (from the Rotten class) = 37, Ripe: 8 (from the Raw class) + 5 (from the Rotten class) = 13, and Rotten: 17 (from the Raw class) + 11 (from the Ripe class) = 28. The resulting false negatives (FN) were Raw: 8 (from Ripe class) + 5 (from Rotten class) = 13, Ripe: 15 (from Raw class) + 11 (from Rotten class) = 26, and Rotten**: 17 (from Raw class) + 11 (from Ripe class) = 28. The confusion matrix yielded an accuracy value of 74.6%.

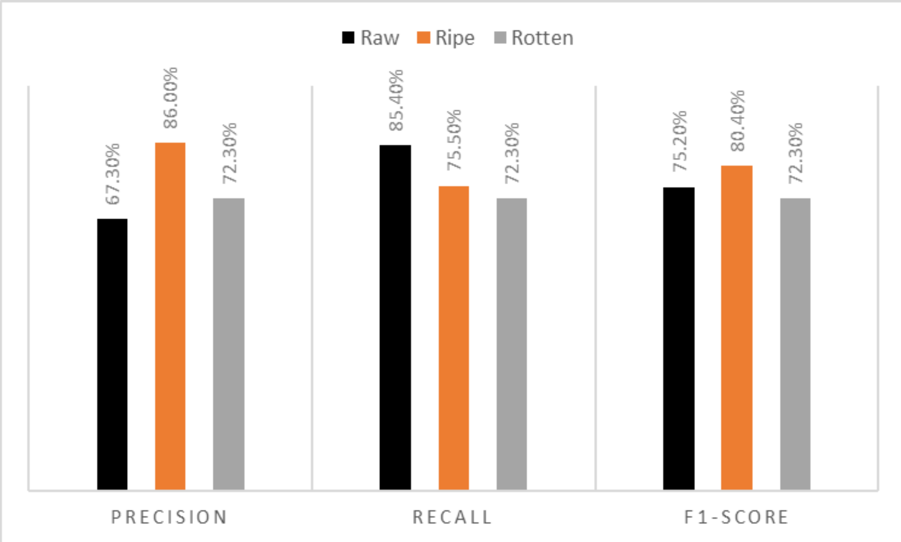

Fig. 5. Precision, Recall, and F1-score of the CNN Model

Figure 5 displays the results of measuring the CNN model's performance in detecting oil palm fruit maturity based on digital color images. The graph presents three main evaluation metrics: precision, recall, and F1-score, for three categories of palm fruit maturity, namely raw, ripe, and rotten. For the raw fruit category, the precision value is 67.30%. This shows that of all the unripe fruit predictions made by the model, 67.30% of them were actually unripe. The ripe fruit category's precision reached 86.00%, the highest value among the three categories. This means that the model is excellent at predicting ripe fruit with a high degree of accuracy. For rotten fruit, precision was 72.30%, showing quite good performance in identifying rotten fruit. The model correctly identified 85.40% of the actual unripe fruit. The model correctly detected 75.50% of the actual ripe fruit. The recall for rotten fruit is 72.30%, which indicates that the model can identify 72.30% of all actual rotten fruit. Unripe fruit has an F1 score of 75.20%. The F1-score is the harmonic mean of precision and recall, providing an idea of the balance between the two. This value indicates excellent overall performance for the raw fruit category. The F1-score for ripe fruit is 80.40%, the highest value among the three categories. This shows that the model has excellent performance in detecting ripe fruit. The F1-score for rotten fruit is 72.30%, indicating that performance is consistent with the precision and recall values previously described.

The ripe fruit category has the highest precision and F1-score, indicating that the model is very effective in detecting ripe fruit with high accuracy and a satisfactory balance between precision and recall. The recall for ripe fruit was marginally lower than the precision, suggesting that the model may have overlooked some ripe fruit. The raw fruit category has a very high recall (85.40%) compared to precision (67.30%). This shows that the model tends to identify fruit as unripe more often, but with a higher error rate. The lower F1-score compared to ripe fruit indicates that although recall is high, low precision reduces the overall performance for this category. The rotten fruit category has a balanced precision, recall, and F1-score of around 72.30%. This shows that the model has a fairly stable performance in detecting rotten fruit, although not as good as ripe fruit.

IV. CONCLUSION

According to the analysis above, the CNN model used in this research performs quite well, with an accuracy of around 74.6%. However, several classes, particularly the Ripe and Rotten classes, exhibit prediction results that require improvement due to their high false positives and false negatives. Precision and Recall also show different performance variations within each class, indicating that the model is better at recognizing some classes than others. By understanding this confusion matrix, researchers can evaluate and improve the model further, such as by adjusting parameters or using data enhancement techniques to obtain more accurate results. Overall, the CNN model applied in this study showed excellent performance in detecting the maturity of oil palm fruit, especially for the ripe fruit category. However, there is room for improvement in detecting unripe and rotten fruit, particularly in increasing precision for unripe fruit and balanced performance for rotten fruit. Further research could focus on optimizing the model to improve precision and recall simultaneously, especially for categories with lower performance.

REFERENCES

[1] B. Indonesia, "**Statistik Kelapa Sawit Indonesia 2022**," 2023.

[2] Ditjenbun, "**Kelapa Sawit (Palm Oil)**," 2019.

[3] I. U. P. Rangkuti and A. Syahputra, "Warna Minyak Sawit Mentah dan Stabilitas Warna Berdasarkan Tingkat Kematangan Buah yang Berasal dari Kebun dengan Ketinggian 800 mdpl," **Agro Fabr**., vol. 1, no. 2, 2019, doi: 10.47199/jaf.v1i2.136.

[4] A. Septiarini, H. R. Hatta, H. Hamdani, A. Oktavia, A. A. Kasim, and S. Suyanto, "Maturity Grading of Oil Palm Fresh Fruit Bunches Based on a Machine Learning Approach," in **2020 Fifth International Conference on Informatics and Computing (ICIC)**, 2020, pp. 1–4. doi: 10.1109/ICIC50835.2020.9288603.

[5] H. F. Muttaqin and A. P. Wahyu, "Innovation Computer Vision Technology With Internet Of Things (Iot) For Support Entrepreneurs In Fishery Sector," **JUMANJI (Jurnal Masy. Inform. Unjani)**, vol. 2, no. 2, p. 125, 2019, doi: 10.26874/jumanji.v2i2.39.

[6] D. Surya Prabha and J. Satheesh Kumar, "Assessment of banana fruit maturity by image processing technique," **J. Food Sci. Technol**., vol. 52, no. 3, pp. 1316–1327, 2013, doi: 10.1007/s13197-013-1188-3.

[7] P. Valdez, "Apple Defect Detection Using Deep Learning Based Object Detection For Better Post Harvest Handling," **ArXiv**, vol. abs/2005.06089, 2020, [Online]. Available: https://api.semanticscholar.org/CorpusID:218614050

[8] Seema, A. Kumar, and G. S. Gill, "Automatic Fruit Grading and Classification System Using Computer Vision: A Review," in **2015 Second International Conference on Advances in Computing and Communication Engineering**, 2015, pp. 598–603. doi: 10.1109/ICACCE.2015.15.

[9] S. Cubero *et al.*, "Optimised computer vision system for automatic pre-grading of citrus

fruit in the field using a mobile platform," **Precis. Agric**., vol. 15, no. 1, pp. 80–94, 2014, doi: 10.1007/s11119-013-9324-7.

[10] E. F. Himmah, M. Widyaningsih, and M. Maysaroh, "Identifikasi Kematangan Buah Kelapa Sawit Berdasarkan Warna RGB Dan HSV Menggunakan Metode K-Means Clustering," **J. Sains dan Inform**., vol. 6, no. 2, pp. 193–202, 2020, doi: 10.34128/jsi.v6i2.242.

[11] N. Sari, M. Shiddiq, R. H. Fitra, and N. Z. Yasmin, "Ripeness Classification of Oil Palm Fresh Fruit Bunch Using an Optical Probe," **J. Aceh Phys. Soc**., vol. 8, no. 3, pp. 72–77, 2019, doi: 10.24815/jacps.v8i3.14122.

[12] M. S. Morshed, S. Ahmed, T. Ahmed, M. U. Islam, and A. B. M. A. Rahman, "Fruit Quality Assessment with Densely Connected Convolutional Neural Network," in **2022 12th International Conference on Electrical and Computer Engineering (ICECE)**, 2022, pp. 1–4. doi: 10.1109/ICECE57408.2022.10088873.

[13] F. F. Maulana and N. Rochmawati, "Klasifikasi Citra Buah Menggunakan Convolutional Neural Network," **J. Informatics Comput. Sci**., vol. 01, no. 02, pp. 104–108, 2019, doi: 10.26740/jinacs.v1n02.p104-108.

[14] L. Shen *et al.*, "Fusing attention mechanism with Mask R-CNN for instance segmentation of grape cluster in the field," **Front. Plant Sci**., vol. 13, 2022, doi: 10.3389/fpls.2022.934450.

[15] A. Kamilaris and F. X. Prenafeta-Boldú, "A review of the use of convolutional neural networks in agriculture," **J. Agric. Sci**., vol. 156, no. 3, pp. 312–322, 2018, doi: DOI: 10.1017/S0021859618000436.