Deep Learning for Paddy Leaf Disease Segmentation: An Exploratory Study

Muhammad Amirul Aiman Asri, Wenjunliang Zhang, Norrima Mokhtar*, Afdhal Haziq Noramly
Department of Electrical Engineering, Faculty of Engineering, Universiti Malaya, Lembah Pantai, Kuala Lumpur, Malaysia.

Raza Ali
Balochistan University of Information Technology, Engineering and Management Sciences (BUITEMS), Quetta, Pakistan

Takao Ito
Graduate School of Advanced Science and Engineering, Hiroshima University

M. Aziz Muslim
Department of Electrical Engineering, Faculty of Engineering, Universitas Brawijaya, Malang, East Java Indonesia

Siti Sendari
Faculty of Engineering, Universitas Negeri Malang, Indonesia

Pringgo Widyo Laksono
Industrial Engineering, Universitas Sebelas Maret, Indonesia

Tsutomu Ito
Ube National College of Technology, Japan

Email: norrimamokhtar@um.edu.my
*Corresponding Author

Abstract

Pixel accurate segmentation of paddy leaf lesions is vital for field diagnostics, yet many prior studies examine only one disease and use inconsistent pipelines, which limits comparability. We present an experimental study and a unified benchmark on the public Kaggle New Paddy Doctor dataset with 2,444 RGB images at 480×640 covering Bacterial Leaf Blight (804 images), Brown Spot (936 images), and Hispa (704 images). We manually annotated pixel level masks for all images. Under one protocol, we evaluate U-Net, U-Net++, and DeepLabV3 with a ResNet-50 backbone in both per disease training and pooled training. DeepLabV3 with ResNet-50 attains the best average per disease Dice of about 0.70, with stronger scores on Bacterial Leaf Blight at about 0.775 and Hispa at about 0.730, while Brown Spot remains more challenging at about 0.599. Training a single pooled model gives slightly lower performance than training per disease, with a similar pattern for IoU. We also examine how stable the model rankings are across diseases and provide qualitative examples of typical failure modes such as background leakage and boundary errors. These results offer clear baselines and practical guidance on when pooling is helpful, and which architecture provides the best balance between accuracy and efficiency for deployment.

Keywords: Pixel accurate segmentation, Paddy leaf disease, Pooled vs per-disease, Accuracy-efficiency trade off

1. Introduction

Proper and timely diagnosis of paddy leaf diseases is crucial for crop protection and food security [1], yet in practice farmers still rely heavily on visual inspection, which is labour-intensive, subjective and difficult to scale to large plantations. This leads to inconsistent diagnoses, delayed interventions and unnecessary yield loss. Deep learning-based classification and detection systems for rice diseases have shown promising results using RGB leaf images and convolutional neural networks but most operate at the image or bounding-box level [2]. Pixel-level lesion segmentation can further support downstream tasks such as lesion area quantification, disease severity assessment and can be integrated with mobile or edge devices and UAV platforms for large-scale field assessment.

Deep convolutional encoder–decoder architectures, such as U-Net and its variants, as well as DeepLab-type networks, have demonstrated excellent performance in plant disease segmentation and other agricultural vision applications. However, many existing studies rely on zoom-in datasets in which individual leaves are

photographed in controlled environments with relatively clean backgrounds. In such settings, lesions occupy a significant portion of the image, the background is simple, and annotation is relatively straightforward. In contrast, real farms often present crop-level scenes, with multiple leaves, stems and cluttered backgrounds appearing in one image. Under these conditions, lesions may be small, thin, partially occluded and severely imbalanced with respect to background pixels. This substantially increases the difficulty of segmentation but also makes the problem more representative of real-world deployment scenarios.

For paddy leaves specifically, most studies focus on a single disease or use self-collected, private datasets, which restricts reproducibility and makes it difficult to compare architectures fairly across works. Experimental pipelines also differ in terms of preprocessing, input resolution, data augmentation, training, validation and test splits, and evaluation metrics, so reported Dice and Intersection-over-Union (IoU) scores are rarely directly comparable. The Paddy Doctor dataset introduced by Petchiammal *et al.* [3] has become a widely used public resource for paddy disease research, while a recent experimental review of deep learning-based paddy disease models [4] highlights that such datasets are typically used at the image level, with limited work on dense lesion masks and systematic multi-disease segmentation benchmarks. In realistic multi-disease scenarios, it therefore remains unclear whether separate per-disease models are necessary, or whether a single pooled model that shares capacity across diseases can achieve comparable performance under a unified segmentation pipeline.

In this work, we focus exclusively on pixel-level segmentation of three common and visually distinct paddy leaf diseases specifically Bacterial Leaf Blight (BLB), Brown Spot and Hispa by using crop-level images from the public Paddy Doctor dataset. We manually annotate lesion masks for all selected images and conduct a controlled comparison of three encoder–decoder architectures, U-Net, U-Net++ and DeepLabV3 with a ResNet-50 backbone, under a single, consistent pipeline. Each model is trained both in per-disease mode, where an individual binary model is learned for each disease, and in pooled mode, where one model is trained jointly on all three diseases by treating any lesion pixel as foreground. This design allows us to quantify the trade-off between per-disease and pooled training and to identify which architecture offers the best balance between segmentation accuracy and efficiency for practical deployment in paddy fields.

2. Literature Review

Deep learning has been widely applied to rice disease recognition at the image level. Early work uses artificial neural networks and cross-validation strategies to stabilize performance on relatively small paddy disease datasets. Recent studies propose deep CNN frameworks for multi-class rice disease detection and classification using RGB images [5]. These works show the feasibility of deep learning for practical disease recognition but operate at the image or bounding-box level, without providing pixel-level lesion masks.

Semantic segmentation extends this line of research by predicting spatially explicit lesion maps. U-Net and its variants remain dominant baselines for plant disease segmentation. Improved U-Net architectures have been developed for potato, tomato, corn and pear leaf diseases [6], [7], [8], while DeepLabv3+ variants have been adapted to rice and other crops [9], [10], [11]. These works mainly differ in backbone choice and attention or multi-scale modules but still follow the general U-Net and DeepLabv3+ families. Representative segmentation models, their target crops and key design elements are summarized in Table 1.

Table 1 Semantic segmentation studies for paddy leaf diseases and related crops.

Model	Variants	Dataset	View type
Deeplabv3+	CBAM-CARAFE-DeepLabv3+	Paddy disease	Zoom-in
	LT-DeepLab	General leaves	Field
	GS-DeepLabV3+	Tea-leaf disease	Field
U-Net	U-Net with Receptive Field Block (RFB) and dual attention	Corn-leaf disease	Zoom-in
	MFBP-UNet	Pear-Leaf disease	Field
	FFAE-UNet	Pear-Leaf disease	Field

The choice of dataset strongly constrains what can be evaluated. Many paddy leaf disease classification works rely on self-collected image sets that are not publicly released while public multi-crop repositories are mainly used for image-level classification and rarely provide dense lesion masks. In the rice domain, the Paddy Doctor dataset has become a popular benchmark because it contains multiple disease classes and field-like imagery. It underpins several classification models in the experimental review of deep learning-based paddy disease models. However, existing work generally treats Paddy Doctor at the image level, reporting accuracy or F1 scores, and does not release pixel-level lesion annotations for lesion segmentation.

Segmentation quality further depends on the annotation strategy. Most U-Net and DeepLabv3+ based plant disease segmentation studies rely on fully supervised pixel-wise masks obtained via manual annotation tools. This yields accurate lesion boundaries but is time-consuming, especially when lesions are small or numerous in cluttered crop-level scenes. Some recent works reduce annotation cost using weakly supervised approaches [12], [13], which infer lesion regions from image-level labels via class activation maps and post-processing on UAV imagery, at the expense of noisier masks [14].

Finally, most existing leaf lesion segmentation models are evaluated in a single-disease or single-dataset setting. Improved U-Net and DeepLabv3+ variants are usually trained and tested for one disease or crop at a time, while multi-class rice disease models such as PaddyNet remain at the image-level classification or detection stage. In the experimental review by Tasfe et al. [4], multiple segmentation models are compared, but experiments are conducted on a small segmentation dataset constructed via automatic mask extraction and do not explicitly benchmark per-disease versus pooled training on a multi-disease rice dataset with dense, manually drawn lesion masks. Therefore, it remains unclear whether separate per-disease segmentation models are necessary for realistic multi-disease rice scenarios, or whether a single pooled model can achieve comparable performance on crop-level imagery. This is an open question we address in this study.

3. Methodology

We use the public Paddy Doctor dataset from Kaggle, which contains RGB images of paddy leaves captured in field-like conditions with several disease types. In this study, we focus on three diseases with clear visible lesions: Bacterial Leaf Blight (BLB), Brown Spot and Hispa (Table 2). All images are crop-level views in which multiple leaves, stems and background elements appear in the frame. All 2,444 images belonging to BLB (804), Brown Spot (936) and Hispa (704) at 480×640 resolution are selected without additional cropping so that the segmentation task reflects the original acquisition conditions. For each image, a binary lesion mask is manually drawn using a polygon-based annotation tool (LabelMe), labelling diseased regions on paddy leaves as foreground and all remaining pixels as background. Example images and masks for the three diseases are illustrated in Figure 1, and these pixel-level masks serve as ground truth for all experiments.

Table 2 Disease types used in dataset

Types of diseases	Full name	Pathogen / Agent
BLB	Bacterial leaf blight	*Xanthomonas oryzae* pv. *oryzae*
Brown Spot	Paddy leaf brown spot	*Bipolaris oryzae*
Hispa	Paddy leaf hispa damage	*Dicladispa armigera* (insect)

a. Original Image b. Disease Masks

Figure 1 Mask for diseases

Before training, image–mask pairs are used at their original composition; images are converted to RGB if needed and resized only when minor deviations from the nominal 480×640 resolution occur. Pixel values are scaled to [0, 1] and normalized channel-wise using the mean and standard deviation computed from the training set. To increase robustness while preserving scene structure, we apply a small set of geometric augmentations during training: each image–mask pair may be randomly mirrored horizontally, mirrored vertically or rotated by 180°, with the same transform applied to both image and mask. We manually hold out 10% of images per disease as test data and use the remaining 90% for training and validation, with final split sizes summarized in Table 3. The same split indices are reused for all architectures and for both the per-disease and pooled regimes.

Table 3 Dataset split configurations.

Disease	Images	Train	Validation	Test
BLB	804	644	80	80
Brown Spot	936	748	94	94
Hispa	704	564	70	70
Pooled	2,444	1,956	244	244

We benchmark three encoder–decoder architectures: U-Net [15], U-Net++ [16] and DeepLabV3 with a ResNet-50 backbone [17], all configured for binary lesion-versus-background segmentation with a single-channel output. U-Net follows the original encoder–decoder design with skip connections, U-Net++ introduces nested skip connections to refine feature fusion, and DeepLabV3 uses a ResNet-50 backbone with atrous convolutions and an Atrous Spatial Pyramid Pooling (ASPP) module for multi-scale context. We use standard configurations without custom attention blocks to keep the comparison focused on core architectural differences.

Each architecture is trained under two distinct regimes: per-disease and pooled. In the per-disease regime, we train three independent binary segmentation models for BLB, Brown Spot and Hispa, respectively, using only the images and masks of the corresponding disease and the fixed training, validation and test splits in Table 3. In the pooled regime, we trained a single binary model per architecture on the union of BLB, Brown Spot and Hispa, treating all lesion pixels as a common foreground class. Evaluation for

the pooled models is still reported both overall and separately per disease by restricting the test metrics to each disease subset. This design allows a direct comparison between disease-specific and pooled training while keeping the data splits and optimization settings identical.

Model performance is evaluated on the held-out test sets using pixel-wise metrics computed from the predicted and ground-truth lesion masks. After training, we load the checkpoint with the best validation Dice score for each architecture and training regime and generate predicted masks for all test images. We report Dice and IoU as the primary metrics, and also record precision, recall, F1-score and pixel accuracy for completeness, although accuracy is less informative under strong background imbalance. Training and evaluation share the same binary formulation where each network outputs a single-channel logits map per image, we optimize a loss defined as the sum of BCEWithLogitsLoss and soft Dice loss applied to the sigmoid outputs, and during testing the probability maps are binarized with a fixed threshold of 0.5. No additional post-processing such as morphological filtering is applied. All models are implemented in PyTorch and trained with the AdamW optimizer, an initial learning rate of 1×10^{-4}, a batch size of 4 and 100 epochs. A fixed random seed is used for PyTorch, CUDA, NumPy and Python's random module, and the same train/validation split indices are reused across all runs. The main training hyperparameters and implementation settings are summarized in Table 4.

Table 4 Training and evaluation settings.

Aspect	Setting / description
Optimizer	AdamW
Initial learning rate	1×10−4
LR scheduler	ReduceLROnPlateau (monitoring validation Dice)
Loss function	BCEWithLogitsLoss + soft Dice loss (equal weight)
Epochs	100
Batch size	4
Input resolution	480 × 640
Framework & hardware	Python 3.9.13, PyTorch 2.5.1+cu121; NVIDIA GeForce RTX 3060 Ti (8 GB VRAM, CUDA 12.6)

4. Results

Table 5 reports the per-image macro Dice, IoU and recall scores for U-Net, U-Net++ and DeepLabV3-ResNet50 on BLB, Brown Spot, Hispa and the pooled setting. Across all diseases, DeepLabV3-ResNet50 consistently achieves the highest Dice and IoU, followed by U-Net, with U-Net++ slightly behind. For BLB, DeepLabV3-ResNet50 reaches a Dice of about 0.78 and an IoU of around 0.66, compared with a Dice of roughly 0.73 and an IoU of about 0.59 for U-Net, and a Dice of approximately 0.71 and an IoU of about 0.57 for U-Net++. A similar pattern is observed for Hispa, where DeepLabV3-ResNet50 attains a Dice of about 0.73 and an IoU of 0.60, together with recall above 0.75, marginally higher than both U-Net variants.

Table 5 Per-image macro Dice, IoU and recall for U-Net, U-Net++ and DeepLabV3-ResNet50 on each disease and the pooled setting.

Dataset	Model	Dice	IoU	Recall
BLB	U-Net	0.7250	0.5941	0.8062
	U-Net++	0.7100	0.5746	0.7546
	DeepLabv3-R50	0.7752	0.6567	0.8214
Brown Spot	U-Net	0.5619	0.4046	0.5822
	U-Net++	0.5396	0.3872	0.5410
	DeepLabv3-R50	0.5990	0.4390	0.5904
Hispa	U-Net	0.7222	0.6004	0.7391
	U-Net++	0.7191	0.5988	0.7214
	DeepLabv3-R50	0.7300	0.6011	0.7541
Pooled	U-Net	0.6597	0.5195	0.6657
	U-Net++	0.6501	0.5127	0.6832
	DeepLabv3-R50	0.6809	0.5418	0.6956

Among the three diseases, Brown Spot is clearly the most challenging for all models. Dice scores for Brown Spot drop to around 0.56 to 0.60 and IoU to 0.39 to 0.44, whereas BLB and Hispa generally lie in the 0.71 to 0.78 Dice and 0.59 to 0.66 IoU range. Recall also decreases for Brown Spot, reflecting the difficulty of capturing its small, scattered lesions, which are easily confused with background texture. In contrast, BLB and Hispa form more elongated, contiguous lesions that are easier to delineate. Overall pixel accuracy remains very high for all settings due to the strong class imbalance toward background, so Dice, IoU and recall provide a more informative comparison of segmentation quality.

To study whether a single model can handle multiple diseases as effectively as disease-specific models, we compare the per-disease scores in Table 5 with the pooled scores for each architecture. For U-Net, the average Dice across the three individual diseases is about 0.67 and decreases slightly to 0.66 in the pooled setting; IoU shows a similar drop of roughly 1 to 2 percentage points. U-Net++ behaves similarly, with average Dice around 0.66 for per-disease training and 0.65 for pooled training. DeepLabV3-ResNet50 shows the strongest per-disease performance, with an average Dice of about 0.70, and its Dice decreases to about 0.68 when trained on all three diseases jointly. Overall, the performance gap between per-disease and pooled training is modest for all three architectures, and the ranking of models remains unchanged, DeepLabV3-ResNet50 performed best, followed by U-Net, with U-Net++ slightly behind.

In addition, we also compute dataset-level Dice and IoU per class where background scores are near 1.0 for all models due to heavy class imbalance, while lesion scores follow the same trends as Table 5, with DeepLabV3-ResNet50 performing best and Brown Spot remaining the most difficult disease.

To complement the accuracy results, we profile the three architectures in terms of model size, computational cost and runtime. Table 6 summarizes the number of parameters, multiply–accumulate operations (MACs), inference time per 640×480 image and peak GPU memory usage on the RTX 3060 Ti. Although DeepLabV3-ResNet50 has the largest parameter count (39.6 M), it is also the fastest at inference (about 33 ms per image) and uses the least GPU memory (290 MB). U-Net has slightly fewer parameters but the highest MAC count and a slower inference time (42 ms), while U-Net++ is the lightest in parameters yet somewhat slower still (48 ms) and consumes similar memory to U-Net. Taken together with the segmentation metrics, these results indicate that DeepLabV3-ResNet50 offers the best balance between segmentation quality and computational efficiency for deployment on commodity GPUs.

Table 6 Model complexity and runtime characteristics for a 1×3×640×480 input.

Model	Params (M)	MACs (G)	Inference time (ms/img)	Peak VRAM
U-Net	31.038	256.583	41.81 ms/img	835.7 MB
U-Net++	9.163	163.610	47.85 ms/img	795.4 MB
DeepLabv3-R50	39.634	192.326	33.12 ms/img	290.0 MB

Figure 2 illustrates representative segmentation examples for BLB, Brown Spot and Hispa, comparing ground-truth lesion masks with predictions from U-Net, U-Net++ and DeepLabV3-ResNet50. For BLB, all three models are generally able to recover the long, stripe-like lesions along the leaf, but U-Net and U-Net++ sometimes produce thicker masks or small gaps along the lesion centre, whereas DeepLabV3-ResNet50 tracks the lesion boundaries more tightly and suppresses isolated false positives in the surrounding foliage. For Brown Spot, the limitations of all models become more apparent. In U-Net and U-Net++ often miss faint spots or fragment them, while DeepLabV3-ResNet50 captures more of the true spots but still underestimates very small or low-contrast lesions. For Hispa, which produces elongated feeding scars, all three models perform better than on Brown Spot, though U-Net and U-Net++ tend to over-segment edges or drop thin branches, while DeepLabV3-ResNet50 produces cleaner, more contiguous masks. Across all diseases, the typical failure modes are background leakage in dense crop-level scenes, missed thin or low-contrast lesions, and fragmentation of small spots, particularly for Brown Spot, aligning with the quantitative trends in Table 5 and Table 6.

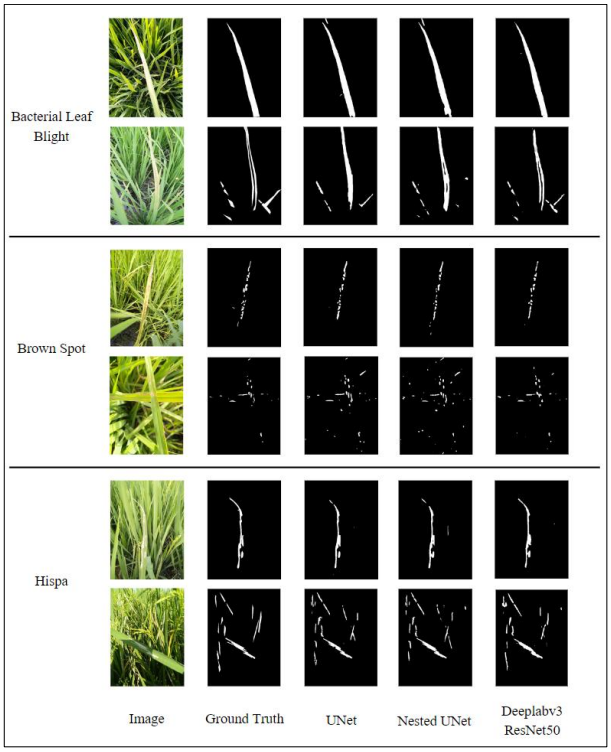

Figure 2 Qualitative segmentation results for three paddy leaf diseases.

5. Conclusion

This work presented a unified benchmark for pixel-level segmentation of three paddy leaf diseases which are BLB, Brown Spot and Hispa, on crop-level images from the public New Paddy Doctor dataset. We manually annotated 2,444 images with lesion masks and evaluated U-Net, U-Net++, and DeepLabV3-ResNet50 under a single, reproducible pipeline, considering both per-disease and pooled training regimes. Across all settings, DeepLabV3-ResNet50 consistently achieved the highest Dice and IoU scores while also offering the best balance between inference speed and memory usage on a commodity GPU.

Our experiments show that Brown Spot is substantially harder to segment than BLB and Hispa, due to its small, scattered lesions and cluttered backgrounds, and this difficulty persists across architectures and training regimes. Training a single pooled model on all three diseases leads to only a modest reduction in Dice or IoU compared to separate per-disease models, while preserving the same ranking among architectures. This suggests that a pooled DeepLabV3-ResNet50 model is a practical choice when deployment simplicity and efficiency are important, whereas per-disease models offer a small but consistent accuracy gain.

Future work will extend this benchmark in several directions, including incorporating class-balanced or focal losses to better handle small lesions, and evaluating cross-dataset generalization to other paddy fields and acquisition conditions. Although our experiments focus on ground-

level RGB images, lightweight segmentation and detection models have already been deployed on UAV-based rice imagery [18] and on edge devices for real-time leaf disease monitoring in other crops [19]. Building on these advances, we plan to compress and adapt the best-performing architecture from this study for on-board or handheld deployment in realistic field settings.

Acknowledgements

The author would like to thank Kurk Akakpo from CESI Graduate School of Engineering, France and the lab members of the Applied Control and Robotics Laboratory, Universiti Malaya, for their support and constructive discussions during this work. The experiments in this study were conducted using the public "New Paddy Doctor" image dataset made available on Kaggle, and the authors gratefully acknowledge its contributors for releasing the data to the research community.

References

1. Malathi V, Gopinath MP, Kumar M, Bhushan S, Jayaprakash S. Enhancing the Paddy Disease Classification by Using Cross-Validation Strategy for Artificial Neural Network over Baseline Classifiers. Journal of Sensors. 2023;2023(1).
2. Li Y, Chen X, Yin L, Hu Y. Deep Learning-Based Methods for Multi-Class Rice Disease Detection Using Plant Images. Agronomy. 2024;14(9): 1879.
3. Petchiammal, Kiruba B, Murugan, Arjunan P. Paddy Doctor: A Visual Image Dataset for Automated Paddy Disease Classification and Benchmarking. Proceedings of the 6th Joint International Conference on Data Science & Management of Data (10th ACM IKDD CODS and 28th COMAD). New York, NY, USA: ACM; 2023. pp. 203–207.
4. Tasfe M, Nivrito A, Al Machot F, Ullah M, Cheikh FA, Ullah H. Deep Learning-Based Models for Paddy Disease Classification and Segmentation: An Experimental Review. IEEE Access. 2025;13: 152962–152996.
5. A P, D M, S BK. PaddyNet: An Improved Deep Convolutional Neural Network for Automated Disease Identification on Visual Paddy Leaf Images. International Journal of Advanced Computer Science and Applications. 2023;14(6).
6. Mu Y, Li K, Sun Y, Bao Y. Semantic Segmentation of Corn Leaf Blotch Disease Images Based on U-Net Integrated with RFB Structure and Dual Attention Mechanism. Agronomy. 2024;14(11): 2652.
7. Wang H, Ding J, He S, Feng C, Zhang C, Fan G, et al. MFBP-UNet: A Network for Pear Leaf Disease Segmentation in Natural Agricultural Environments. Plants. 2023;12(18): 3209.
8. Wang W, Ding J, Shu X, Xu W, Wu Y. FFAE-UNet: An Efficient Pear Leaf Disease Segmentation Network Based on U-Shaped Architecture. Sensors. 2025;25(6): 1751.
9. Zeng W, He M. Rice disease segmentation method based on CBAM-CARAFE-DeepLabv3+. Crop Protection. 2024;180: 106665.
10. Yang T, Wei J, Xiao Y, Wang S, Tan J, Niu Y, et al. LT-DeepLab: an improved DeepLabV3+ cross-scale segmentation algorithm for Zanthoxylum bungeanum Maxim leaf-trunk diseases in real-world environments. Frontiers in Plant Science. 2024;15.
11. Zhou H, Peng Y, Zhang R, He Y, Li L, Xiao W. GS-DeepLabV3+: A mountain tea disease segmentation network based on improved shuffle attention and gated multidimensional feature extraction. Crop Protection. 2024;183: 106762.
12. Kakinoki, K., Katayama, T., Kita, Y., Yamaba, H., Aburada, K., & Okazaki, N. (2024). ASLA: Automatic Segmentation and Labeling by Deep Learning for Document Pictures. Journal of Robotics, Networking and Artificial Life, 10(4), 362-367.
13. Ru, Y., Du, H., Li, S., & Chang, H. (2017). Action recognition based on binocular vision. Journal of Robotics, Networking and Artificial Life, 4(1), 5-9.
14. Chen S, Zhang K, Wu S, Tang Z, Zhao Y, Sun Y, et al. A Weakly Supervised Approach for Disease Segmentation of Maize Northern Leaf Blight from UAV Images. Drones. 2023;7(3): 173.
15. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. 2015. pp. 234–241.
16. Zhou Z, Rahman Siddiquee MM, Tajbakhsh N, Liang J. UNet++: A Nested U-Net Architecture for Medical Image Segmentation. 2018. pp. 3–11.
17. Chen L-C, Zhu Y, Papandreou G, Schroff F, Adam H. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. 2018.
18. Zhang P, Sun X, Zhang D, Yang Y, Wang Z. Lightweight Deep Learning Models for High-Precision Rice Seedling Segmentation from UAV-Based Multispectral Images. Plant Phenomics. 2023;5: 0123.
19. Karim MdJ, Goni MdOF, Nahiduzzaman Md, Ahsan M, Haider J, Kowalski M. Enhancing agriculture through real-time grape leaf disease classification via an edge device with a lightweight CNN architecture and Grad-CAM. Scientific Reports. 2024;14(1): 16022.

Authors Introduction

Mr. Muhammad Amirul Aiman Asri

He received the B.Eng. in Electrical Engineering from Universiti Malaya, Malaysia, in 2020, where he is currently completing the M.Eng.Sc. in Applied Control and Robotics (ACR Lab). His research interests include computer vision and deep learning for agricultural and biomedical imaging, with a focus on image quality assessment and semantic segmentation.

Mr. Wenjunliang Zhang

He received his Master degree in Engineering in 2024 from the Faculty of Engineering, Newcastle University in United Kingdom. He is currently a PhD student in Universiti Malaya, Malaysia. His current research field include multi-agent collaborative control and precision agriculture.

Mr. Afdhal Haziq Noramly

He is currently pursuing his B.Eng. in Electrical Engineering at Universiti Malaya. His research interests include AI-based detection systems, robotics, and automation technologies. He is also active in a robotics NGO, where he contributes to STEM outreach initiatives.

Associate Prof. Ir. Dr. Norrima Mokhtar

She earned a Bachelor of Engineering (B.Eng) in Telecommunication Engineering from Universiti Malaya in 2000. She was awarded the Panasonic Scholarship to pursue her Master of Engineering (M.Eng) at Oita University, Japan (2003-2006). Additionally, she received the SLAB/SLAI scholarship to complete her Ph.D. in Electrical Engineering at Universiti Malaya (2008-2012). Between 2000 and 2002, she worked as a Telecommunication Engineer at Echobroadband Sdn. Bhd., where she contributed to upgrading cable TV networks to hybrid fiber coaxial networks in Köln, Germany. Since 2003, she has built a distinguished academic career at Universiti Malaya, where she currently serves as an Associate Professor.

Dr. Raza Ali

He received the B.S. degree in telecommunication engineering from Balochistan University of Information Technology, Engineering and Management Sciences (BUITEMS), Quetta, Pakistan, the M.S. degree in electrical engineering (communication) from UET, Lahore, and the Ph.D. degree from the University of Malaya, Malaysia, in 2022. During the Ph.D. studies, he was associated with the VIP Laboratory, University of Malaya. He is currently an Assistant Professor with the Faculty of Information and Communication Technology, BUITEMS. His research interests include signal processing, computer vision, machine learning, and deep learning.

Professor Takao Ito

He received his M.S., and Ph.D. He is Professor of Management of Technology (MoT) in Graduate School of Engineering at Hiroshima University. He is serving concurrently as Professor of Harbin Institute of Technology (Weihai) China. He has published numerous papers in referred journals and proceedings, particularly in the area of management science, and computer science. He has published more than eight academic books including a book on Network Organizations and Information (Japanese Edition). His current research interests include automata theory, artificial intelligence, systems control, quantitative analysis of inter-firm relationships using graph theory, and engineering approach of organizational structures using complex systems theory.

Dr. M. Aziz Muslim

He received his Bachelor and Master degrees in Electrical Engineering from Institut Teknologi Sepuluh Nopember (ITS), Indonesia, and later earned his Ph.D. in Brain Science and Engineering from the Kyushu Institute of Technology, Japan. His research interests include robotics, artificial intelligence, deep learning-based control systems, and smart navigation technologies. He is currently a faculty member of the Electrical Engineering Department, Faculty of Engineering, Universitas Brawijaya, Indonesia.

Professor Siti Sendari

She received her Bachelor's degree in Electronic Engineering from Universitas Brawijaya, Malang, Indonesia, in 1996. She obtained her M.S. degree in Computer Engineering and Informatics from Universitas Gadjah Mada, Yogyakarta, in 2005, and her Ph.D. degree in System Engineering from Waseda University, Japan, in 2013. She is currently a Professor in the Department of Electrical and Informatics Engineering, Universitas Negeri Malang, with research interests in Human–Machine Interaction and Reinforcement Learning.

Dr. Pringgo Widyo Laksono

He is a researcher and lecturer at Universitas Sebelas Maret (UNS), Indonesia. He earned his Doctor of Engineering (Dr.Eng) degree in Intelligent Machine from Gifu University, Japan. He specializes in industrial engineering, focusing on data analytics, smart systems in industrial applications, intelligent machines & robotics, smart manufacturing systems, human-machine automation, bio-signal control, design working system, management science in industry, and renewable energy.

Dr. Tsutomu Ito

Dr. Tsutomu Ito is Assistant Professor of the Department of Business Administration at National Institute of Technology, Ube College, Japan. He has published many papers in refereed journals and proceedings, particularly in the area of industrial management, and computer science. His current research interests include internet of things (IoT), mechanical engineering, artificial intelligence (AI), automata theory, quantitative analysis of Japanese Keiretsu. Dr. Ito earned his doctor degree of Engineering from Hiroshima University, Japan in 2018.