A hybrid atrous CNN-SVM model for rice leaf disease classification

H. Huy Nguyen[1], Quang-Thinh Bui[2]
[1]Faculty of Information Technology, HUTECH University, Ho Chi Minh City, Vietnam
[2]Office of Scientific Research, Technology Management and International Cooperation, Tien Giang University, Dong Thap, Vietnam
nguyenhuuhuy6262@gmail.com, buiquangthinh@tgu.edu.vn

Abstract

Rice leaf diseases can significantly impact crop yields, so identifying them early and accurately is crucial for maintaining control. In this work, we propose a hybrid deep learning approach called Atrous CNN-SVM, which leverages a pre-trained VGG19 model and enhances it with atrous (or dilated) convolutional layers to extract features across different scales. This approach is then paired with a Support Vector Machine for reliable classification. We put the model through its paces using a practical dataset gathered from Dong Thap, Vietnam, and double-checked it against an external dataset. The experiments revealed that Atrous CNN-SVM achieved an accuracy of 86.49% when trained on 90% of the local data, outperforming plain CNN-SVM, MobileNet, and classic models that rely on hand-engineered features. Overall, this highlights the real value of combining atrous convolutions with SVM for automatic rice disease detection, and it underscores the importance of incorporating location-specific data for precision farming.

Received	02/07/2025
Accepted	06/09/2025
Published	28/09/2025

Keywords
Rice leaf disease; CNN-SVM; Atrous convolution; Transfer learning; Smart agriculture

1. Introduction

The health of rice leaves is particularly important for the total productivity of the crop, as leaves are where photosynthesis occurs, which directly affects plant development and grain yield. Among the many threats to rice production, leaf diseases are particularly alarming due to their prevalence and severe impact. These diseases, which are usually caused by fungi or bacteria, can significantly reduce the effectiveness of photosynthesis, resulting in lower yields and crop losses. In places like Vietnam, where rice is grown, identifying leaf diseases early and accurately is crucial for maintaining farm productivity and ensuring sufficient food supplies.

Traditional methods for detecting rice leaf diseases often depend on farmers or experts checking. This method is available small-scale applications, but it has several disadvantages: it requires considerable effort and time, highly subjective, and challenging to implement effectively on a large scale [1], [2]. To address these limitations, early computer techniques were developed to accelerate the process of finding diseases. These techniques employed manually generated features, including color, form, and texture, in conjunction with traditional machine learning models like Support Vector Machines (SVM) [3], K-Nearest Neighbors (KNN) [4], or Naive Baye [5].

In the last several years, deep learning, particularly Convolutional Neural Networks (CNNs), has become a powerful technique for diagnosing plant diseases from images. This is because they can automatically extract characteristics with little manual preparation [6]. However, standard CNNs often face challenges in capturing broad contextual information, especially when disease symptoms appear at multiple scales [7]. To overcome constraints in spatial resolution and computing efficiency, researchers have investigated hybrid CNN-SVM models, wherein the CNN functions as a feature extractor and the SVM operates as a robust classifier, demonstrating notable efficacy for small or imbalanced datasets [8].

To tackle these issues, particularly the challenge of recognizing signs of illness of varying sizes (from small spots in the early stages to large lesions in the later stages) without sacrificing important spatial information, this work combines tried-and-true methods in a smart way. We adopt and refine a hybrid architecture that combines the power of Atrous convolution within a pre-trained VGG19 network, coupled with an SVM classifier. The reason for this choice is that Atrous convolution allows the model to widen its receptive field, capturing context at different scales without using pooling procedures that lower resolution. This is especially useful for diagnosing

https://doi.org/10.55401/hyzp8z13

diseases, because keeping the small details of symptoms is quite important.

The main contributions of this study are as follows:

(1) We present the application and evaluate the effectiveness of a hybrid architecture that integrates Atrous convolutional layers into a VGG19 network with an SVM classifier, demonstrating that this combination is exceptionally suited for extracting multi-scale features in the task of rice leaf disease detection on local data.

(2) We conduct extensive empirical validation on a real-world dataset collected locally (Dong Thap, Vietnam) and an external dataset, showing that the proposed Atrous CNN-SVM model consistently outperforms conventional CNN-SVM models, MobileNet, and traditional handcrafted-feature models.

(3) We conduct a comprehensive analysis of data localization, elucidating the vital significance of the local data fraction during training and underscoring that domain-specific datasets can substantially improve model performance in practical precision agricultural applications.

The remainder of this paper is organized as follows: Section 2 presents the theoretical framework and pertinent literature, covering conventional machine learning methodologies, current CNN-SVM hybrid models, and the fundamentals of Atrous convolution. In Section 3, we discuss the describe model architecture, which focuses on utilizing CNN with Atrous convolution to extract features and SVM for classification. Section 4 provides detailed information about the dataset, preprocessing pipeline, experimental setting, evaluation metrics, and outcomes. Finally, Section 5 concludes the work by summarizing the main points and suggesting areas for future investigation.

2. Background Knowledge

2.1. Traditional and Classical Machine Learning Methods

Initial efforts in plant disease detection were predominantly based on manual observation by experts or farmers. However, this approach is time-consuming, labor-intensive, and subjective [9]. These limitations motivate the application of image processing methods such as thresholding, colour space analysis, and handcrafted feature based on texture and morphology. After that, these attributes were utilized in traditional machine learning approaches, such as SVM and KNN, for classification [10]. Concurrently, numerous studies have experimented with statistical methods, such as Naive Bayes, to automate the diagnostic process [11]. While laying an important basis for automated systems, these methods often struggled to adapt to the inherent diversity of disease symptoms, varying light conditions, and complex image backgrounds, largely due to the high demand for deep domain expertise in designing suitable features.

2.2 The Hybrid CNN-SVM Model

To merge the strong automatic feature extraction capabilities of CNNs with the effective classification performance of SVMs, particularly in cases with limited data or high-dimensional feature, a number of studies have proposed hybrid CNN-SVM models [12]. In these models, a CNN (typically using standard convolutional layers) extracts the features, and its output vector becomes the input for the SVM classifier. However, a majority of current CNN-SVM models still rely on standard convolutional layers for feature extraction, which may lead to inheriting the limitations of traditional convolutions in capturing a wide-ranging context [13].

2.3. Atrous Convolution: Expanding Context While Preserving Detail

In the task of rice leaf disease recognition, CNNs that use standard convolutions face an inherent challenge. These convolutional layers are great at picking out small details, their receptive fields expand quite slowly with network depth. To capture the broader context—which is essential for distinguishing large-scale disease symptoms—models typically resort to pooling layers or strided convolutions. However, these methods lower the spatial resolution of the feature maps, which increases the risk of losing fine-grained details, such as disease spots in their early stages.

To resolve this trade-off, this study employs Atrous Convolution (also known as dilated convolution), a technique proven effective in various domains that require a deep contextual understanding, such as semantic image segmentation [14] and audio processing [15]. By adding regulated gaps (specified by a "dilation rate") between the kernel weights, it significantly expands the receptive field without increasing the number of parameters or reducing resolution.

This allows the model to simultaneously analyze the overall context of the leaf and accurately identify the smallest disease symptoms—a critical requirement for early and effective diagnosis.

3. Proposed methodology

3.1. Architectural Overview

The proposed Atrous CNN-SVM model has two main parts working together in order: a feature extractor based on a CNN with Atrous convolution,and a classifier using SVM. When an image of a rice leaf is input, the model first processes it through the Atrous convolutional network, which extracts a concise and information-rich feature vector that clearly reflects the crucial visual characteristics related to disease identification. This feature vector then goes into a trained SVM classifier, which makes the final decision, determine the specific type of disease or to confirm that the leaf is healthy. A general overview of the model's architecture is illustrated in Figure 1.

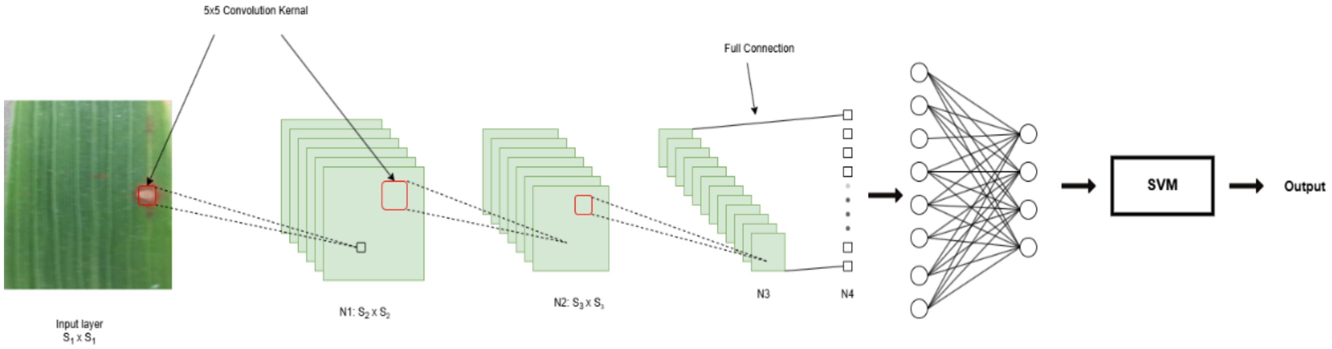


Figure 1. The proposed architecture of the AtrousCNN-SVM model.

3.2. Atrous CNN Feature Extractor

The main part of the model is the feature extractor, which turns the input image into a useful vector of features. The VGG19 was chosen as the backbone for this model for its strong performance. Although more modern architectures, such as ResNet or EfficientNet, exist, the sequential and relatively simple structure of VGG19 provides an ideal baseline for studying the impact of inserting Atrous Convolutional layers in a clear manner. This method helps to separate and test the performance of Atrous convolution, eliminating the confusing effects of complicated features such as skip connections in ResNet. This makes it easier to see what each part does. The model uses the ImageNet dataset to get the pre-trained weights of VGG19. This leverages the advantages of transfer learning, enabling the model to harness features learned from a large-scale dataset, thereby improving performance and accelerating convergence.

To address the problem of identifying disease features at different scales, we incorporate an Atrous Spatial Pyramid Pooling (ASPP) block into the architecture. This block is made to capture contextual information at multiple scales efficiently.

Specifically, this ASPP block is placed after the final convolutional layer of the VGG19 architecture, serving as a powerful multi-scale feature extractor. The ASPP block comprises several parallel components to capture information at different levels. It begins with a 1×1 convolutional layer to reduce the feature depth. This is followed by three 3×3 Atrous convolutional layers with dilation rates set to $r = 6$, $r = 12$, and $r = 18$, respectively. We selected these rates so that the model could identify illness spots and large lesion regions at medium, large, and very large scales. These layers enable the model to gather context at both medium and large scales without increasing the number of parameters or compromising resolution. Lastly, an Image Pooling block performs Global Average Pooling on the feature map, followed by a 1×1 convolution and upscaling to the original size to capture global information.

The outputs from these parallel branches are combined and sent through a final 1×1 convolutional layer to merge the multi-scale features. This approach lets the model to simultaneously analyze the image at various levels of detail, from fine-grained local features to the global context of the leaf, creating an extremely rich and powerful feature representation before it is passed to the classifier. An example of modifying a VGG19 convolutional block to an Atrous convolutional form is illustrated in Figure 2.

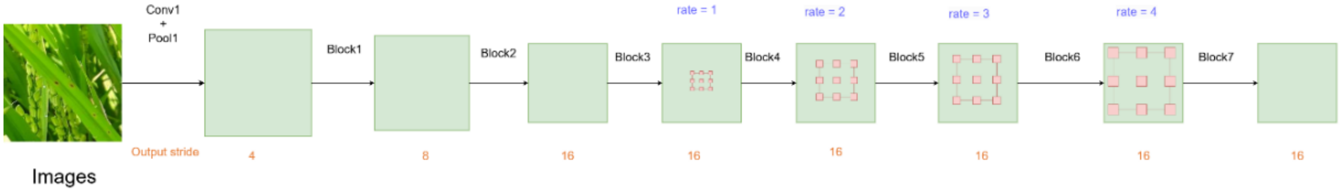


Figure 2. Detailed architecture of the Atrous CNN feature extractor based on a modified VGG19.

3.3. SVM for Multi-Class Classification

SVM is a popular supervised learning method praised for its solid theory and strong results, especially with many features or limited training data. In essence, SVM is a binary classifier to find an optimal hyperplane that separates two classes with the largest possible margin in the feature space.

In this study, the input to the SVM classifier is a feature vector generated by the ASPP block. Specifically, the output feature map from the ASPP is passed through a Global Average Pooling layer to create a fixed-length feature vector of 256 dimensions. This vector, which has condensed multi-scale information, serves as the final representation of the input image. The Radial Basis Function (RBF) kernel was chosen. The main reason is that the RBF kernel can effectively handle complicated, non-linear interactions between features well, which is common in image recognition. It maps the original feature into a higher-dimensional space where the disease classes can be separated linearly, thereby enhancing the model's generalization ability.

However, the rice leaf disease classification problem is a multi-class task. To extend SVM from binary to multi-class classification, strategies such as "one-vs-one" (OVO) or "one-vs-rest" (OVR) are commonly used. In this research, we employ the OVO strategy to support multi-class classification, as implemented in the scikit-learn SVC library. According to the OVO strategy, for a problem with K classes, a total of $K * (K-1)/2$ binary SVM classifiers are constructed, with each classifier trained to distinguish between a specific pair of classes. During the prediction process, each binary classifier casts a vote for its chosen class, and the new input is assigned to the class that receives the most votes.

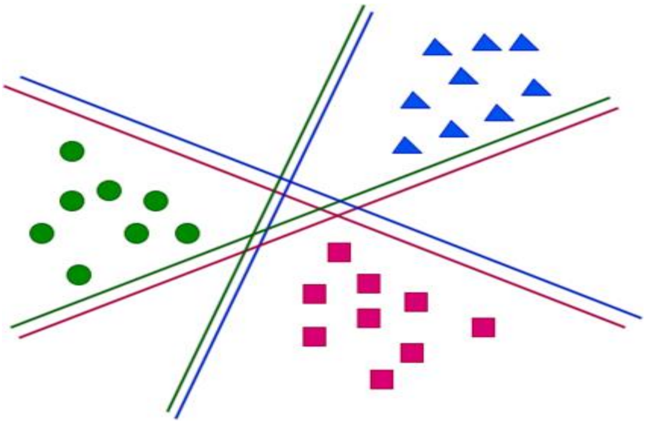


Figure 3. The one-vs-one SVM architecture.

For the RBF kernel, the decision function for a new input x_n in a specific binary pair classification is typically:

$$f(x) = sign\left(\sum_{i=1}^{N} \alpha_i y_i K(x_i, x_n) + b\right),$$

where $K(x_i, x_n) = \exp\left(-\gamma * ||x_i - x_n||^2\right)$ is the RBF kernel function, x_i are the support vectors, y_i are the corresponding labels , α_i are the learned weights, b is the bias term and γ is a kernel parameter that defines the influence of a single training example. The final multi-class classification decision is based on aggregating the votes from all binary OVO classifiers.

4. Experiments

4.1. Datasets and Configuration

4.1.1. Data

This study was conducted on an original dataset of 1,314 diseased rice leaf images collected in Dong Thap, Vietnam, featuring four common diseases: bacterial blight, Blast, Brown spot, and Tungro. This homegrown collection gives a solid, real-world basis for testing disease classification models.

To test how well the model generalizes and handles diverse data, we mixed in an external set “Rice Leaf Disease Image Samples”[1] from Mendeley Data, put together by Prabira Kumar Sethy and released on July 18th, 2020 with 5932 images of the same diseases. Its broader variety beefs up the training data and helps evaluate across different sources.

As shown in Table 1, we conducted experiments on training sets with various blends, including 90% from Dong Thap and 10% external data. The total image count varies with each mix, limited by the available content. For unbiased results, we set aside some Dong Thap images as a pure test set, kept out of all training.

4.1.2. Preprocessing

All images from Dong Thap or the external Mendeley dataset underwent a standard pipeline, which differed for training versus test/validation sets. For training, we boosted generalization with augmentations like random ±30 degrees, 300×300 random crops (80-100% scale, aspect tweaks), and 50% chance horizontal/vertical flips, applied on-the-fly during loading.

Test and validation images were resized to 256 pixels on the shorter side, then center-cropped to 224×224, without augmentations. Everything was converted to PyTorch tensors and normalized using ImageNet's mean and standard deviation for consistent scaling and training stability.

We held out a pure Dong Thap test set upfront, used only for final eval post-training. The rest is split 80/20 into training/validation for reproducibility, with validation guiding training, hyperparameter tuning, and early stopping. The results in Table 1 are based on this untouched Dong Thap test set.

---

[1] https://www.kaggle.com/datasets/nirmalsankalana/rice-leaf-disease-image

4.1.3. Environment

The experiments were conducted on a server equipped with an NVIDIA Tesla T4 GPU (16 GB VRAM) and 32 GB of RAM. The software environment included Python 3.10, PyTorch 2.1, Torchvision 0.18, and CUDA 12.2.

4.1.4. Metrics

The Atrous CNN , built upon the modified VGG19 architecture, was fine-tuned with optimization strategies to maximize performance and enhance model stability. The training process utilized the AdamW optimizer with a weight decay of 0.01 to mitigate overfitting. Different learning rates were applied to different layers: the pre-trained layers used a lower learning rate (e.g., 1×10^{-5}) to gently adjust the learned weights, while newly added layers were trained with a higher rate (e.g., 1×10^{-3}) to adapt more quickly to the specific data, helping the model adapt effectively without disrupting the previously learned features.

To handle data imbalance, class weights were calculated based on the inverse frequency of each class and incorporated into the Cross-Entropy Loss function. Additionally, label smoothing was employed to reduce the model's overconfidence and improve its generalization ability.

The model's performance was evaluated based on accuracy the primary metric in this study, defined as the proportion of correctly classified samples in the test set, applied to the various data mixing configurations as presented in Table 1.

Table 1. Accuracy of the models with different data mixing ratios

Model	Mixing ratios(%)						
	90	**85**	**80**	**75**	**70**	**65**	**60**
CNN	76.06	80.00	71.83	76.06	76.06	77.78	69.01
SVM	80.28	81.69	73.61	77.78	73.61	71.83	76.22
MOBILENET	83.78	83.64	84.44	83.92	81.48	76.39	79.37
GLCM-SVM	67.57	72.73	58.33	57.78	60.19	53.97	55.24
ALXNET-SVM	81.08	80.29	80.58	81.27	77.26	75.69	73.13
LBP-SVM	62.16	61.82	62.50	58.89	57.41	55.56	55.94
CNN-SVM	83.78	81.82	83.10	71.83	79.63	76.98	73.24
Atrous CNN-SVM	**86.49**	**83.10**	**86.11**	**84.44**	**84.26**	**78.57**	**78.32**

4.2. Experimental Results

4.2.1. Evaluation of Results

The experimental results confirmed the effectiveness of the proposed Atrous CNN-SVM model in classifying rice leaf diseases. The model consistently demonstrated superiority over comparative methods across multiple data configurations, achieving a peak accuracy of 86.49% when trained with 90% local data from Dong Thap. This impressive performance validates the study's initial hypothesis: integrating Atrous convolutional layers significantly enhanced the multi-scale feature extraction capability. By effectively capturing both small disease spots in early stages and larger lesion areas without losing spatial information, the model generated richer and more accurate feature representations, leading to the superior performance of the SVM classifier in the final step.

A comparison analysis with different models further underscores the benefits of the Atrous CNN-SVM architecture. The typical CNN-SVM model with regular convolutional layers got results that were mostly steady, but it was still beaten most of the time. This shows that adding Atrous convolution helps create more useful features for showing disease symptoms. Likewise, while MobileNet and other deep learning models such as ALXNET-SVM produced competitive results in specific setups, they were unable to sustain better performance as consistently as the suggested hybrid model. On the other hand, models that used hand-crafted features, including GLCM-SVM and LBP-SVM, got much worse results. This illustrates how typical feature engineering can't capture the wide range of complex visual patterns associated with leaf diseases.

Another crucial observation relates to the data mixing ratio. When the model was trained with a high proportion (80-90%) of local data, performance reached optimal or near-optimal levels. This suggests that domain-specific information—particularly from the geographical region where the model will be deployed—plays an essential role in improving accuracy. Conversely, incorporating too much external data tended to decrease performance. The characteristics of the SVM classifier can explain this decline. External data (from Kaggle), despite having the same disease labels, may contain variations in lighting conditions, rice varieties, and camera angles that differ from the local data. This discrepancy can be perceived as a form of "noise" or "domain shift" by the model. The SVM classifier is inherently sensitive to noisy data points, as they can directly influence the position of the support vectors and the separating hyperplane. Especially in the OVO multi-class strategy, if noise appears between a specific pair of classes, it will degrade the performance of the corresponding binary classifier, leading to errors in the "voting" process and reducing the overall accuracy of the model.

A noteworthy point is that the MobileNet model achieved its best results with 80% local data, suggesting that the optimal data ratio may depend on the specific model architecture. This underscores the need for designing a suitable data strategy, especially in practical application scenarios.

Table 2. Detailed results of the leading models on the 90% local data set

Model	Accuracy	Precision	Recall	F1-Score
CNN	76.06	76.62	75.52	76.05
SVM	80.28	80.59	80.14	80.28
MOBILENET	83.78	84.15	83.52	83.78
GLCM-SVM	67.57	67.95	67.02	67.57
ALXNET-SVM	81.08	81.21	81.18	81.08
LBP-SVM	62.16	62.43	61.85	62.15
CNN-SVM	83.78	83.61	83.79	83.78
Atrous CNN-SVM	**86.49**	**86.48**	**86.50**	**86.49**

To gain a deeper insight into the classification performance of the best model, we conducted an analysis of the confusion matrix for the Atrous CNN-SVM, as shown in Figure 4. The results indicate that the model achieved very high performance in classifying the two classes, Brown Spot and Tungro, with very few misclassifications. However, some significant confusion was noted between the Bacterial Blight and Blast classes.

The cause of this confusion can be explained from both a visual and an agricultural perspective. In agriculture, Blast disease in its late stages often presents symptoms such as large, dry, grayish-white lesions, that are very similar to the appearance of Bacterial Blight. During data collection, some images of Blast disease were captured in advanced stages of development, resulting in a dataset containing samples with high visual similarity between the two disease classes, which makes it difficult for the model to distinguish between them.

This finding suggests a potential avenue for improvement in future research, which is to focus on extracting more subtle distinguishing features or supplementing the dataset for class pairs that are easily confused to enhance the model's classification capability.

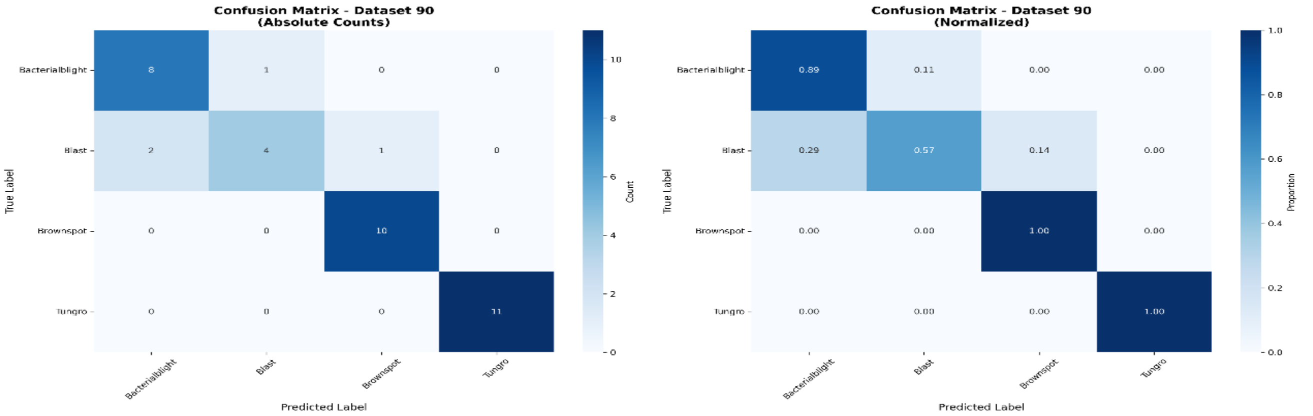


Figure 4. Confusion matrix of the Atrous CNN-SVM model on the 90% local data set.

In Figure 5, we used Grad-CAM to generate heatmaps that highlight the most important regions in the images. The heatmaps show that the Atrous CNN-SVM model learned to accurately focus on areas with disease symptoms (spots, lesions, discoloration) on the rice leaves, rather than on the background or healthy parts of the leaves. This demonstrates that the model not only achieves high accuracy but also learns relevant visual features, which increases its reliability.

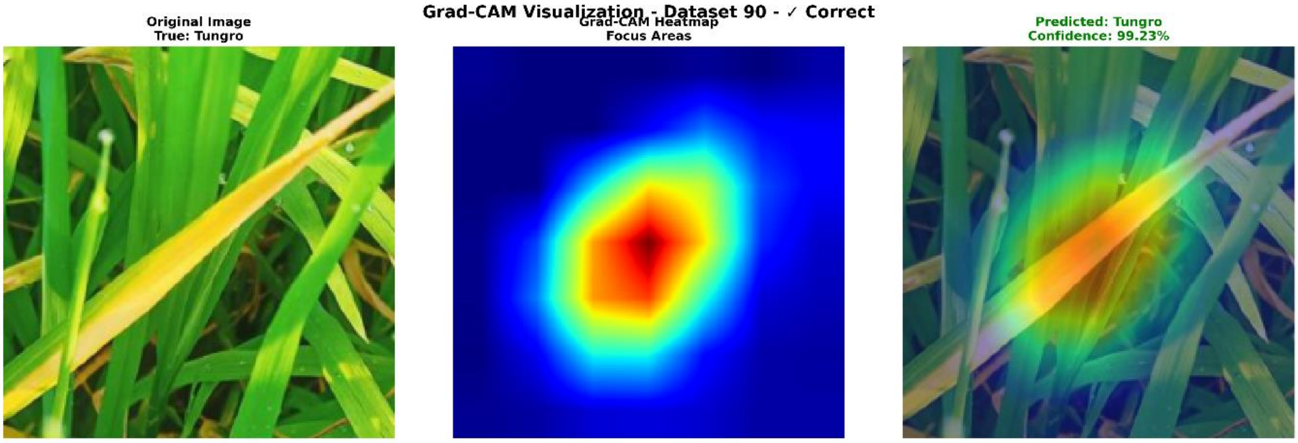
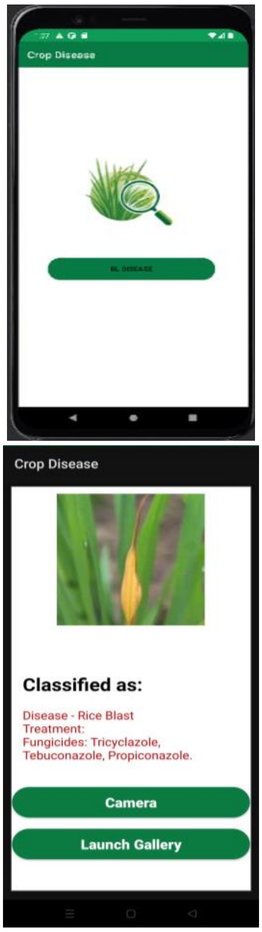


Figure 5. Grad-CAM visualization on the 90% local data set.

In summary, the proposed Atrous CNN-SVM model demonstrates architectural superiority in capturing disease-related features and adapts well to real-world data conditions. This study emphasizes the significance of local data quality and the seamless integration of model architecture and data sources in developing reliable diagnostic systems for precision agriculture.

4.2.2. Mobile Applications



Classified as:

Disease - Rice Blast
Treatment:
Fungicides: Tricyclazole, Tebuconazole, Propiconazole.

Figure 6. Application interface and recognition result.

To illustrate the practical application potential of the research, a mobile application named "Crop Disease" was developed as a proof-of-concept, as shown in Figure 6. However, deploying a powerful and highly accurate model, such as Atrous CNN-SVM, directly onto every mobile device presents a real-world challenge of balancing performance with system resources. While being the top choice for accuracy, the VGG19 base architecture and the Atrous convolutional layers demand significant processing power and storage, which can overload or degrade the user experience on devices with limited specifications.

Therefore, a flexible deployment strategy was devised to address this issue. For high-spec devices or when the application can connect to a processing server (cloud-based), the Atrous CNN-SVM model would be the optimal choice to deliver the highest diagnostic accuracy. Conversely, to ensure the application can operate independently (offline) and be compatible with the majority of devices with limited resources, a more lightweight architecture, such as MobileNetV2 is prioritized. The version of "Crop Disease" presented in this paper was built with the second approach in mind, prioritizing broad accessibility and immediate operation. This choice does not diminish the value of the core model but rather confirms a practical deployment roadmap, where the application serves as a ready platform to integrate higher-performance models as conditions permit.

5. Conclusion

This research has effectively introduced and assessed a hybrid deep learning model, Atrous CNN-SVM, for the classification of rice leaf diseases. The results have shown that incorporating Atrous convolutional layers into a CNN architecture and combining them with an SVM classifier is a highly effective solution, enabling the extraction of disease features at multiple scales without compromising spatial resolution. Experiments on the dataset collected in Dong Thap, Vietnam, showed that the proposed model achieves superior accuracy (up to 86.49%) compared to other baseline methods. The study also demonstrates the importance of local data in improving performance.

Although deployment on practical mobile devices requires a balance between accuracy and resources, the Atrous CNN-SVM model has affirmed its potential as a reliable, automated disease diagnosis tool. Future work will focus on optimizing the model architecture (using techniques such as quantization and pruning) to reduce computational requirements. The ultimate goal is to close the gap between research and real-world use, allowing the most accurate model may be added to be integrated into mobile tools, thereby contributing to the advancement of sustainable agriculture.

References

[1] C. Jackulin and S. Murugavalli, “A comprehensive review on detection of plant disease using machine learning and deep learning approaches,” *Measurement: Sensors*, vol. 24, p. 100441, Dec. 2022, doi: 10.1016/j.measen.2022.100441.

[2] S.-E. Lee and J.-O. Kim, “Multi-Scale Attention Based Plant Disease Segmentation Network,” in *2023 International Technical Conference on Circuits/Systems, Computers, and Communications (ITC-CSCC)*, 2023, pp. 1–4. doi: 10.1109/ITC-CSCC58803.2023.10212849.

[3] D. Das, M. Singh, S. S. Mohanty, and S. Chakravarty, “Leaf Disease Detection using Support Vector Machine,” in *2020 International Conference on Communication and Signal Processing (ICCSP)*, 2020, pp. 1036–1040. doi: 10.1109/ICCSP48568.2020.9182128.

[4] G. V, S. P. T, D. J, I. N. M, and S. Suganya, “Plant Leaf Diseases Detection Using KNN Classifier,” in *2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS)*, 2023, pp. 2157–2162. doi: 10.1109/ICACCS57279.2023.10112901.

[5] D. Mohapatra, J. Tripathy, and T. K. Patra, “Rice Disease Detection and Monitoring Using CNN and Naive Bayes Classification,” in *Soft Computing Techniques and Applications*, S. Borah, R. Pradhan, N. Dey, and P. Gupta, Eds., Singapore: Springer Singapore, 2021, pp. 11–29.

[6] D. S. Joseph, P. M. Pawar, and R. Pramanik, “Intelligent plant disease diagnosis using convolutional neural network: a review,” *Multimedia Tools and Applications*, vol. 82, no. 14, pp. 21415–21481, June 2023, doi: 10.1007/s11042-022-14004-6.

[7] B. Nagachandrika, R. Prasath, and I. R. Praveen Joe, “An automatic classification framework for identifying type of plant leaf diseases using multi-scale feature fusion-based adaptive deep network,” *Biomedical Signal Processing and Control*, vol. 95, p. 106316, Sept. 2024, doi: 10.1016/j.bspc.2024.106316.

[8] D. J. Chaudhari and K. Malathi, “Detection and Prediction of Rice Leaf Disease Using a Hybrid CNN-SVM Model,” *Optical Memory and Neural Networks*, vol. 32, no. 1, pp. 39–57, Mar. 2023, doi: 10.3103/S1060992X2301006X.

[9] I. Ahmed and P. K. Yadav, “Plant disease detection using machine learning approaches,” *Expert Systems*, vol. 40, no. 5, p. e13136, June 2023, doi: 10.1111/exsy.13136.

[10] D. Radovanović and S. Đukanovic, “Image-Based Plant Disease Detection: A Comparison of Deep Learning and Classical Machine Learning Algorithms,” in *2020 24th International Conference on Information Technology (IT)*, Feb. 2020, pp. 1–4. doi: 10.1109/IT48810.2020.9070664.

[11] A. K. Singh and N. Nand Dwivedi, “A Comprehensive Review of Machine Learning and Deep Learning Techniques for Plant Disease Identification,” in *2024 Second International Conference Computational and Characterization Techniques in Engineering & Sciences (IC3TES)*, 2024, pp. 1–7. doi: 10.1109/IC3TES62412.2024.10877549.

[12] H. A. Santoso, B. Fandhi Safsalta, N. Febrianto, G. Wilujeng Saraswati, and S.-C. Haw, “Comparative analysis of convolutional neural network and DenseNet121 transfer learning in agriculture focusing on crop leaf disease identification,” *Applied Computing and Informatics*, vol. ahead-of-print, no. ahead-of-print, Jan. 2024, doi: 10.1108/ACI-03-2024-0132.

[13] I. Attri, L. K. Awasthi, T. P. Sharma, and P. Rathee, “A review of deep learning techniques used in agriculture,” *Ecological Informatics*, vol. 77, p. 102217, Nov. 2023, doi: 10.1016/j.ecoinf.2023.102217.

[14] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, “Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 40, no. 4, pp. 834–848, 2017.

[15] C. Jiang, K. Miao, Z. Hu, F. Gu, and K. Yi, “Image Recognition Technology in Smart Agriculture: A Review of Current Applications Challenges and Future Prospects,” *Processes*, vol. 13, no. 5, 2025, doi: 10.3390/pr13051402.

Mô hình kết hợp Atrous CNN-SVM để phân loại bệnh lá lúa

Nguyễn Hữu Huy[1], Bùi Quang Thịnh[2]
[1]Khoa Công nghệ Thông tin, Trường Đại học HUTECH, Thành phố Hồ Chí Minh, Việt Nam
[2]Phòng Quản lý Khoa học Công nghệ và Hợp tác Quốc tế, Trường Đại học Tiền Giang, Đồng Tháp, Việt Nam
nguyenhuuhuy6262@gmail.com, buiquangthinh@tgu.edu.vn

Tóm tắt Bệnh lá lúa là một trong những nguyên nhân hàng đầu làm suy giảm năng suất cây trồng, đòi hỏi phát hiện chính xác và sớm để quản lý hiệu quả. Nghiên cứu này đề xuất một mô hình học sâu lai, Atrous CNN-SVM, kết hợp kiến trúc VGG19 huấn luyện trước được tăng cường bằng các lớp tích chập Atrous để trích xuất đặc trưng đa quy mô và một máy học hỗ trợ vector (SVM) để phân loại mạnh mẽ. Mô hình được đánh giá trên một bộ dữ liệu thực tế thu thập tại Đồng Tháp, Việt Nam, và tiếp tục được xác thực bằng một bộ dữ liệu bên ngoài. Kết quả thực nghiệm cho thấy Atrous CNN-SVM đạt được độ chính xác cao nhất là 86.49% khi được huấn luyện với 90% dữ liệu cục bộ, vượt trội so với các mô hình CNN-SVM tiêu chuẩn, MobileNet và các mô hình đặc trưng được thiết kế thủ công truyền thống. Những phát hiện này nhấn mạnh hiệu quả của việc tích hợp tích chập Atrous với SVM để chẩn đoán bệnh lá lúa tự động và nhấn mạnh tầm quan trọng của việc kết hợp dữ liệu có liên quan tại địa phương trong các ứng dụng công nghệ nông nghiệp chính xác.

Từ khóa Bệnh lá lúa; CNN-SVM; Tích chập Atrous; Học chuyển giao; Nông nghiệp thông minh.