KOMPUTASI: JURNAL ILMIAH ILMU KOMPUTER DAN MATEMATIKA
VOL. 22 (2) (2025), 20-27, p-ISSN: 693-7554, e-ISSN:2654-3990

Shallot Growth Stage Monitoring with Multispectral Imagery using MobileNetV2

Yanuar Firmansyah [1],**Nugraha Priya Utama**[2]

[1,2] School of Electrical Engineering and Informatics, Institut Teknologi Bandung, Bandung, 40132, Indonesia

Abstract

Shallots are a strategic horticultural commodity, playing a pivotal role in the Indonesian food supply as a staple food. It is anticipated that the demand for shallots will increase in the future. Nonetheless, the production level remains inadequate, necessitating the importation of shallots. The success of the production process is contingent upon the monitoring of the crops by the farmers themselves. The current manual approach to this process has proven to be inadequate, necessitating the development of agricultural technology. Remote sensing has the potential to become a valuable technology in agriculture and represents a feasible methodology for accurate land mapping. However, its application in Indonesia remains constrained by the heterogeneity of the mapped areas and the quality of the images obtained. Furthermore, Indonesia's agricultural land is extensive and heterogeneous. This research presents the use of multispectral cameras for the precise monitoring and mapping of agricultural land. The imaging system employs a multispectral camera attached to a UAV to perform land mapping. RGB imagery and vegetation indices, including NDVI, NDRE, and OSAVI, are calculated to ascertain the crop status within a single shallot planting period. The shallot growth stage prediction is proposed using MobileNetV2 architecture. The proposed method demonstrated an accuracy of 84.78% using RGB and 44.78% using a combination of vegetation indices.

Keywords: *Shallot, Unmaned Aerial Vehicle, Multispectral Imagery, Remote Sensing, MobileNetV2*

1. Introduction

Shallots are a strategic horticultural commodity due to their function as a staple food in Indonesia. The demand for shallots is increasing in line with population growth. It is necessary to balance these efforts with initiatives to increase shallot production on a national scale. However, the gap between production and consumption remains significant. This gap is not merely quantitative; it also manifests in terms of temporality with the consequence that imports of shallots remain a feature of the market. The planting of shallots typically occurs during the dry season, whereas in the rainy season farmers are reluctant to cultivate shallots due to the elevated risk of pest infestation [1].

Remote sensing represents a viable methodology for the accurate mapping of terrestrial environments [2], [3]. The application of this approach in Indonesia remains limited. The primary challenge arises from the heterogeneous characteristics of mapped areas and the quality of the images obtained. Additionally, the vast and diverse agricultural landscapes of Indonesia further complicate the use of remote sensing techniques.

The most common method of land mapping is the use of multispectral satellite or UAV (Unmanned Aerial Vehicle) imagery. The use of portable and easily operated UAVs has become a widely adopted approach in recent times [4]. Furthermore, the variety of payload options that can be carried by UAV according to their flight missions provides additional support for this assertion. The accuracy of drone imagery, which can reach 2 cm, is a notable improvement over satellite imagery. However, with a flight time of approximately 15–20 minutes, the land that can be mapped in a single flight is limited to 8–15 hectares, depending on the UAV's flight altitude [5]. The use of UAV is well-suited to mapping relatively small areas that require regular observation. The utilization of UAV in the agricultural sector is intricately linked to the implementation of multispectral camera payloads, which are employed in numerous remote sensing applications. These include vegetation and land mapping, leaf area index measurement to assess plant canopy, land estimation, assessment of fertilizer and biomass effects, pest management, detection of

*Corresponding author. *E-mail address: 23522054@std.stei.itb.ac.id*
Received: 19 April 2025, Accepted: 31 July 2025 and available online 31 July 2025
DOI: https://doi.org/10.33751/komputasi.v19i2.5260

plant stress levels, and land drought levels [6]. The utilization of drones in image analysis has emerged as a compelling alternative to satellite imaging, particularly in scenarios where the study focuses on small objects, such as plants, habitat conditions, and even animals [5].

Multispectral imagery is a type of digital imagery that consists of data from multiple spectral bands of the electromagnetic spectrum, extending beyond the visible light range. This type of imagery is frequently employed in remote sensing and other applications to capture and analyze data on various environmental factors. In this study, multispectral cameras were utilized to generate vegetation indices such as Normalized Difference Vegetation Index (NDVI), Normalized Difference Red Edge Index (NDRE), and Optimized Soil-Adjusted Vegetation Index (OSAVI), which facilitate the monitoring of shallot plant conditions. The NDVI index was selected due to its sensitivity to plant greenness [7], while the NDRE index is employed for the detection of leaf chlorophyll content [8]. The selection of OSAVI was driven by its capacity to mitigate the impact of background soil in areas with minimal vegetation coverage [9]. The selection of these indices was predicated on their demonstrated relevance in the context of shallot plant condition monitoring during the growth period.

MobileNetV2 is a convolutional neural network (CNN) architecture that builds upon the original MobileNet, designed specifically for efficient mobile and embedded vision applications. It was introduced by Google [10] in 2018. The key objective of MobileNetV2 is to provide lightweight, fast, and resource-efficient architecture for mobile devices while maintaining high performance for tasks like image classification, object detection, and segmentation. The potential of MobileNetV2 for land mapping has been the subject of recent research. This architectural model has been demonstrated to be lightweight and to perform well in a variety of land mapping applications, including the vine disease detection [11], mapping banana disease [12], plant counting [13][14], and paddy seedling monitoring [15]. Research on monitoring onion fields using multispectral imagery to determine the effects of nitrogen fertilizer has been conducted [16]. Another study employs multispectral imagery to ascertain NDVI (Normalized Difference Vegetation Index) for estimating the growth phase of shallots through a deep learning approach [17].

In this study, a multi-rotor unmanned aerial vehicle (UAV) equipped with a multispectral camera is employed to obtain spectral data over the course of a single shallot planting period. A convolutional neural network (CNN) model, MobileNetV2, is utilized to predict the growth stage of the shallots based on a combination of vegetation indices derived from multispectral drone imagery. The case study for this topic is situated in Ngantang District, Malang Regency, Indonesia. The findings of research are expected to support the implementation of smart agriculture in Indonesia, especially for farmers and related stakeholders.

2. Methods

This section presents the details of the proposed methodology. Figure 1 illustrates the designed system architecture. The multispectral images obtained from the UAV are processed into orthomosaic maps. The RGB maps and vegetation index combination maps generated are then used as datasets for the MobileNetV2 model. Ground truth annotating is conducted to create growth stage labels. This combination will result in the production of a shallot growth stage monitoring model.

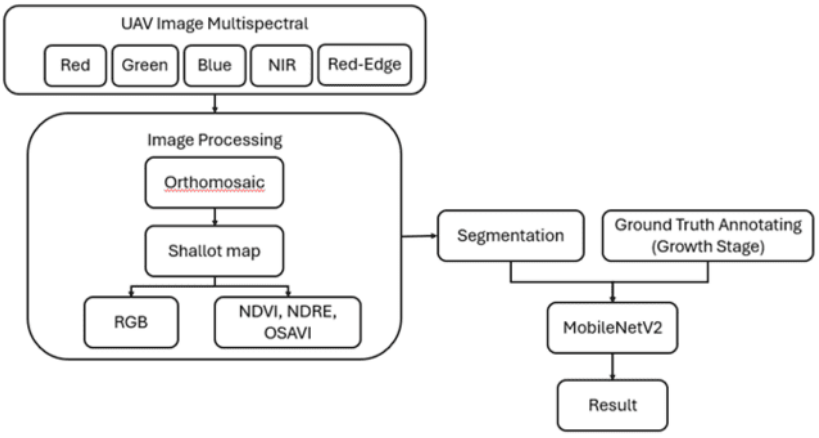

Figure 1. System architecture

2.1. Experimental Design

The field experiment was conducted in 2022 in agricultural land in Ngantang Village, Malang Regency, East Java Province, Indonesia (7°52'06.6"S, 112°22'18.4"E), as shown in Figure 2.

Figure 2. Experiment location

The experimental area is located in the shallot agricultural area on a mountain slope with an average altitude of 625 meters above sea level. The land used for the experiment has an area of 6.3 hectares. Land mapping was carried out on six occasions over the duration of one planting period (two months), with an interval of approximately one week between each mapping survey.

2.2. Field data acquisition

Field data acquisition was conducted for the construction of labels for the model. This data collection was conducted simultaneously with the image data collection in Ngantang, Malang. A total of 198 plant sampling points that are illustrated in Figure 3 were conducted over the course of a single planting period, with each point being sampled six times.

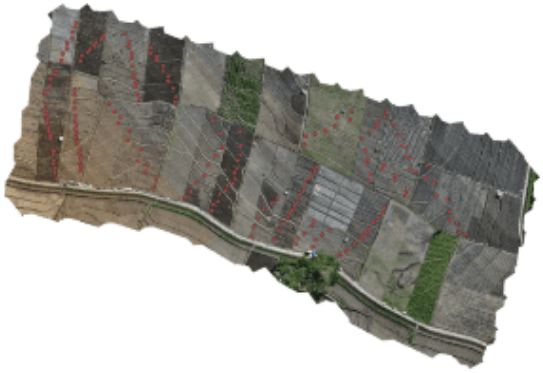

Figure 3. Field data acquisition point

The data obtained was the age of the plant, which was used as the basis for the classification of the shallot growth stage, as shown in Table 1. There are six classifications of the shallot growth stage: germination, seedling, bulb initiation, bulb growth, maturation, and harvest. This classification was used as the label for the model and was added with one more category, namely soil.

Table 1. Classification of the shallot growth stage

Class_id	Age (days)	Classification
0	0	Soil
1	1 – 9	Germination
2	10 – 19	Seedling
3	20 – 29	Bulb Initiation
4	30 – 39	Bulb growth
5	40 – 49	Maturation
6	50 – 65	Harvest

2.3. Acquisition of UAV Images

Multispectral imagery was obtained using the DJI P4 multispectral RTK drone, which is a variation of the DJI Phantom 4 drone from the manufacturer DJI. It has been modified with the addition of a multispectral camera and RTK (real-time kinematic) device. The RTK device enables the drone to perform missions with a position accuracy of up to 0.1 meters. Figure 4 illustrates the devices used in data collection.

Figure 4. DJI P4 multispectral

The multispectral camera embedded in the DJI P4 Multispectral RTK drone has a 5 band + 1 RGB image configuration. Each camera uses a 1 / 2.9 "CMOS sensor of 2.08 MP. This camera has a FOV (Field of View) of 62.7 ° with a focal length of 5.74 mm and an aperture of f / 2.2. The specifications of each multispectral camera band are as follows:

- Blue (B) : 450 nm +- 16 nm
- Green (G) : 560 nm +- 16 nm
- Red (R) : 650 nm +- 16 nm
- Red-edge (RE) : 730 nm +- 16 nm
- Near-infrared (NIR) : 840 nm +- 26 nm

The mapping of shallot fields was conducted at an flight height of 50 metres and a airspeed of 4.3 metres per second. The front and side overlap ratios were 75 per cent. The ground sampling distance (GSD) was obtained at 3 centimetres. The flight was conducted in sunny and windy weather between 8 a.m. and 12 p.m.

2.4. Selection of Vegetation Index

A variety of vegetation indices can be derived through the combination of bands from multispectral imagery. The indices that will be used in this study are the Normalized Difference Vegetation Index (NDVI), the Normalized Difference Red Edge Index (NDRE), and the Optimized Soil-Adjusted Vegetation Index (OSAVI).

The NDVI is a measure of the value of plants, calculated by subtracting the red band (absorbed by plants) from the near-infrared band (reflected by vegetation) [18]. The NDVI index is the most commonly utilised index in remote sensing for the assessment of health indicators or levels of green vegetation density [7]. NDVI values range from -1 to 1. The calculation of NDVI is illustrated in equation 1.

$$NDVI = \frac{(NIR - Red)}{(NIR + Red)} \quad (1)$$

The NDRE index is derived by combining the NIR band with the Red-Edge band. The term 'Red-Edge' is used to describe this band because it is in the transition frequency band between the red band and the near-infrared (NIR) band. The index is sensitive to the chlorophyll value of plants, which is the primary indicator of internal conditions in leaves. This is due to the Red-Edge band's ability to penetrate leaves at a deeper level than the red light utilized in the NDVI index, thereby providing insights into soil conditions or pests in plants [8]. The formula for calculating NDRE is presented in Equation 2:

$$NDRE = \frac{(NIR - Red\ Edge)}{(NIR + Red\ Edge)} \quad (2)$$

The OSAVI is a remote sensing vegetation index used to assess and monitor vegetation, particularly in areas where soil is more visible, such as regions with sparse plant cover. It is designed to reduce the influence

of soil brightness on the vegetation signal, providing more accurate vegetation measurements under these conditions [19]. The formula for calculating OSAVI is presented in Equation 3:

$$OSAVI = \frac{(NIR - Red)}{(NIR + Red + 0.16)} \quad (3)$$

2.5. Processing of UAV imagery

The geometric correction process is conducted with the objective of guaranteeing the absence of errors or shifts in position (geotag) for each image obtained. Subsequently, the orthorectification process is carried out with the objective of eliminating the distortion effects caused by the image perspective (typically on the side of the image) and the topographic relief of the mapped location. To mitigate the impact of image perspective, the flight mission is conducted with an overlap of 75% between image captures, ensuring the acquisition of at least 10 images at each point. The aim of these processes is to guarantee the optimal quality of the resulting orthomosaic image. Orthomosaic is the process of combining multiple images into a single, comprehensive representation of an area. The entire image preprocessing procedure is conducted using DJI Terra software. Two orthomosaic maps are generated, which will serve as the basis for the model dataset: the RGB map and the vegetation index combination map (NDVI, NDRE, OSAVI).

2.6. Modelling method

MobileNetV2 is composed of an initial convolutional layer followed by a series of inverted residual blocks and ends with a fully connected layer for classification. The basic building block of the architecture is the inverted residual block, which consists of three layers shown in figure 5:

- Expansion layer (1x1 convolution)
- Depthwise convolution (3x3 spatial convolution)
- Projection layer (1x1 convolution)

In the output layer, changes are made so that the number of classes classified by the network is 7 according to the growth stage of the shallot [20].

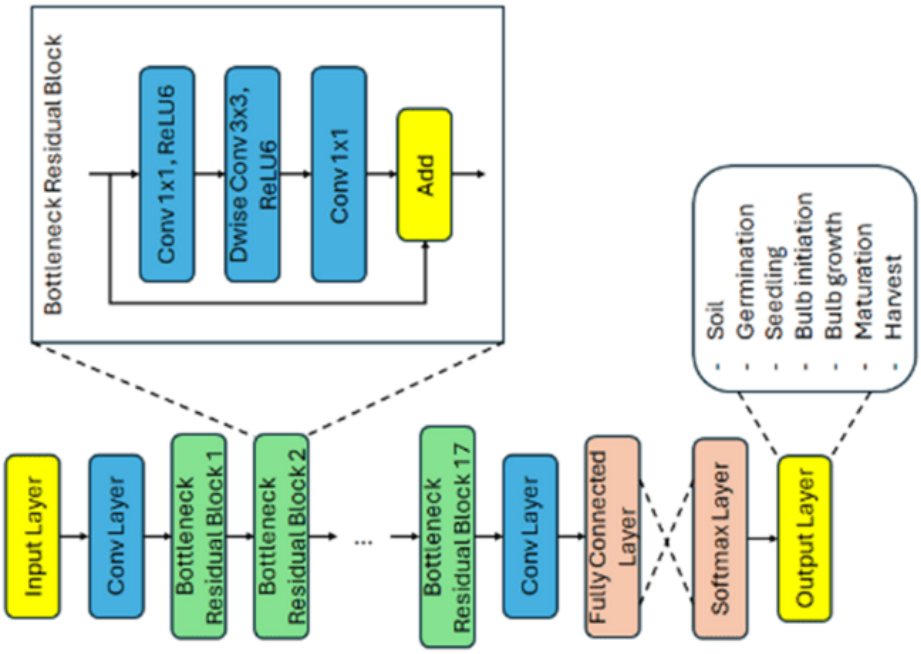

Figure 5. A modified MobileNetV2 architecture

3. Result and Discussion

3.1. Orthomosaic map and Pre-processing

The multispectral imagery obtained from drones is processed into two orthomosaic maps. The first map is an RGB map composed of the conventional red, green, and blue bands. The range of band values in this map is between 0 and 255. The second map is a combination of three vegetation indices, namely NDVI, NDRE, and OSAVI, which form three bands of an orthomosaic map. The range of band values for this map is between -1 and 1. In general, values closer to -1 indicate non-vegetated land, such as rocks, water, or other non-vegetated areas. Soil will typically exhibit a value of approximately 0, while plants will display a positive value dependent on their health status. A value closer to 1 indicates a higher level of health. Figure 6

illustrates the results of the orthomosaic map in RGB format (1) and vegetation index format (2).

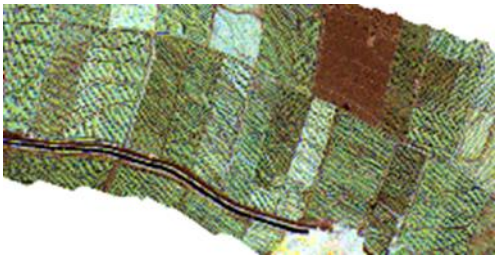

Figure 6. RGB map (1) and vegetation index map (2)

Six RGB maps and vegetation indexes, each measuring 14602x10256, are available for analysis. The area covered by each map is 6.2 hectares. The image was resized to 4867x3418 to facilitate the computation process. The feature extraction is conducted through the implementation of the Meanshift filter segmentation algorithm. This process is semi-automatically executed by comparing the segmentation results with the field data results, thereby producing seven shallot plant growth stage classes. The process is conducted utilizing the QGis 3.28.15-Firenze application. Subsequently, image chips are generated to create a dataset. The resulting image chips are 7x7 in size, with a total of 998,124,360 data.

3.2. Shallot growth stage classification

The constructed dataset comprises a total of 998,124,360 image chips. The preprocessing of the data is conducted to ensure that the number of datasets for each class is identical. The rescaling process is carried out due to the different value ranges between the RGB map and the vegetation index map, where the RGB map has a scale of 0 to 255 and the vegetation index map has a scale of -1 to 1. The data normalisation process is carried out so that both maps have the same value range between 0 and 1, to ensure that each feature is treated equally. Subsequently, the dataset is partitioned into a training set of 60% and a testing set of 40%.

The MobileNetV2 algorithm with 10 epoch parameters and the SGD optimiser was implemented for the execution of the testing phase. The algorithm has 3.5 million parameters. The principal benefit of this approach is that it features inverted residuals, linear bottlenecks, and depthwise separable convolutions, which collectively contribute to its high efficiency and power, while also ensuring a relatively lightweight network. The training process was conducted on a system configured with an Intel Xeon Silver 4208 CPU, a Nvidia Quadro RTX 5000 GPU, and 128 gigabytes of RAM.

Testing accuracy is a prediction that demonstrates the precision and accuracy of any selected deep learning model. 4 parameters are used to evaluate the results, namely accuracy, precision, recall, and F1 score. Accuracy is the proportion of correctly classified instances out of the total number of instances.

$$Accuracy = \frac{TP + TN}{Total\ Population\ (TP + TN + FP + FN)} \tag{4}$$

In which :

- True Positives (TP): Correctly predicted positive instances.
- True Negatives (TN): Correctly predicted negative instances.
- False Positives (FP): Incorrectly predicted positive instances.
- False Negatives (FN): Incorrectly predicted negative instances.

Precision measures the proportion of correctly predicted positive instances out of all instances that were predicted as positive. High precision means the model produces few false positives.

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

Recall measures the proportion of correctly predicted positive instances out of all actual positive instances. High recall means the model produces few false negatives.

$$Recall = \frac{TP}{TP + FN} \tag{6}$$

The F1 score is the harmonic mean of precision and recall. It provides a balance between precision and recall. The F1 score ranges from 0 (worst) to 1 (best). High F1 score indicates a good balance between

precision and recall.

$$F1\ score = 2\ x\ \frac{Precision\ x\ Recall}{Precision\ +\ Recall} \quad (7)$$

Table 2. Model performance.

Map	Accuracy	Precision	Recall	F1-score
RGB	84.78%	88%	84.78%	84.85%
NDVI, NDRE, OSAVI	44.78%	56%	44.78%	41.99%

Table 2 shows the results of the model's performance in terms of accuracy, precision, recall, and F1 score. The evaluation results for the model are identical for all four metrics, with the RGB map achieving an accuracy of 85% and the vegetation index map reaching 44.78%.

4. Conclusion

This study explores the potential of utilizing a multispectral camera mounted on a multi-rotor unmanned aerial vehicle (UAV) for the monitoring of the growth stage of shallot plants. A comparison was carried out between the NDVI, NDRE, and OSAVI vegetation index maps and the RGB map. The outcomes showed that the RGB images demonstrated superior performance in comparison to the combination of vegetation indices. This is likely related to the parameters used in the preprocessing stage or to the parameters applied in the agricultural land objects under study. The UAV sensing system model presented in this study can provide additional references and technical support for management and decision-making in medium-scale agriculture. Moreover, the findings of this research offer a practical solution for monitoring shallot plants in Indonesia, a domain that has received scant attention from previous studies. However, the limited dataset and the sparse coverage of the research locations represent significant challenges that must be addressed in future research endeavors. The application of a combination of other vegetation indices and a variety of agricultural land types will enhance the versatility of this method for a multitude of applications.

References

[1] Pusat Data dan Informasi Kementrian Pertanian, "Analisis Kinerja Perdagangan Bawang Merah Semester 1 2021," pp. 1–58, 2021.

[2] Q. Yang, M. Liu, Z. Zhang, S. Yang, J. Ning, and W. Han, "Mapping plastic mulched farmland for high resolution images of unmanned aerial vehicle using deep semantic segmentation," Remote Sens., vol. 11, no. 17, 2019, doi: 10.3390/rs11172008.

[3] Z. Fu et al., "Wheat growth monitoring and yield estimation based on multi-rotor unmanned aerial vehicle," Remote Sens., vol. 12, no. 3, 2020, doi: 10.3390/rs12030508.

[4] Y. Unpaprom, N. Dussadeed, and R. Ramaraj, "Modern Agriculture Drones," Mod. Agric., vol. 1, no. July, pp. 13–19, 2018.

[5] L. J. Mangewa et al., "Comparative Assessment of UAV and Sentinel-2 NDVI and GNDVI for Preliminary Diagnosis of Habitat Conditions in Burunge Wildlife Management Area, Tanzania," Earth (Switzerland), vol. 3, no. 3, pp. 769–787, 2022, doi: 10.3390/earth3030044.

[6] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," Comput. Electron. Agric., vol. 147, no. February, pp. 70–90, 2018, doi: 10.1016/j.compag.2018.02.016.

[7] S. Huang, L. Tang, J. P. Hupy, Y. Wang, and G. Shao, "A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing," J. For. Res., vol. 32, no. 1, pp. 1–6, 2021, doi: 10.1007/s11676-020-01155-1.

[8] B. Boiarskii and M. Sinegovskii, "Application of NDVI and NDRE vegetation indices in the assessment of soybean productivity under nitrogen controlled-release fertilizer," 2022 8th Int. Conf. Inf. Technol. Nanotechnology, ITNT 2022, pp. 5–10, 2022, doi: 10.1109/ITNT55410.2022.9848588.

[9] H. Ren and G. Feng, "Are soil-adjusted vegetation indices better than soil-unadjusted vegetation

indices for above-ground green biomass estimation in arid and semi-arid grasslands?," Grass Forage Sci., vol. 70, no. 4, pp. 611–619, Dec. 2015, doi: https://doi.org/10.1111/gfs.12152.

[10] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., pp. 4510–4520, 2018, doi: 10.1109/CVPR.2018.00474.

[11] M. G. Aruna, E. Silvia, R. R. Al-Fatlawy, H. K. Rao, and M. Sowmya, "Vine Disease Detection UAV Multi Spectral Image using Segnet and Mobilenet Method," Int. Conf. Distrib. Comput. Optim. Tech. ICDCOT 2024, pp. 1–4, 2024, doi: 10.1109/ICDCOT61034.2024.10515972.

[12] R. Linero-ramos, C. Parra-rodr, A. Espinosa-valdez, G. Jorge, and M. Gongora, "Assessment of Dataset Scalability for Classification of Black Sigatoka in Banana Crops Using UAV-Based Multispectral Images and Deep Learning Techniques," 2024.

[13] Y. Prabowo et al., "Palm Trees Counting Using MobileNet Convolutional Neural Network in Very High-Resolution Satellite Images," 2022 IEEE Asia-Pacific Conf. Geosci. Electron. Remote Sens. Technol. Underst. Interact. Land, Ocean. Atmos. Smart City Disaster Mitig. Reg. Resilience, AGERS 2022 - Proceeding, pp. 79–83, 2022, doi: 10.1109/AGERS56232.2022.10093287.

[14] Z. Lin and W. Guo, "Cotton stand counting from unmanned aerial system imagery using mobilenet and centernet deep learning models," Remote Sens., vol. 13, no. 14, pp. 1–16, 2021, doi: 10.3390/rs13142822.

[15] M. M. Anuar, A. A. Halin, T. Perumal, and B. Kalantar, "Aerial Imagery Paddy Seedlings Inspection Using Deep Learning," Remote Sens., vol. 14, no. 2, 2022, doi: 10.3390/rs14020274.

[16] G. Messina, S. Praticò, G. Badagliacca, S. Di Fazio, M. Monti, and G. Modica, "Monitoring onion crop 'cipolla rossa di tropea calabria igp' growth and yield response to varying nitrogen fertilizer application rates using uav imagery," Drones, vol. 5, no. 3, 2021, doi: 10.3390/drones5030061.

[17] N. U. Din, B. Naz, S. Zai, Bakhtawer, and W. Ahmed, "Onion Crop Monitoring with Multispectral Imagery using Deep Neural Network," Int. J. Adv. Comput. Sci. Appl., vol. 12, no. 5, pp. 303–309, 2021, doi: 10.14569/ijacsa.2021.0120537.

[18] Kriegler F.J., M. W.A., N. R.F., and W. Richardson, "Preprocessing transformations and their effects on multispectral recognition," Remote Sens. Environ., vol. VI, p. 97, 1969.

[19] Ł. Jełowicki, K. Sosnowicz, W. Ostrowski, K. Osińska-Skotak, and K. Bakuła, "Evaluation of rapeseed winter crop damage using UAV-Based multispectral imagery," Remote Sens., vol. 12, no. 16, 2020, doi: 10.3390/RS12162618.

[20] U. Seidaliyeva, D. Akhmetov, L. Ilipbayeva, and E. T. Matson, "Real-time and accurate drone detection in a video with a static background," Sensors (Switzerland), vol. 20, no. 14, pp. 1–18, 2020, doi: 10.3390/s20143856.