Proceeding International Conference on Religion, Science and Education (2024) **3**, 645-653

Identification of Diseases on Corn Leaves Using CNN Denoising (DeCNN)

Dzaki Nashrullah Suharto[1,*], Rila Mandala[2]
[1]Informatics Engineering Department; [2]Electrical Engineering and Informatics Department, Faculty of Computing, President University, Jl. Ki Hajar Dewantara Bekasi 17550, Indonesia. Tel. 021-89109763.

Corresponding author*
dzakinash@gmail.com

Abstract: Indonesia is a country with large agricultural commodities, especially corn. The need for corn increases every year both for food and feed. In agricultural production, if it cannot be identified correctly and quickly it will have a negative impact on crop yields, especially on the amount of harvest and crop quality. Currently every country is focusing on agricultural automation to increase crop yields and crop quality and accuracy to meet market demand. To respond to this, this research was conducted to reduce failed yields caused by diseases that attack corn. The aim of this study is to improve the accuracy of the Convolutional Neural Network algorithm in identifying diseases on corn leaves by combining Gaussian filters. This research was conducted with 4 classes on 4000 images, this research will produce data based on the class of healthy leaves, gray diseased leaves, blight diseased leaves. and leaf rust disease. the accuracy results obtained by combining the gaussian filter in this study are 99%, and it is hoped that further development in this research will be in the form of a mobile application that can help farmers.

Keywords: Identification of Leaves, Corn, Gaussian Filter, Cnn, Image Processing.

Introduction

Indonesia is one of the tropical countries in the world that is crossed by the equator, thus making Indonesia a country that has great potential for agricultural commodities, one of which is corn. The need for corn every year increases both for food and feed. In 2000, corn contributed IDR 9.4 trillion to Indonesia's income and increased drastically in 2003 to IDR 18.2 trillion. Based on data owned by the ministry of agriculture, in January - December 2021 the national corn harvest area is 15.79 million hectares. Meanwhile, the national demand for corn for feed, industry and consumption for a year is 14.37 million hectares (Kementerian pertanian republik indonesia, 2021). With so many national needs that have an impact on national corn availability (Kompas, 2008). In agricultural production, if you can't identify it properly and quickly it will have a negative impact on crop yields, especially on the amount and quality of the crop. Currently every country is focusing on agricultural automation to increase crop yields and crop quality and accuracy to meet market demand (Chhillar et al. 2020). This research was conducted based on the previous background. This research was conducted to be able to assist in increasing crop yields. Previous studies have estimated that there are around 220,000 to 420,000 plant species in the world (Scotland et al. 2003). Plants are one of the commodities needed at this time, both for primary needs and industrial needs. As well as the spread of disease in plants contributing to crop yields which will have an impact on the fulfillment of needs (Thangaiyan et al. 2019). In some countries, corn is the main food source. Food security is affected by climate change and plant diseases (Tai et al. 2014). In developed countries, agricultural modernization has received more attention in the development of a combination of technology and agricultural science, this can be seen in differences with developing countries. In developing countries, plant diseases can be disastrous, as many livelihoods depend on crop

ISSN 2829-3355 | EISSN 2828-8467

yields (Harvey et al. 2014). Based on previous research, this study proposes to use Denoising combined with the CNN algorithm to improve the performance of the CNN algorithm for predicting disease in maize.

Related Research and Methods

Related Research

This research is bases on previous important studies which can be seen in table 1.

Table 1. Related Work.

Title	Author	Year	Topic	Method
A CNN Approach for Corn Leaves Disease Detection to Support Digital Agricultural System	Kshyanaprava Panda Panigrahi, Abhaya Kumar Sahoo, and Himansu Das	2020	The corn leaf disease detection system uses CNN to support digital agricultural systems	CNN
Identificatio n of Diseases in Corn Leaves using Convolution al Neural Networks and Boosting	Prakruti Bhatt, Sanat Sarangi, Anshul Shivhare, Dineshkumar Singh and Srinivasu Pappula	2019	Disease identification system for corn leaves using CNN + Adaboost by augmenting training data to improve accuracy.	CNN + AdaBoost
Identification of Plant-Leaf Diseases Using CNN and Transfer-Learning Approach	SK Mahmudul Hassan, Arnab Kumar Maji, Michal Jasinski	2021	Development of a system to identify diseases in plant leaves using CNN with various architectures.	CNN
Dev of Efficient CNN model for Tomato crop disease identification	Mohit Agrawal, Sunnet Kr, Gupta, K.K Biswas	2020	Development of an efficient CNN model for identifying diseases in tomato plants	CNN

Methods

A. Dataset

The *dataset* used in this study is real data taken on corn leaves, namely *PlantDoc*. The total images in this dataset are 2,598 color images (RGB) consisting of 13 plant species and more than 17 types of plant diseases. In this dataset, the author only takes corn leaf data. The data used amounted to 4000 image data (1000 image data of healthy corn leaves, 1000 image data of corn leaf affected by leaf rust, 1000 image data of corn leaf affected by leaf spot and 1000 data image of corn leaf affected by leaf blight). The dataset used will be divided into test data and training data, with the presentation of the division being 80% training data and 20% test data.

1. Common Rust

Leaf rust has a characteristic, namely the presence of small lessio on the leaves which then circle lengthwise. Subsequent symptoms are reddish-brown discoloration on the top and bottom of the leaves, which then turn blackish brown. If the infection is not controlled, the leaves will dry out, resulting in the death of the plant.

Figure 1. *Corn Leaf affected common rust*

2. Gray Spot

Gray leaf spot disease (Gray spot) is a disease caused by the fungus *Cercospora zeaemaydis* which attacks corn leaves. This fungus can be found for a long time in the soil due to plant residues. Unlike leaf blight which develops between 20°C-30°C, this disease has a favorable temperature between 25°C-30°C. Symptoms that appear caused by the fungus *Cercospora zeaemaydis* are the appearance of brown or brown dots and a yellow circle of *chlorosis*. These lesions turn gray and infect younger leaves. This disease grows to be long and sometimes quadrangular parallel to the leaf veins, if this disease occurs before the formation of seeds can result in huge losses for farmers.

Figure 2. *Corn Leaf affected gray spot.*

3. Blight

This disease is caused by *Helminthosporium sp* which is one of the most important diseases of maize after common rust. The development of this disease is caused by non-optimal environmental conditions, one of which is temperature. The optimal temperature for the development of the disease ranges between 20°C-30°C, because this temperature is often found by Helminthosporium sp in maize areas. Symptoms that can be seen are slightly elongated spots in the middle of the leaves that widen and get smaller on the edges, grayish brown surrounded by yellow along the veins.

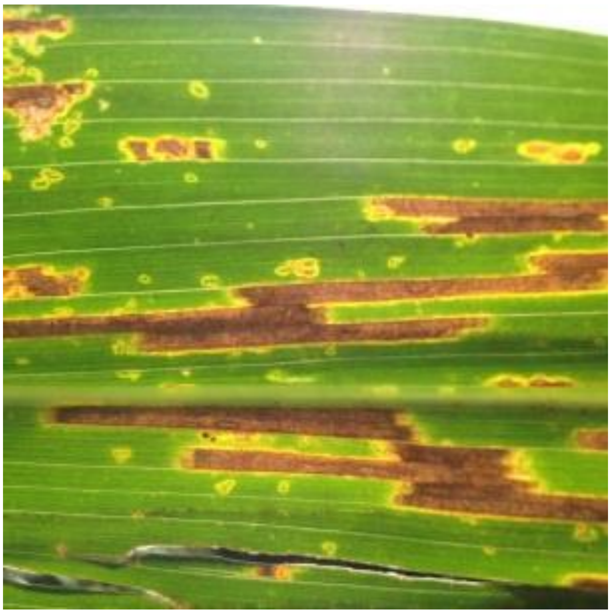


Figure 3. *Corn Leaf affected blight.*

B. Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) is an algorithm that is often used in Deep Learning, a development of Multilayer Perceptron (MPL), which is designed as an algorithm for processing data in the form of 2D data such as images. CNN is used to classify labeled data used in the supervised learning method (Hidayat & Darusalam, 2019). CNN consists of a convolutional layer structure and a pooling layer used for feature extraction and a fully connected layer as a classifier (Zhu et al. 2018).

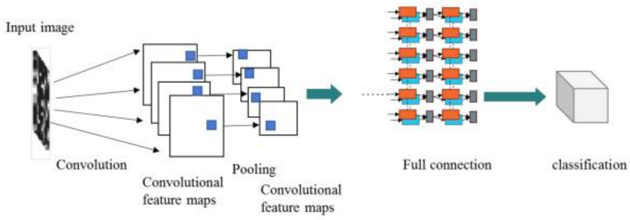


Figure 4. *Convolutional Neural Network.*

1. Pooling

Pooling is the process of reducing the size of an image which is then used as input. This process makes it easy for CNN to recognize an object. The pooling layer will take the largest value (max-pooling) of the image pixels. The method that is widely used in CNN research is the max-pooling method (Liu et al.2022).

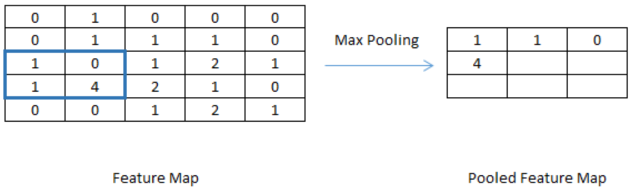


Figure 5. *Pooling.*

2. Full Connection

The full connection layer is the layer where all the activating neurons of the previous layer are connected to the neurons of the next layer. All activations in the previous layer are converted into one-dimensional data before it can be connected to all connected neurons in all layers.

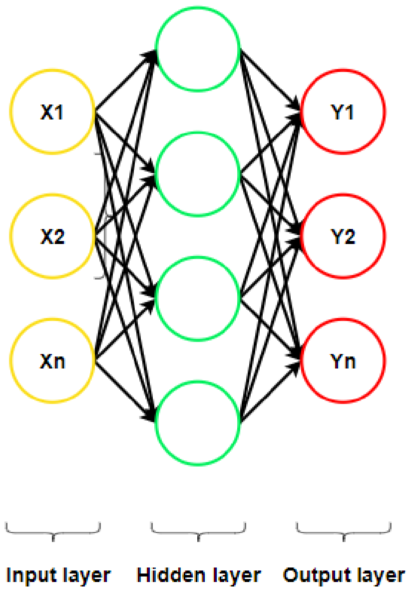


Figure 6. *Full Connection*

3. Clasification / Output

After the training process and the classification process are complete, the last layer is the output layer, in the output layer there are 4 types of images, namely healthy leaves, rusty leaves, leaf spots and leaf blight.

C. Gaussian Filter

Gaussian filter is a denoising method that can be used for image processing to make it smoother. The purpose of the Gaussian filter is to remove noise from the image and improve the quality of the image detail. By using a Gaussian filter on the image to be used in this study, it is hoped that there will be an increase in the accuracy of CNN (Kumar et al. 2020).

D. Research Framework

The method proposed for this study is a Convolutional Neural Network with Gaussian denoising (DeCNN). The method used will focus on using the Gaussian filter as a filter which will reduce noise in the image data, and get better image data. The following are the stages of the proposed method process, Figure 7

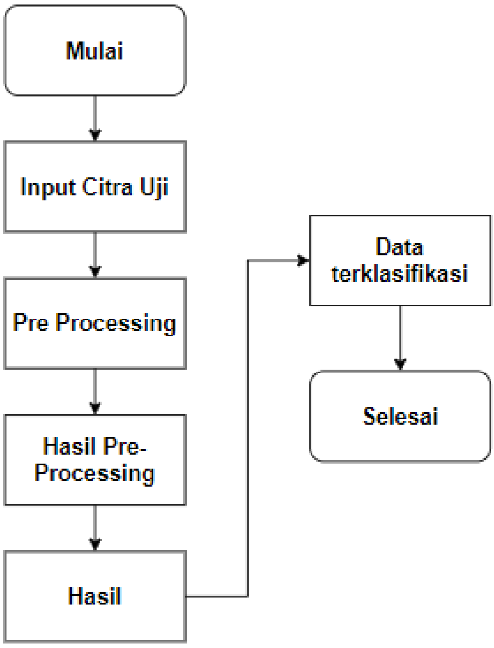


Figure 7. *Research Framework.*

E. Preprocessing

At the preprocessing stage, the dataset used will be processed. The dataset used will go through the Gaussian filter, Resize, Greyscale and Augmentation processes. After the process, there will be labeling on each image. For the ratio of training data and testing data from the data set is 80% as training data and 20% as data testing (Kc et al2021).

1. Gaussian Filter

Image data that will be used in research will be through the provision of Gaussian filters to remove noise from image data and improve the quality of image data (Mafi et al. 2019).

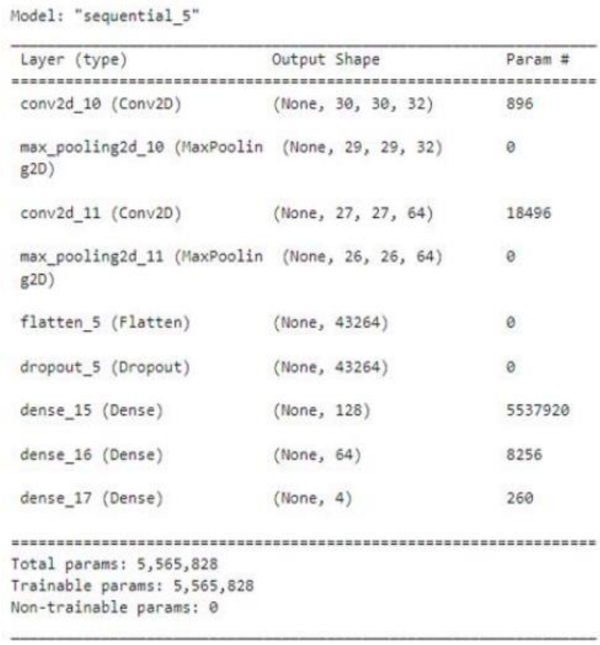
2. Resize
At this stage, preparing the image data to be used is to equate the size of the image data to be used to 28x28 pixels. The image data that will be used will go through a resizing process so that the same image data size will be obtained (Hidayat et al. 2019).

3. Greyscale
After the image data has gone through the resizing process, it will then go through the greyscale process. Image data that was previously in RGB format will be changed to gray (Lesmana et al. 2022).

4. Augmentasi
Augmentation is the process of modifying or manipulating an image data, so that the original image in the standard form will be changed in shape and position. The augmentations that will be performed on the image data that will be used are horizontal flip, vertical flip, rotation and color jitter (Iswantoro et al.2022).

F. Training Process Analysis
The training process is carried out using the Convolutional Neural Network (CNN) algorithm combined with an image dataset that has gone through a Gaussian filter pre-processing process. The Convolutional Neural Network (CNN) model used in training is in the form of 2 convolutional layers with a kernel size in each layer of 3, using a strides/pixel shift value of 1, and using Relu activation.

In this study using pooling, namely maxpooling. After the dataset has gone through the convolutional and pooling layers, the dataset will be changed using flatten, the values in the data will be sorted into one row. After the flatten process, then a dropout layer is added / removes pixel values that have a value below the predetermined value. The last stage is the data entering the classification layer.



```
Model: "sequential_5"
_________________________________________________________________
Layer (type)                Output Shape              Param #
=================================================================
conv2d_10 (Conv2D)          (None, 30, 30, 32)        896

max_pooling2d_10 (MaxPoolin  (None, 29, 29, 32)       0
g2D)

conv2d_11 (Conv2D)          (None, 27, 27, 64)        18496

max_pooling2d_11 (MaxPoolin  (None, 26, 26, 64)       0
g2D)

flatten_5 (Flatten)         (None, 43264)             0

dropout_5 (Dropout)         (None, 43264)             0

dense_15 (Dense)            (None, 128)               5537920

dense_16 (Dense)            (None, 64)                8256

dense_17 (Dense)            (None, 4)                 260

=================================================================
Total params: 5,565,828
Trainable params: 5,565,828
Non-trainable params: 0
_________________________________________________________________
```

Figure 8. *Training Methods.*

G. Evaluation Model
In this study, to evaluate the model, researchers used Accuracy, Precision and Recall to obtain performance evaluation results from the model developed with the proposed model. The Accuracy, Precision and Recall are made based on the Confusion Matrix table, in table 2 below

		Actual	
		Positive	Negative
Predicted	Positive	**TP (True Positive)**	**FP (False Positive)**
	Negative	**FN (False Negative)**	**TN (True Negative)**

Figure 9. *Table Confusion Matrix.*

Accuracy is the percentage of correct detection based on the amount of data used

$$Accuracy = \frac{TP + TN}{FP + FN + TP + TN}$$

Precision is the percentage of correct detection among all detected disease categories based on the model used. The following is the precision formula based on the confusion matrix table.

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

Recall is the percentage of correct detection of all test images for each category. The following is the recall formula based on the confusion matrix table.

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

Results and Discussion

The testing process was carried out using the Plantdoc dataset with 4000 corn leaf image data, consisting of 1000 healthy corn leaf image data, 1000 leaf rust affected leaf image data, 1000 corn leaf image data affected by leaf spots, and 1000 affected corn leaf image data. blight. The environment used for testing is as follows:

- Laptop Windows 11 Pro 64-bit
- CPU Intel i5-1135G7 @ 2.40GHz
- RAM 16 GB
- NVME 512 GB
- Jupyter Notebook

A. Evaluation Results

After training using the Convolutional Neural Network (CNN) algorithm, the next step is evaluating the results of the Convolutional Neural Network (CNN) algorithm classification using the Confusion Matrix. The evaluation of the Confusion Matrix used is the Confusion Matrix. From the evaluation of the classification using the Convolutional Neural Network (CNN), it can be seen in Figure 10

		Actual Value			
		Blight	Gray	Healthy	Rust
Predict Value	Blight	196	2	0	2
	Gray	1	198	1	0
	Healthy	3	1	196	0
	Rust	1	0	0	199

Figure 10. *Result Confusion Matrix*

1. Confusion Matrix Blight

The confusion matrix in the Blight class produces True Positive 196, False Positive 4, False Negative 5, True Negative 595. it can be seen in Figure 11

		Actual Value	
		Blight	Non Blight
Predict Value	Blight	196	4
	Non Blight	5	595

Figure 11. *Result Confusion Matrix Blight.*

2. Confusion Matrix Gray

The confusion matrix in the Gray class produces True Positive 198, False Positive 2, False Negative 3, True Negative 597. it can be seen in Figure 12

		Actual Value	
		Gray	Non Gray
Predict Value	Gray	198	2
	Non Gray	3	597

Figure 12. *Result Confusion Matrix Gray.*

3. Confusion Matrix Healthy

The confusion matrix in the Healthy class produces True Positive 196, False Positive 4, False Negative 1, True Negative 599. it can be seen in Figure 13

		Actual Value	
		Healthy	Non Healthy
Predict Value	Healthy	196	4
	Non Healthy	1	599

Figure 13. *Result Confusion Matrix Healthy.*

4. Confusion Matrix Rust

The confusion matrix in the Rust class produces True Positive 199, False Positive 1, False Negative 2, True Negative 598. it can be seen in Figure 14

		Actual Value	
		Rust	Non Rust
Predict Value	Rust	199	1
	Non Rust	2	598

Figure 14. *Result Confusion Matrix Rust.*

B. Comparison with Previous Research

After obtaining the values of TP, FP, FN and TN using the evaluation of the confusion matrix, then the values for Accuracy, Precision and Recall can be identified.

Table 2. Evaluation Confusion Matrix

	Accuracy	*Precision*	*Recall*
Result	99%	99%	99%

With the values obtained from the confusion matrix, it can be seen that the use of Gaussian Filters on image data can increase accuracy. As for the comparison of Accuracy values with previous studies (Bhatt et al. 2019) (Kaya et al. 2019) (Yang et al. 2020) (Panigrahi et al. 2020) (Li et al. 2020) as in table 4

Table 3. Previous Research.

Title	Author	Year	Accuracy
Identification of Diseases in Corn Leaves using Convolutional Neural Networks and Boosting	Prakruti Bhatt et al	2019	98%
Analysis of transfer learning for deep neural network based plant classification models	Aydin Kaya et al	2019	90%
The Implementation of A Crop Diseases APP Based on Deep Transfer Learning	Mengji Yang et al	2020	92%
A CNN Approach for Corn Leaves Disease Detection to support Digital Agricultural System	Kshyanaprava Panda et al	2020	98%
Do we really need deep CNN for plant diseases identification?	Yang Li et al	2020	94%

Conclusions

Using a Gaussian filter on the image data used can improve the accuracy of the Convolutional Neural Network (CNN) algorithm compared to previous studies. The results of the evaluation of the image data used are Accuracy 99%, Precision 99%, and Recall 99%.

The dataset in this study is a dataset that has been collected by previous researchers, it is necessary to add datasets from each region in Indonesia to increase the accuracy of the data.

This research is limited in improving the algorithm from previous research, it is hoped that with this research there will be mobile development that can help corn farmers.

Conflict of Interest: During the writing and course of this research, the author has no conflicts with any parties. Any information submitted is original as obtained when conducting research and is not influenced by personal opinion or interests.

References

Agarwal, M., Gupta, S. K., & Biswas, K. K. (2020). Development of Efficient CNN model for Tomato crop disease identification. *Sustainable Computing: Informatics* *and Systems*, *28*. https://doi.org/10.1016/j.suscom.2020.100407

Assefa, F., & Ayalew, D. (2019). Status and control measures of fall armyworm (Spodoptera frugiperda) infestations in maize fields in Ethiopia: A review. *Cogent Food and Agriculture*, *5*(1). https://doi.org/10.1080/23311932.2019.1641902

Bhatt, P., Sarangi, S., Shivhare, A., Singh, D., & Pappula, S. (2019). Identification of diseases in corn leaves using convolutional neural networks and boosting. *ICPRAM 2019 - Proceedings of the 8th International Conference on Pattern Recognition Applications and Methods*. https://doi.org/10.5220/0007687608940899

Burhanudin. (2015). Prefensi Penyakit Karat Daun (Puccinia polysora Undrew) pada Tanaman Jagung. *Prosiding Seminar Nasional Serealia*.

Chhillar, A., Thakur, S., & Rana, A. (2020). Survey of Plant Disease Detection Using Image Classification Techniques. *2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, 1339–1344. https://doi.org/10.1109/ICRITO48877.2020.9197933.

Harvey, C. A., Rakotobe, Z. L., Rao, N. S., Dave, R., Razafimahatratra, H., Rabarijohn, R. H., Rajaofara, H., & MacKinnon, J. L. (2014). Extreme vulnerability of smallholder farmers to agricultural risks and climate change in Madagascar. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1639). https://doi.org/10.1098/rstb.2013.0089

Hidayat, A., & Darusalam, U. (2019). Detection of Disease on Corn Plants Using Convolutional Neural Network Methods. *Jurnal Ilmu Komputer Dan Informasi (Journal of a Science and Information)*, *12*(1), 51–56.

Iswantoro, D., & Handayani UN, D. (2022). Klasifikasi Penyakit Tanaman Jagung Menggunakan Metode Convolutional Neural Network (CNN). *Jurnal Ilmiah Universitas Batanghari Jambi*, *22*(2). https://doi.org/10.33087/jiubj.v22i2.2065

Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). A review of the use of convolutional neural networks in agriculture. In *Journal of Agricultural Science* (Vol. 156, Issue 3, pp. 312–322). Cambridge University Press. https://doi.org/10.1017/S0021859618000436

Kaya, A., Keceli, A. S., Catal, C., Yalic, H. Y., Temucin, H., & Tekinerdogan, B. (2019). Analysis of transfer learning for deep neural network based plant classification models. *Computers and Electronics in Agriculture*, *158*, 20–29. https://doi.org/10.1016/j.compag.2019.01.041

Kc, K., Yin, Z., Li, D., & Wu, Z. (2021). Impacts of background removal on convolutional neural networks for plant disease classification in-situ. *Agriculture (Switzerland)*, *11*(9). https://doi.org/10.3390/agriculture11090827

Kementerian pertanian republik indonesia. (2021, September 21). *Panen jagung nusantara, bukti pasokan jagung melimpah*. Www.Pertanian.Go.Id.

Kompas. (2008, April 25). *Puluhan hektar tanaman jagung diserang penyakit bulai daun*. Www.Nasional.Kompas.Com.

Kumar, A., & Sodhi, S. S. (2020). Comparative Analysis of Gaussian Filter, Median Filter and Denoise Autoenocoder. *46 7th International Conference on Computing For Sustainable Global Development (INDIACom)*, 45–51.

Lesmana, A. M., Fadhillah, R. P., & Rozikin, C. (2022). Identifikasi Penyakit pada Citra Daun Kentang Menggunakan Convolutional Neural Network (CNN). *Jurnal Sains Dan Informatika*, *8*(1). https://doi.org/10.34128/jsi.v8i1.377

Li, Y., Nie, J., & Chao, X. (2020). Do we really need deep CNN for plant diseases identification? *Computers and Electronics in Agriculture*, *178*. https://doi.org/10.1016/j.compag.2020.105803

Liu, X., Zhou, S., Chen, S., Yi, Z., Pan, H., & Yao, R. (2022). Buckwheat Disease Recognition Based on Convolution Neural Network. *Applied Sciences (Switzerland)*, *12*(9). https://doi.org/10.3390/app12094795

Mafi, M., Martin, H., Cabrerizo, M., Andrian, J., Barreto, A., & Adjouadi, M. (2019). A comprehensive survey on impulse and Gaussian denoising filters for digital images. In *Signal Processing* (Vol. 157, pp. 236–260). Elsevier B.V. https://doi.org/10.1016/j.sigpro.2018.12.006

Panigrahi, K., Sahoo, A., & Das, H. (2020). A CNN Approach for Corn Leaves Disease Detection to support Digital Agricultural System. *2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184)*, 678–683. https://doi.org/10.1109/ICOEI48184.2020.9142871

Scotland, R. W., & Wortley, A. H. (2003). How Many Species of Seed Plants Are There? In *Source: Taxon* (Vol. 52, Issue 1).

Tai, A. P. K., Martin, M. V., & Heald, C. L. (2014). Threat to future global food security from climate change and ozone air pollution. *Nature Climate Change*, *4*(9), 817–821. https://doi.org/10.1038/nclimate2317

Thangaiyan, J., Vinoth Kumar, K., & Jayasankar, T. (2019). An Identification of crop disease using image segmentation. *International Journal of Pharmaceutical Sciences and Research*, *10*(3), 1054. https://doi.org/10.13040/IJPSR.0975-8232.10(2).1000-11

Tugrul, B., Elfatimi, E., & Eryigit, R. (2022). Convolutional Neural Networks in Detection of Plant Leaf Diseases: A Review. In *Agriculture (Switzerland)* (Vol. 12, Issue 8). MDPI. https://doi.org/10.3390/agriculture12081192

Yang, M., Li, D., Chen, M., Bouras, A., Tang, Y., & Yu, X. (2020). The Implementation of A Crop Diseases APP

Based on Deep Transfer Learning. *2020 3rd International Conference on Artificial Intelligence and Big Data (ICAIBD)*, 22–28. https://doi.org/10.1109/ICAIBD49809.2020.9137469

Zhu, N., Liu, X., Liu, Z., Hu, K., Wang, Y., Tan, J., Huang, M., Zhu, Q., Ji, X., Jiang, Y., & Guo, Y. (2018). Deep learning for smart agriculture: Concepts, tools, applications, and opportunities. *International Journal of Agricultural and Biological Engineering*, *11*(4), 21–28. https://doi.org/10.25165/j.ijabe.20181104.4475

THIS PAGE INTENTIONALLY LEFT BLANK