Proceedings of the 8th International Conference on Electronics, Communication and Aerospace Technology (ICECA 2024)
IEEE Xplore Part Number: CFP24J88-ART; ISBN: 979-8-3503-6790-4

2024 8th International Conference on Electronics, Communication and Aerospace Technology (ICECA) | 979-8-3503-6790-4/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICECA63461.2024.10801067

Paddy Plant Disease Detection using Hybrid Deep Learning Algorithms

R Suguna
Assistant Professor
Department of Computer Science and Engineering
Bannari Amman Institute of Technology
Sathyamangalam, Erode-638401, Tamilnadu, India
sugunar@bitsathy.ac.in

S Deepa
Assistant Professor (Sr.G)
Department of Computer Technology - PG
Kongu Engineering College
Perundurai -638060
Erode, Tamilnadu
profsdeepa2020@gmail.com

T Kanimozhi
Assistant Professor
Department of Artificial Intelligence and Data Science
Sri Eshwar College of Engineering
Coimbatore - 641202
kani16.cse@gmail.com

M Chandru
Assistant Professor
Department of Computer Science
VET- Institute of Arts and Science (Co-Education) College
Erode
emchandru@gmail.com

G Arun
Assistant Professor
Department of Computer Science and Engineering
School of Computing, Vel Tech Rangarajan Dr Sagunthala R&D Institute of Science and Technology
Chennai
arung@veltech.edu.in

Harshini M
PG Scholar
M.Sc (Software Systems)
Kongu Engineering College
Perundurai – 638060
Erode, Tamilnadu
mharshini56@gmail.com

***Abstract* - Agriculture is the backbone of India's economy, supporting a large portion of the population and establishing the country as a key global player in food production. India ranks as the second-largest producer of both paddy and wheat. However, agricultural productivity is highly susceptible to various plant diseases, which can significantly reduce crop yields and quality. Paddy, one of the most important staple crops, is particularly vulnerable to diseases caused mainly by viral and bacterial related diseases. The climate conditions are also play a vital role in paddy plant diseases. This research work proposes a deep learning-based approach for early detection and classification of paddy diseases. The proposed research devises a hybrid model that combines the VGG19 and ResNet50 algorithms, the system analyses images of both healthy and disease-affected paddy leaves, sourced from the Kaggle repository. This research focuses on predicting four major paddy diseases that can cause yield losses of up to 80%, and in severe cases, can destroy an entire crop: (i) Paddy Blast, (ii) Bacterial Leaf Blight, (iii) Sheath Blight, and (iv) Ufra Disease. The proposed model achieves a high accuracy rate of 98% in predicting paddy diseases and offers a rapid classification process, showcasing the potential of advanced image processing and deep learning techniques to mitigate the impact of paddy plant diseases on agricultural productivity.**

Keywords - Hybrid Deep learning, Paddy Plant Disease, VGG19 architecture, ResNet50

I. INTRODUCTION

Agriculture is a cornerstone of economic sustenance for numerous nations, providing a vital source of income for countless individuals. Farmers cultivate a variety of food plants according to the specific needs and environmental conditions of their region. The economy of our nation would be directly impacted by the annual losses in paddy production caused by pest attacks and diseases that affect the paddy leaves. While paddy diseases are spread by air and water to cover an entire farm, farmers must take appropriate precautions to prevent them. Once they have, it becomes tough to stop the spread of these diseases, which can result in significant losses while cultivating paddy. The scope of this project extends to the implementation of a hybrid algorithm incorporating deep learning techniques. The goal is to detect diseases in paddy leaves and promptly notify farmers about the specific types of diseases affecting their crops. By integrating deep learning into the algorithm, the system can learn and adapt to various disease patterns, enhancing the accuracy of disease detection. Employing CNN has proven the efficiency in analyzing images of paddy leaves, enabling precise detection of signs of bacterial condition of the given images. The algorithm can process large datasets, recognizing subtle patterns indicative of disease presence. These diseases have a significant negative economic impact on the agricultural industry. This innovative approach not only aids in the early detection of diseases but also reduces the need for excessive fertilizer application. Identifying diseases in their early stages allows farmers to implement targeted interventions, minimizing economic losses and promoting sustainable agricultural practices. Integrating a hybrid algorithm, and incorporating deep learning techniques, enhances the capability of disease detection in paddy cultivation. This proactive approach empowers farmers to protect their crops more effectively, contributing to the overall resilience and sustainability of agriculture in Tamil Nadu.

II. LITERATURE SURVEY

[1] The work utilized machine learning techniques for the classification and spotting of diseased plant leaves. [2] The study proposes an Internet of Things (IoT)-centered strategy for advancing smart agriculture, by seamlessly amalgamating sensor data with machine learning techniques to proficiently tackle crop diseases. [3] presented an

innovative ensemble learning methodology that merges various classification models to elevate the accuracy and robustness of plant disease classification. [4] pioneered a CNN algorithm tailored for diagnosing and distinguishing plant diseases. While their work was groundbreaking, further advancements are imperative to enhance the model's accuracy in identifying plant illnesses. To train the model for identifying five apple leaf illnesses, [5] suggested a CNN technique that used the GoogLeNet Inception and Rainbow concatenation methods. Despite attaining a 74% detection accuracy, this technology recognizes only one thing at a time is a drawback, suggesting that object identification skills could be improved. [6] focused on employing image processing techniques to identify and classify the diseases in citrus plants, addressing challenges of real-world scenarios. Using generative adversarial networks, [7] authors demonstrated a unique transfer learning-based approach for classifying paddy diseases. Although the procedure produces a high accuracy rate, there is a downside. Cheng, J. et al. investigated the fusion of data from multiple sensors to enhance the true prediction rate of plant diseases [8]. Wu, Q., et al. introduced an enhanced monitoring system that can be achieved using the author's adaptive deep learning framework, which is capable of dynamically modifying its parameters in response to the changing landscape of plant diseases. [9]. [10] The study reviewed the challenges and opportunities in implementing artificial intelligence for plant disease management, considering socioeconomic factors [10]. The researchers [11] explored incorporating Satellite Imagery with Machine Learning techniques to predict crop diseases, offering a remote sensing perspective for disease monitoring. Tian, Long, et al. presented an Internet of Things (IoT)-centered strategy for advancing smart agriculture, by seamlessly amalgamating sensor data with machine learning techniques to proficiently tackle crop diseases [13]. Authors effectively utilize the benefits of deep learning algorithms for image classification based paddy disease detection algorithm for early and accurate detection [14],[15],[16],[19],[20]. A new methodology is devised: an attention-embedded MobileNet-V2 model for crop disease identification [17]. A web based application for the detection of diseased paddy plant using Chatbot interface was proposed by the researcher [18].

III. PROBLEM DEFINITION

Conventional approaches for detecting leaf diseases depend on visual inspection by humans, making the process time-consuming and expensive due to the requirement for expert guidance. These methods have inherent limitations, such as dependency on the eyesight of individuals, leading to variations in accuracy and precision. To overcome these drawbacks, deep learning-based approaches, deep learning and innovative solutions have risen to the forefront as a promising method for effectively pinpointing and categorizing plant diseases. Methods offer several advantages, including the ability to accurately identify disease types, make informed decisions, and recommend appropriate treatments. Unlike human vision-based approaches, deep learning models exhibit consistent performance, eliminating variability associated with individual expertise. Consequently, there is a growing demand. In the domain of identifying diseases in paddy plant leaves, the advancement of classification methods based on deep learning is under development. Figure 1 shows the proposed paddy plant disease detection technique. It involves the following steps.

Step 1 : Obtain the data set.
Step 2 : Perform the data preprocessing.
Step 3 : Perform the image augmentation to expand the data set.
Step 4 : Input data set is divided into training and testing purpose.
Step 5 : Apply the hybrid model (VGG19 and ResNet50) for image classification and prediction.

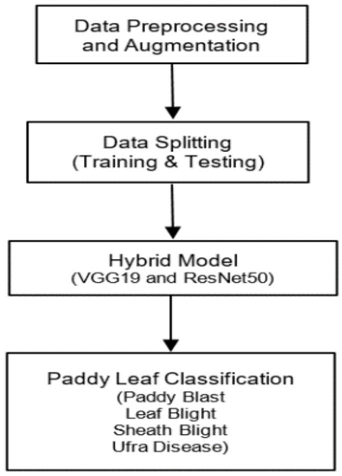

Figure 1: Proposed paddy plant disease detection

IV. PROPOSED METHODOLOGY

The proposed hybrid deep learning algorithm leverages the strength of VGG19 and RestNet model for improved classification of paddy diseases. Figure 2 depicts the paddy plant leaves of bacterial and virus affected images.

Figure 2: Disease affected paddy plant leaves

Algorithm : Hybrid Model (VGG19 + ResNet50)
Step 1: Input I (Images of healthy and affected leaves) Step 2: Extract the feature of input image (I) : F_{VGG19} = VGG19 (I) Step 3: Extract the feature of input image (I): $F_{ResNet50}$ = ResNet50(I) Step 4: Feature fusion : $F_{combined}$ = [F_{VGG19}, $F_{ResNet50}$] Step 5 : Classification Y = FC ($F_{combined}$), followed by softmax or sigmoid for final prediction

Figure 3: Algorithm for proposed hybrid model

VGG19 uses a straightforward architecture with stacked convolutional layers of the same filter size (usually 3x3). This design helps in feature extraction, capturing fine details. ResNet can train extremely deep networks without suffering from the vanishing gradient problem. It excels at learning complex features in very deep layers. Figure 3 shows the algorithm of proposed system.

A. VGG19 Algorithm

VGG19 is a CNN architecture, familiar for its simplicity and uniformity in design. As the name suggests, it boasts 19 layers, structured with alternating convolutional and max-pooling layers.

- Convolutional Layers – The 19 layers of convolution comprise filters/kernels that works on extracting the features of the given input images.
- Max-Pooling Layers – The layers focus on preserving the important features by applying the dimension reduction.

$$Y_{i,j,k} = max_{m,n} X_{p.i+m,p.j+n,k} \quad (1)$$

where X – input, Y-output

- Fully Connected Layers - The feature maps of the input images which are obtained from the convolutional and max-pooling layers are given to this layer for further classification.

$$Y = W. X + b \quad (2)$$

where

W – weight matrix, X – input vector, B – bias term and Y – output vector
The softmax function is applied to the output of the above to transform the raw logits into a probability distribution across the possible classes.

$$P\ (Yi) = \frac{e^{yi}}{\sum_j e^{yi}} \quad (3)$$

where, y_i – i^{th} logit and $P(y_i)$ – predicted probability of class *i*.

- Paddy Plant Disease Detection - To apply VGG19 for detecting diseases in paddy plants, the model uses the healthy and four types of diseases affected leaves for its training phase. Once trained, the model can classify new paddy plant images into various disease categories based on the features it has learned. Figure 3 depicts the architecture of VGG19.

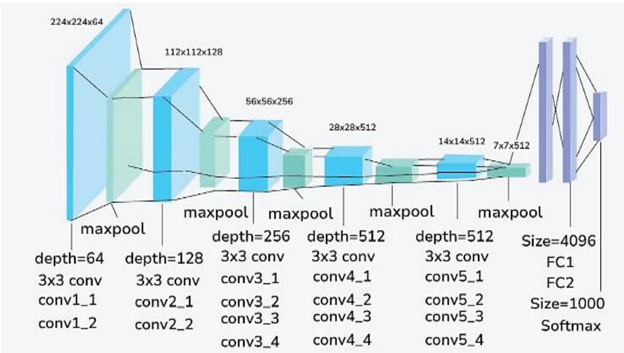

Figure 3: VGG19 architecture

B. Resnet 50 Algorithm

ResNet (Residual Network) is an architecture known for its use of residual blocks, ResNet50, a variant with 50 layers, employs various techniques to tackle challenges like the vanishing gradient problem during training: Figure 4 depicts the architecture of Restnet 50.

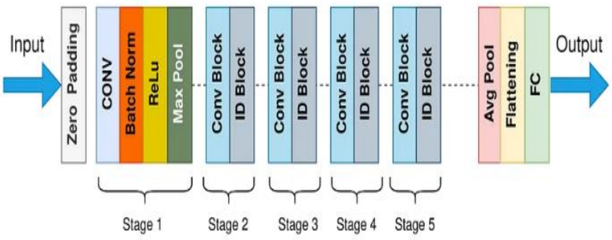

Figure 4 : ResNet50 architecture

- Residual Blocks –By enabling gradients to pass directly through the network, the shortcut connections support more efficient training of deeper architectures.

$$Y_{res} = F(X) + X \quad (5)$$

where *X* – input, *F(X)* – transformation and Y_{res} – final output of the residual block

- Convolutional Layers - Following the footsteps of VGG19, ResNet50 utilizes convolutional layers for effective feature extraction.
- Global Average Pooling - Rather than using fully connected layers, ResNet architectures frequently utilize global average pooling at the end.
- Softmax Layer - Finally, a softmax layer is employed for classification, providing probabilities for each class in the output.
- Paddy Plant Disease Detection - Similar to VGG19, ResNet50 also trained with the input images. ResNet's effectiveness in managing deeper architectures enhances its capability to capture more complex patterns within the data.

$$F(X) = f_{res}\ (f_{conv}\ (X)) \quad (6)$$

where

f_{conv} – convolution and pooling operations

f_{res} – residual connections and transformations

C. Image Preprocessing

The first stage is pre-processing, employed for removing undesirable noise from the loaded image of an infected paddy leaf. To make training the CNN network easier, this step involves shrinking the image to a 50-by-50-pixel size.

Image data undergo preprocessing through several techniques:

1) Resizing: Input images are transformed into a standard dimension for consistency, with common sizes being 224x224 or 256x256 pixels.
2) Normalization: The input images are normalized into a range from 0 to 1.
3) Feature Extraction: Input images are analyzed to extract color, texture, and shape features.

4) GAN Image Augmentation: Augmentation methods such as rotation, blurring, and flipping are used to expand the dataset. Initially containing 580 images of paddy diseases, the dataset is augmented to approximately 4,000 images.

D. Classification

Training Module – Transfer learning VGG19

The CNN-VGG19 model is designed to recognize brown spot paddy leaf diseases, employing transfer learning for accurate identification. Achieving a maximum accuracy, precision, and sensitivity of 93.0%, 92.4%, and 89.9% respectively, this model surpasses others due to its simpler image pre-processing and network architecture, featuring fewer layers compared to ResNet50 or InceptionV3. Despite not being initially trained for disease prediction, applying pre-trained weights enables the model to detect diseases effectively. The process involves capturing images of paddy plant diseases using a digital camera and optimizing further data processing through pre-processing to enhance the acquired image's features.

Testing Module

During the testing phase, a technique known as testing is employed to predict the classifier's output accurately, distinguishing between healthy leaves and various diseases affecting paddy, such as (i) Paddy Blast, (ii) Bacterial Leaf Blight, (iii) Sheath Blight, and (iv) Ufra Disease. This assessment helps determine the effectiveness of the input in correctly identifying different types of paddy leaf diseases.

Implementing the algorithm

To embark on developing a CNN algorithm for disease prediction, the initial stride involves procuring a pertinent dataset. This dataset is crucial for effective training and subsequently evaluating the model's performance. Following this, the necessary libraries are imported to facilitate the coding process. Subsequently, the labels in the dataset are read to ensure accurate classification during training. Pre-processing steps are then applied to the data to enhance its quality and suitability for training. The dataset is a collection of training and testing stages. To gain insights into the dataset, images are displayed using the Matplotlib library. Moving forward, the model construction phase begins, where layers for the CNN algorithm are created. With the model built, the training process is initiated using the prepared dataset. Post-training, the CNN model is ready for predictions. By applying the trained model to new data, the algorithm can accurately predict diseases with a certain level of accuracy. The accuracy rate functions as a pivotal metric for assessing the model's effectiveness in disease prediction. The entire process encompasses dataset handling, model construction, and training, culminating in a reliable algorithm for disease prediction.

IV. RESULT

A. Dataset Description

The proposed system involves a comprehensive two-phase approach encompassing training and testing. The training dataset is robust, featuring a substantial number of images portraying paddy affected by prevalent diseases, specifically images featuring of paddy blast, bacterial leaf blight, sheath blight, and ufra disease. The images for this research work are obtained from the Kaggle data repository. The 70 – 30 percentage splitting of data set comprising of 2800 images for training and a validation dataset consisting of 1200 images, ensuring a well-balanced and representative sample for model development and evaluation.

The VGG19 algorithm has been utilized for classifying diseases in paddy plants. In this approach, the input layer of VGG19 receives an image and applies a series of filters to capture features at varying levels, including edges, textures, and patterns pertinent to paddy plants and their diseases. Through training on a dataset, VGG19 model classify new images into distinct disease categories. To ensure the model's ability to generalize, it is crucial to validate it using a separate dataset that was not included in the training process. Additionally, employing the ResNet50 deep learning-based CNN algorithm for detecting diseases in paddy leaves has shown promising results, achieving high accuracy while also requiring less time for disease classification.

Table 1 shows the attributes considered for the proposed work.

Table 1: Attributes used in paddy plant disease detection

Attributes	Explanation
Leaf texture	Texture as smoothness
Leaf color	Presence of disease-related spot
Leaf shape	Diseases cause deformities
Leaf vein patterns	Patterns formed by veins on leaves
Lesion size and distribution	Related to lesions on leaves on shape

Figure 5 shows that the model exhibits commendable performance with a training accuracy of 96.5%. During rigorous testing, it demonstrates robust capabilities, achieving an accuracy of 99.8% for brown spot, 83.2% for bacterial leaf blight, and 92.4% for leaf smut.

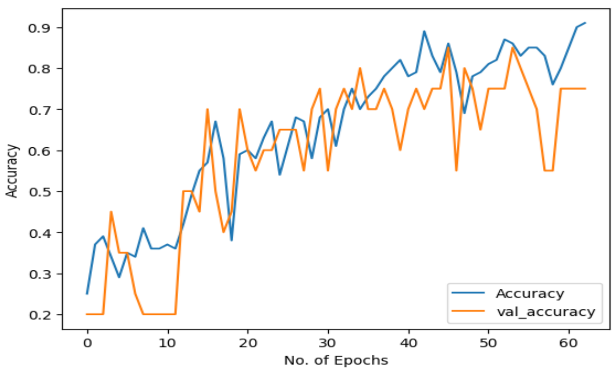

Figure 5: Analysis of hybrid technique

B. Evaluation Metrics

Standard assessment metrics are utilized to measure the model's performance. These metrics include precision, recall, and F-measure. The positively predicted performance of the machine learning model is evaluated by precision metrics. Recall specifically measures the model's ability to identify true positives. A combined score of precision and recall is evaluated with f-score measure. The performance criteria are as follows:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (7)$$

$$Recall = \frac{TP}{TP+FN} \quad (8)$$

$$Precision = \frac{TP}{TP+FN} \quad (9)$$

$$F-Score = \frac{2*Precision*Recall}{Precision+Recall} \quad (10)$$

C. Performance Evaluation

Table 2 shows the comparison measure of proposed methodology with the Densenet, Efficient net and Inception v3 algorithms. The resultant values proved that the proposed hybrid deep learning method outperforms in terms of the evaluation parameters than the other deep learning methodologies.

Table 2: Comparison with various existing accuracy with proposed accuracy

Method	**Parameter Evaluation**			
	Accuracy (%)	**Precision (%)**	**Recall (%)**	**F-measure (%)**
DenseNet-201	78.74	83.27	92.06	91.53
Efficient Net	83.74	84.23	91.44	86.34
InceptionV3	91.29	92.33	94.75	95.63
Proposed VGG19 and ResNet50	**95.59**	**96.34**	**98.36**	**98.57**

From the Table 6 and Figure 6, the hybrid deep learning techniques produced 95.59% accuracy with the f-measure rate of 98.57% which outperforms the existing techniques.

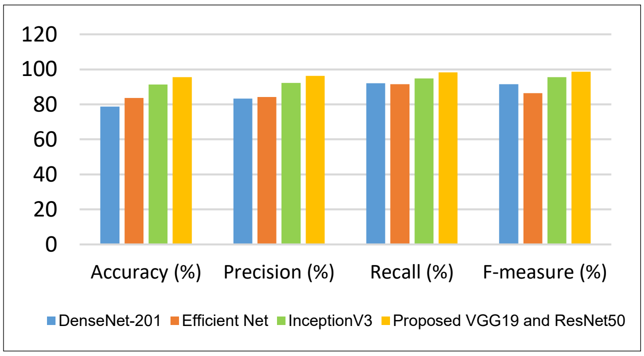

Figure 6: Comparison with existing methods

VI. CONCLUSION

The study concludes that leveraging deep learning methods facilitates a more accessible means for farmers to predict paddy diseases in their early stages, enabling timely intervention before the entire paddy farm succumbs to the ailments. This approach proves to be user-friendly, particularly in its application. The proposed method employs the hybrid CNN algorithm, demonstrating a commendable accuracy rate of 95.59% and the capability to distinguish three prevalent types of paddy leaf diseases, including paddy blast, bacterial leaf blight, sheath blight, and ufra disease. Further enhancements to the proposed approach could involve exploring automatic detection of disease by adding real time image capturing system to automate the prediction techniques.

VII. REFERENCE

[1] Archana, K. S., and Arun Sahayadhas. "Automatic paddy leaf disease segmentation using image processing techniques." vol 6 Int. J. Eng. Technol 7.3.27 (2018): 182-185.

[2] Chen, Junde, et al. "Detection of paddy plant diseases based on deep transfer learning." vol 9 Journal of the Science of Food and Agriculture 100.7 (2020): 3246-3256.

[3] Ferentinos, Konstantinos P. "Deep learning models for plant disease detection and diagnosis." Vol 11 Computers and Electronics in Agriculture 145 (2018): 311-318.

[4] Jiang, Peng, et al. "Real-time detection of apple leaf diseases using deep learning approach based on improved convolutional neural networks." vol 3 IEEE Access 7 (2019): 59069-59080.

[5] Kathiresan, Gugan, et al. "Disease detection in paddy leaves using transfer learning techniques vol 2." Journal of Physics: Conference Series. Vol. 1911. No. 1. IOP Publishing, 2021.

[6] Kawcher, Ahmed, et al. "Paddy leaf disease detection using machine learning techniques." vol 4 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI). IEEE, 2019.

[7] Liang, Wan-jie, et al. "Paddy blast disease recognition using a deep convolutional neural network." vol 7 Scientific reports 9.1 (2019): 1-10.

[8] Li, Dengshan, et al. "A recognition method for paddy plant diseases and pests video detection based on deep convolutional neural network." Sensors 20.3 (2020): 578.

[9] Lu, Yang, et al. vol 22 "Identification of paddy diseases using deep learning.

[10] Prajwa lGowda, B. S., et al. "Paddy crop disease detection using machine learning." International Journal of Engineering Research & Technology 8.13 (2020).

[11] Patil, Nilam Sachin. "Identification of Paddy Leaf Diseases using Evolutionary and Machine Learning Methods." *Turkish* Journal of Computer and Mathematics Education (TURCOMAT) 12.2 (2021): 1672- 1686.

[12] Rahman, Chowdhury R., et al. "Identification and recognition of paddy diseases and pests using convolutional neural networks." Biosystems Engineering 194 (2020): 112-120.

[13] Ramesh, S., and D. Vydeki. "Recognition and classification of paddy leaf diseases using Optimized Deep Neural network with Jaya algorithm." Information processing in agriculture 7.2 (2020): 249-260.

[14] Sethy, Prabira Kumar, et al. "Deep feature based paddy leaf disease identification using support vector machines." Computers and Electronics in Agriculture 175 (2020): 105527.

[15] Shrivastava, Vimal K., et al. "Paddy plant disease classification using transfer learning of deep convolutional neural networks." Vol 6 International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences 42.3 (2019): W6.

[16] Sladojevic, Srdjan, et al. "Deep neural networks based recognition of plant diseases by leaf image classification." vol 10 Computational intelligence and neuroscience 2016 (2016).

[17] Wang, Guan, Yu Sun, and Jianxin Wang. "Automatic image-based plant disease severity estimation using deep learning. vol 11" vol 12 Computational intelligence and neuroscience 2017 (2017).

[18] Wang, Yibin, Haifeng Wang, and Zhaohua Peng. Vol 1"Paddy diseases detection and classification using attention based neural network and Bayesian optimization." Vol 20 Expert Systems with Applications 178 (2021): 114770.

[19] Wang, Jingxian, et al. "CNN Transfer Learning for Automatic Image-Based Classification of Crop." Vol 2 Image and Graphics Technologies and and Applications, IGTA 2018, Beijing, China, April 8–10, 2018, Revised Selected Papers. Vol. 875. Springer, 2018.

[20] Zhou, Guoxiong, et al. "Rapid detection of paddy disease based on FCM-KM and faster R-CNN fusion." Vol 12 IEEE Access 7 (2019): 143190-143206.