

ENHANCED RICE PLANT DISEASE IDENTIFICATION: A HYBRID APPROACH OF TRANSFER LEARNING, SVM AND PCA

SURENDER MOGILICHARLA¹, UPENDRA KUMAR MUMMADI²

¹Research Scholar, Department of Computer Science and Engineering, Osmania University, Hyderabad, Telangana State, India.

²Professor, Department of Computer Science and Engineering, Muffakham Jah College of Engineering & Technology, Osmania University, Hyderabad, Telangana State, India.

E-mail: ¹surender.mogilicharla@gmail.com, ²upendra.kumar@mjcollege.ac.in

ABSTRACT

The accurate and timely identification of diseases and pests impacting rice cultivation is crucial for farmers, allowing for swift intervention and thereby minimizing economic losses. Recent progress in convolutional neural networks (CNNs) has significantly boosted the accuracy of image classification, yet their resource-intensive nature, demanding significant memory and processing power, underscores the necessity of leveraging pre-trained models. Additionally, the paper introduces an ensemble approach combining deep learning with traditional machine learning methods, further enhancing the effectiveness of disease and pest detection in agricultural settings. The paper presents a novel approach leveraging state-of-the-art large-scale architecture, ResNet-50, to propose two distinct models: Model-1 integrates ResNet-50 with SVM, while Model-2 incorporates ResNet-50 with PCA and SVM to effectively detect and identify rice diseases and pests. Through experimentation on authentic and real datasets, the paper demonstrates the effectiveness of these models. Additionally, recognizing the constraints of large-scale architectures, especially concerning their compatibility with mobile or embedded devices due to processing power and memory limitations, the paper introduces and evaluates two proposed models. Notably, Model-2 surpasses Model-1, achieving a superior accuracy of 93.7% compared to 91.6% in Model-1. Moreover, Model-2 significantly reduces the feature set size by 100% compared to Model-1 through dimensionality reduction using PCA.

Keywords: *Deep Learning, Ensemble Methods, Rice Plant Disease Detection, Pretrained Models, Transfer Learning*

1. INTRODUCTION

Automated plant disease detection has become a vital component of precision agriculture. Diseases triggered by fungi, bacteria, and insects greatly affect crop yield and productivity. Classifying plant leaf diseases presents a significant challenge due to striking similarities between classes and complex pattern variations. Moreover, changes in climate conditions can substantially amplify the growth of plant infections. Early detection of diseases affecting plant leaves stands as a paramount challenge for ensuring optimal productivity in the agriculture sector. With diseases accounting for over 50% of the reduction in plant productivity, timely identification facilitates prompt intervention, thereby mitigating crop loss.

In agriculture, rice plants face considerable

challenges from diseases that impact crop quality and quantity. This highlights the pressing requirement for automated plant disease identification and detection to improve yield. Rice holds immense importance as a key cereal crop in our country, with rice farming playing a pivotal role in the agricultural economy. Throughout the growth stages of rice, diseases affecting the plants represent a primary concern for farmers, leading to substantial losses alongside issues such as pests and environmental factors. Despite the availability of various methods for disease detection in crops, such as image processing and remote sensing, these systems often exhibit poor accuracy.

The remarkable growth of the Indian population necessitates significant enhancements in cultivation and agricultural practices. Rice,

being a staple crop, holds unparalleled importance as the most coveted food crop throughout the nation [1]. Nevertheless, rice stands out as a crop highly susceptible to plant diseases, which consequently impacts cultivation and overall profitability. To mitigate such challenges and bolster crop yields, it's imperative to anticipate and prevent plant diseases at their early stages[2]. Forecasting diseases through visual observation alone is often slow, occasionally imprecise, and can result in heightened expenses. Estimating disease types is challenging and susceptible to errors [3]. The challenges stem from a lack of comprehensive understanding of the plant. Consequently, if diseases in rice plants are not anticipated or identified in their early stages, it adversely impacts rice production, a trend observed over recent decades[4]. The image processing technique for disease detection typically involves several steps, including image acquisition, preprocessing, segmentation, feature extraction, and classification [5]. Handcrafted feature extraction for rice plant disease detection comes with inherent limitations[6]. Firstly, crafting effective features manually demands specialized domain knowledge and consumes significant time, posing challenges in capturing all pertinent information comprehensively. Moreover, handcrafted features may lack robustness in generalizing across various types of rice plant diseases and diverse environmental conditions. Furthermore, this approach exhibits limited adaptability, necessitating manual intervention for feature modifications to address evolving disease patterns or incorporate new data. The limitations associated with the utilization of Convolutional Neural Networks (CNNs) for detecting diseases in rice plants include the requirement for extensive labeled data, susceptibility to overfitting, and computational complexity, presenting challenges in resource-limited settings[6].

The current research paper proposes an innovative methodology that integrates a Convolutional Neural Network (CNN) utilizing a ResNet-50 architecture with conventional machine learning methods like Principal Component Analysis (PCA) and Support Vector Machine (SVM). The study utilizes a large dataset comprising images of healthy and unhealthy rice plants collected from real-world environments. By employing a pretrained ResNet model for feature extraction, training time overhead is minimized. Two models are proposed: Model-1 involves feeding ResNet-50 output features directly into an SVM classifier, while Model-2 employs PCA to reduce feature dimensionality before inputting them into an SVM for classification.

Benchmark datasets are utilized to assess the performance of the proposed models. It turned out that Model-2 outperforms Model-1. This approach is assessed using performance metrics including accuracy, precision, recall, and F1 score.

2. RELATED WORK

In agricultural image recognition and classification, two primary technological categories dominate: deep learning and traditional machine learning[7]. Notably, deep learning has shown rapid advancements and notable achievements in this domain. For example, Mohanty et al. have successfully developed a deep learning model with the ability to identify 14 different crop species and detect 26 various crop diseases.[8]. Previous research efforts in the realm of rice plant infection recognition and classification have been limited. Lu et al. [9] introduced a novel method employing Deep Convolutional Neural Network (DCNN) for predicting rice plant diseases. Their research utilized a dataset comprising numerous images depicting both healthy and diseased paddy stems and leaves, resulting in enhanced accuracy when compared to conventional machine learning methods. Furthermore, a segmentation model was developed by Dhingra et al.[10] based on neutrosophic logic, originating from fuzzy set theory theory to estimate Regions of Interest (ROI). This model utilized three Membership Functions (MFs) for segmentation, employing feature subsets for predicting the presence of plant leaf infections based on segregated sites. Islam et al. [11] introduced an innovative method for predicting and classifying rice plant diseases. Their approach employs image processing (IP) techniques to detect diseases by analyzing the proportion of RGB values in the affected area. In the context of paddy leaf disease prediction, automated disease detection is facilitated through the utilization of image processing (IP) techniques [12]. In this investigation, a hybridized methodology integrating grayscale co-occurrence matrix, Discrete Wavelet Transform (DWT), and Scale Invariant Feature Transform (SIFT) was adopted for feature extraction. These extracted features were subsequently employed in conjunction with various machine learning classifiers to distinguish between healthy and diseased crops.

Kaya et al. [13] analyzed the outcomes of implementing four distinct Transfer Learning (TL) strategies for plant classification utilizing Deep Neural Networks (DNNs) on four standard datasets. Their findings underscored the substantial advantages of TL in automated plant prediction, thereby enhancing the performance of plant disease classifiers. In a prior study [14], A Convolutional Neural Network (CNN) model was utilized to forecast weeds in soybean crop images, distinguishing between grass and broadleaf weeds. The image dataset encompassed various soil, soybean, broadleaf, and grass weed images. CNN, employed for Deep Learning (DL), yielded optimal results in image recognition tasks.

3. THE PROPOSED RICE PLANT DISEASE DETECTION MODEL

3.1 Framework Overview

Our research endeavors in rice plant disease identification employ a dual approach, integrating transfer learning and traditional machine learning algorithms, delineated in Fig. 1 and Fig. 2, where Fig. 1 corresponds to Model-1 and Fig. 2 to Model-2. Initially, rice plant images are categorized based on predefined classes as detailed in Table 1, then processed through a pretrained ResNet-50 model via transfer learning. This model, initially trained on the ImageNet dataset featuring 1000 classes, adeptly extracts multi-level features. Leveraging transfer learning enables us to conserve computational resources by retaining the pretrained network parameters.

In Model-1, the pretrained ResNet-50 model automatically extracts features, yielding embeddings utilized for classification via an SVM machine learning classifier. Conversely, in Model-2, the model

again extracts features, followed by dimensionality reduction through Principal Component Analysis (PCA), ensuring retention of 99% variance. PCA serves a dual role: reducing computational overhead and mitigating overfitting risks inherent in high-dimensional embeddings. The streamlined embeddings are then fed into an SVM classifier for classification. Further elucidation on these methodologies is provided in subsequent sections.

3.2 Overview of Model-1

Figure 1 showcases an elaborate block diagram outlining our proposed model for rice plant disease detection. The illustrated flowchart delineates the sequence of steps, commencing with the input of rice plant images into the model. This model incorporates a pretrained ResNet-50 model, proficient at feature extraction from the input images. Subsequently, these extracted features are fed into an SVM classifier to precisely categorize the images into their respective disease classes. A notable advantage of our approach is the integration of a pretrained ResNet-50 model [15], significantly reducing training time while leveraging the model's expertise in feature extraction from images. Additionally, we capitalize on the efficiency of using a Support Vector Machine (SVM) [16] classifier instead of Artificial Neural Networks (ANN) [17], enables our model to operate seamlessly even on systems possessing limited restricted resources, such as mobile devices or embedded systems. This strategic choice ensures that our model remains accessible and applicable across a wide range of platforms, including those with low-end resources.

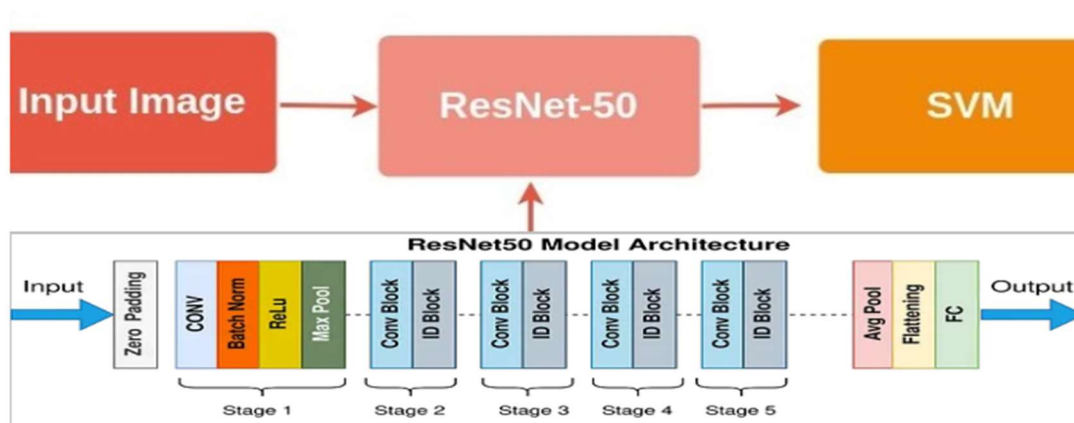


Figure 1: Model-1 Integrated Framework For Disease Detection In Rice Plants.

3.3 Overview of Model-2

Figure 2 depicts a detailed block diagram of our innovative framework designed for the detection of diseases in rice plants. The flowchart visually delineates the sequential process, commencing with the input of rice plant images into the model. An integral component of this model is the inclusion of a pretrained ResNet-50 model, renowned for its

adeptness in feature extraction from images. Following feature extraction, the features undergo dimensionality reduction through PCA, effectively reducing the number of features. Finally, the reduced-dimensional features are obtained inputted into an SVM classifier, ensuring precise classification of the images into their respective disease classes.

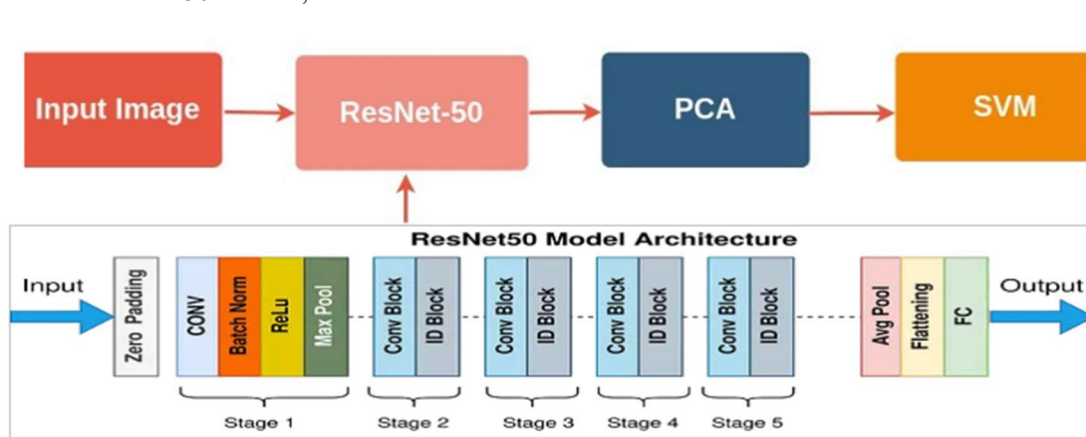


Figure 2 : Model-2 Integrated Framework For Disease Detection In Rice Plants.

An evident strength of our methodology lies in the incorporation of a pretrained ResNet-50 model, markedly diminishing training time while capitalizing on its proficiency in extracting features from images. Moreover, by implementing PCA [21] to condense the feature set, we optimize computational efficiency without sacrificing accuracy. Embracing the efficacy of a Support Vector Machine (SVM) [22] classifier over Artificial Neural Networks (ANN) [23] further enhances our model's adaptability, facilitating seamless operation even on resource-constrained platforms like mobile or embedded systems. This strategic decision ensures the accessibility and applicability of our model across diverse platforms, including those with limited resources.

3.4 ResNet-50

ResNet-50 [24] is a convolutional neural network architecture renowned for its depth and efficacy in image-related tasks. With 50 layers, it excels in feature extraction from input images typically sized at 224x224 pixels in RGB format. Its output dimensions vary depending on the task, often culminating in 1000 classes for ImageNet classification. Notably, ResNet-50's innovation lies in its skip connections, which directly link layers, addressing the issue of vanishing gradients and enhancing smoother gradient flow during the training

process. Further enhancing its efficiency, ResNet-50 incorporates bottleneck blocks, where 1x1 convolutions reduce computational complexity without compromising representation power. Lastly, it adopts global average pooling as its final layer, summarizing feature maps into a compact representation, instead of the traditional fully connected layers, contributing to its widespread adoption and success in computer vision applications.

3.5 Pca

PCA [25], an extensively utilized unsupervised machine learning technique, serves multiple purposes such as dimensionality reduction, visualization, noise filtering, and feature extraction[26]. Consider a dataset comprising grayscale images of rice plants [27], each sized at 200x200 pixels, resulting in a total of 40,000 pixels per image. Upon applying PCA, we observe that the first 25 components retain 80% of the variance, while the first 75 components capture 90% of the variance. Consequently, we can compress the original 40,000 pixels into a 75-dimensional vector, preserving at least 90% of the image information. This condensed 75-dimensional vector can then be referred to as the pixel features of the rice

plant images. We have used PCA in our proposed Model-2, where it has reduced the number of 100352 features generated from ResNet-50 to 941 features reduced to 100% by retaining the main information of the images. To mitigate overfitting during the training process, we'll apply Principal Component Analysis (PCA) to the features extracted from ResNet-50, aiming to retain 99% of the variance. This step is crucial for its potent dimensional reduction effect, ensuring the preservation of essential information [28].

3.6 Svm

Support Vector Machines (SVMs) [29] are utilized for classifying a rice image dataset consisting of 9 classes, encompassing 5 diseases, 3 pests, and 1 healthy class. SVMs operate by determining the hyperplane that optimally segregates these classes in the feature space, aiming to maximize the margin between classes. Employing kernel functions, SVMs accommodate linear and non-linear classification, facilitating the classification of rice images afflicted by various diseases and pests. Their robustness against overfitting makes SVMs well-suited for this task, contributing to their efficacy in identifying and distinguishing between different rice health conditions [30].

4. RESULTS AND DISCUSSION

In this section, a sequence of tests was carried out on a laptop T490, featuring a Core i7 CPU and 16 GB of memory. The experimental configuration employed Jupyter Notebook operating

with Python 3.7. Initially, tests were executed on the dataset utilizing Model-1, succeeded by experiments with Model-2. Following this, an extensive comparison and assessment were undertaken, scrutinizing performance across diverse evaluation metrics including accuracy, precision, recall, F1 score, and feature count [31]

4.1 Dataset

Rice plants are vulnerable to a multitude of diseases and pests, which can impact various parts of the plant. This study encompasses nine classes, including five diseases, three pests, and one healthy plant class. The classification is detailed in Table 1, with certain diseases grouped together due to similar treatment methods and occurrence patterns. Symptoms manifest in different parts of the rice plant, with diseases like Bacterial Leaf Blight and pests like Brown Plant Hopper primarily affecting the leaves. Other diseases such as Sheath Blight and pests like Stemborer target the stem, while Neck Blast and False Smut affect the grains. To prevent confusion between diseased and dead plant parts, images of both have been included in the dataset. Note that Sheath Blight, Sheath Rot and their simultaneous occurrence have been considered in the same class, because their treatment method and place of occurrence are the same. Additionally, some classes exhibit multiple symptom variations, also detailed in Table 1, to encompass all observed variations found in BRRI's paddy fields. Sample images for each class are illustrated in Figure 3.

Table 1: Collection Of Images Representing Various Classes.

Class Name	Type	Number of collected Images	In tra-class variations in symptoms	Images count
Flase smut	Disease	93	Brown symptoms	66
			Black symptoms	27
Brown Plant Hopper (BPH)	Pest	71	Early detection of BPH infestation	50
			Advanced stage of BPH infestation.	21
Bacterial Leaf Blight (BLB)	Disease	138	No noticeable symptoms variation.	138
Neck Blast	Disease	286	No noticeable symptoms variation.	286

Stemborer	Pest	201	Symptoms of stem borer pest infestation on grains.	180
			Symptoms of stem borer pest infestation on stems.	21
Hispa	Pest	73	black pests and also white spots visible on plant leaves	53
			Pronounced spots on leaves with no visible pests suggest a potential issue or infection.	20
Sheath Blight & Sheath Rot	Disease	219	Black stems indicate a potential problem or disease affecting the plant.	70
			White spots on the plant indicate potential issues	77
			Mixed black and white symptoms	72
Brown Spot	Disease	111	No noticeable symptoms variation.	111
Healthy	Healthy	234	Healthy green leaves and stems indicate optimal plant health.	96
			Yellow grains indicate that the plant has reached maturity and is healthy.	71
			Dead leaves and stems	67

The dataset utilized in this research, comprising a collection of 1426 images depicting various rice diseases and pests has been gathered. in real-life scenarios Bangladesh Rice Research Institute's [BRRRI] paddy fields, is made publicly available by Chowdhury R. Rahman et al. [15]. The dataset encompasses nine classes, including eight classes representing various rice diseases and pests,

along with one class denoting healthy rice plants. Access to the dataset can be found through the provided [link](https://drive.google.com/open?id=1ewBesJcguriVTX8sRJseCDbXAF_T4akK) (https://drive.google.com/open?id=1ewBesJcguriVTX8sRJseCDbXAF_T4akK) [15]. Additional information about the dataset is available in the cited publication [15], where the author utilized it for implementing a simple CNN.



Figure 3 : A Sample Image Of Each Detected Class.

4.2 Experimental findings

In this study, we conducted experiments to evaluate the performance of two models, referred to as Model-1 and Model-2, for a specific task. The experiments were carried out using the Rice dataset, wherein two models were proposed. In Model-1, an ensemble model comprising a pretrained ResNet-50 and SVM was employed, while Model-2 utilized an ensemble model consisting of ResNet-50, PCA, and SVM. Specifically, the training and testing sets were divided in a 8:2 ratio. For PCA, components were selected to retain 99% variance.

Model-1 utilizes a pretrained ResNet-50 model for feature extraction. The ResNet-50 model generated a total of 100,352 features from the input data. These features were then fed into a Support Vector Machine (SVM) classifier for classification. The accuracy achieved by Model-1 was measured to be 91.6%, with an F1 score of 81.7. The recall and precision rates were found to be 82.4% and 81.8%, respectively.

Model-2 also utilizes a pretrained ResNet-50 model for feature extraction. Furthermore, Principal

Component Analysis (PCA) was employed to decrease the dimensionality of the extracted features, leading to a notable reduction in the feature count down to 941, representing almost 100% reduction compared to Model-1. The reduced-dimensional features were then fed into an SVM classifier for classification. The final outcomes of Model-2 demonstrated improved performance, with an accuracy of 93.7% and an F1 score of 86.4. The recall and precision rates for Model-2 were measured at 87.8% and 89.6%, respectively.

4.3 Comparative analysis of Proposed Models

To evaluate the performance and generalizability of our proposed models, experiments were carried out using a dataset consisting of 1426 images of paddy. This dataset consists of 9 classes, including 8 diseased classes and 1 healthy class. All images are standardized to a size of 256x256 pixels with 3 channels. The dataset is partitioned into training and testing sets at a ratio of 8:2. Evaluation metrics encompass precision, recall, F1-score, and the number of parameters.

$$\text{Accuracy} = (T_n + T_p) / (T_n + F_p + T_p + F_n)$$

$$\text{Precision} = T_p / (T_p + F_p)$$

$$\text{Recall} = T_p / (T_p + F_n)$$

$$\text{F1 Score} = 2 * (\text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall})$$

In binary classification, T_p : True positive, T_n : True negative, F_p : False positive, and F_n : False negative are technical terms used to evaluate classifier performance. Specifically, T_p refers to positive samples correctly classified, T_n denotes

negative samples correctly classified, F_p indicates positive samples misclassified, and F_n represents negative samples misclassified [32].

The experimental comparison results are presented in Table 2. These experiments were conducted using a designated experimental environment, ensuring consistency and reproducibility in the results obtained. The results offer valuable insights into the efficacy of the proposed models for the given task. Here, we provide a comparative analysis of their performance.

Table 2 : The Metrics From The Comparison Experiments.

Dataset	Indicator	Model-1	Model-2
Rice	Accuracy	0.916	0.937
	Precision	0.818	0.896
	Recall	0.824	0.878
	F1-score	0.817	0.864
	Number of Features	100352	941

Model-2 exhibits superior performance over Model-1 across all metrics tested on the rice image dataset. Despite having a significantly smaller number of features (941 compared to Model-1's 100,352), Model-2 demonstrates higher accuracy (0.937 vs. 0.916), precision (0.896 vs. 0.818), recall (0.878 vs. 0.824), and F1-score (0.864 vs. 0.817). This suggests that Model-2 is more adept at accurately identifying rice instances within the image dataset, outperforming Model-1 even with a substantially reduced feature set.

selection in optimizing performance, as Model-2 showcases that superior results can be achieved with a more streamlined approach.

Additionally, Model-2's higher precision indicates its greater ability to correctly identify true positives, while its higher recall suggests it captures a larger proportion of actual rice instances, ultimately leading to a better balance between minimizing false positives and false negatives. Model-2 emerges as the preferred choice for the rice image dataset due to its consistently higher performance across all evaluated metrics. Its effectiveness in achieving high accuracy, precision, recall, and F1-score despite a significantly smaller number of features underscores its efficiency in rice image classification tasks. This highlights the importance of model architecture and feature

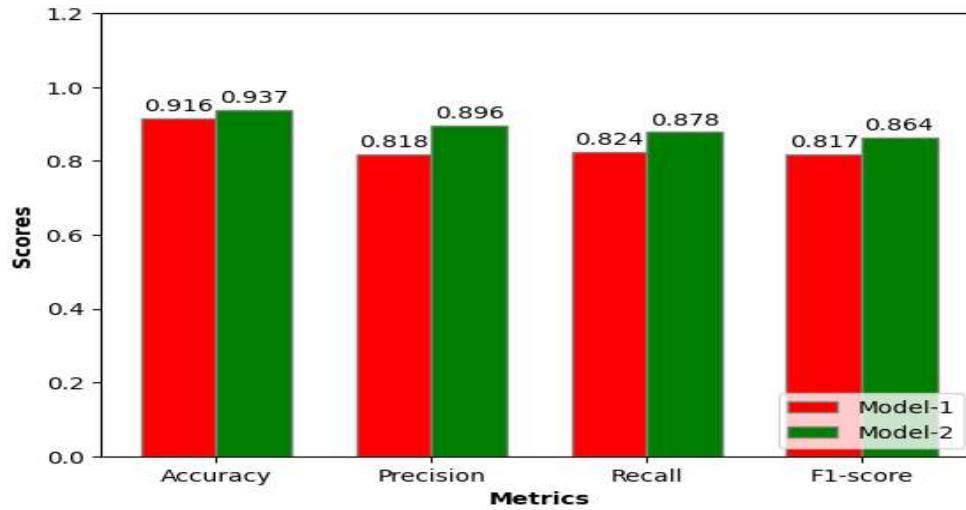


Figure 4: : Comparison Of Model-1 and Model-2 Evaluation Metrics

The below is the confusion matrix [Table 3] generated on test sample's using Model-1. The confusion matrix illustrates how well the model performs in classifying various rice diseases and health states. Strong identification of healthy rice (H) and diseases like Hispa (H_{is}) and Blast (B_{lb}) is

evident, with notable misclassifications in distinguishing between Blast and Blast affected high (B_{phs}), and in classifying False Smut (F_s) and Shudra Rot (S_{br}). These findings highlight both the model's strengths and areas for refinement in rice disease classification.

Table 3 : Confusion Matrix Generated Using Model-1.

True Label	Predicted Label										
	B _{lb}	B _{phe}	B _{phs}	B _s	F _s	H	H _{is}	N _b	S _{br}	S _{tm}	
B _{lb}	23	0	0	0	0	0	0	0	0	0	0
B _{phe}	0	9	0	0	0	1	1	0	0	0	0
B _{phs}	0	3	0	0	0	0	0	0	2	0	0
B _s	0	0	0	3	0	0	0	0	0	0	0
F _s	1	0	0	0	28	1	0	1	0	0	0
H	1	0	0	0	0	37	0	0	2	0	0
H _{is}	2	1	0	0	0	2	16	0	3	0	0
N _b	0	0	0	0	0	0	0	56	0	0	0
S _{br}	0	0	0	0	0	1	0	0	48	0	0
S _{tm}	0	0	0	0	0	0	0	1	1	42	0

The below is the confusion matrix [Table 4] generated on test sample's using Model-2. The updated confusion matrix demonstrates improved performance over the previous one, with fewer misclassifications and enhanced accuracy in disease

classification. Notably, distinctions between Blast and Blast affected high (B_{phs}) and other diseases like Hispa (H_{is}) are clearer, resulting in more accurate predictions. These enhancements highlight the model's refined capability in

accurately identifying various rice diseases, addressing previous limitations and showcasing improved classification accuracy.

This confusion matrix reveals insights into the performance of the classification model across multiple classes representing various plant diseases and conditions. Examining the diagonal elements, which indicate correct predictions, it's apparent that the model performs exceptionally well in identifying Healthy (H) plants, with 36 out of 40 instances correctly classified. Additionally, the model demonstrates strong performance in detecting Neck Blast (Nb) and Brown Spot (Bs), with all instances accurately classified.

However, some misclassifications are evident, particularly in distinguishing between Brown Plant Hopper early symptoms (Bphe) and False Smut (Fs), where three instances of Bphe were incorrectly classified as Fs. Similarly, there are instances where Hispa (His) was misclassified as Brown Plant Hopper severe symptoms (Bphs), suggesting a potential overlap or similarity between these conditions. Despite these misclassifications, the overall performance of the model appears commendable, with the majority of classes accurately predicted.

Table 4 : Confusion Matrix Generated Using Model-2.

True Label	B _{lb}	23	0	0	0	0	0	0	0	0	0
	B _{phe}	0	8	0	0	0	0	3	0	0	0
	B _{phs}	0	1	2	0	0	0	2	0	0	0
	B _s	0	0	0	3	0	0	0	0	0	0
	F _s	0	0	0	0	30	0	0	1	0	0
	H	0	0	0	3	0	36	0	0	1	0
	H _{is}	3	0	0	0	0	0	21	0	0	0
	N _b	0	0	0	0	0	0	0	56	0	0
	S _{br}	0	0	0	0	0	1	0	1	47	0
	S _{tm}	0	0	0	0	0	0	0	1	1	42
		B _{lb}	B _{phe}	B _{phs}	B _s	F _s	H	H _{is}	N _b	S _{br}	S _{tm}
Predicted Label											

B_{lb} :Bacterial leaf blight, B_{phe} :Brown plant hopper early symptoms,B_{phs} : Brown plant hopper severe symptoms, B_s :Brown spot, F_s :Flase smut, H: Healthy, H_{is} : Hispa, N_b : Neck Blast, S_{br} : Sheath blight and/or sheath rot, S_{tm} : Stemborer

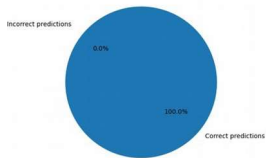
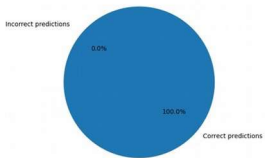
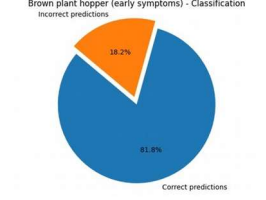
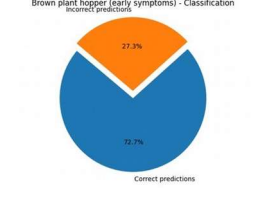
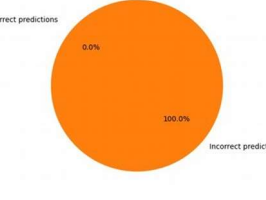
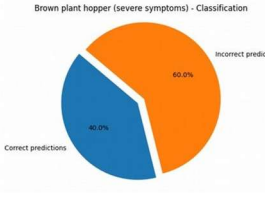
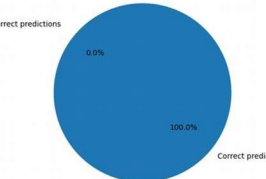
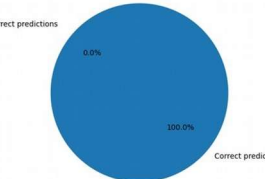
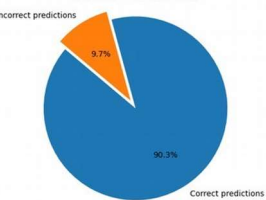
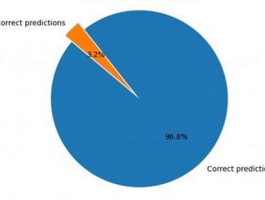
In model-1, although it demonstrated a high accuracy and precision across several classes such as "Healthy" and "Neck Blast", it struggled with the early prediction of "Brown plant hopper severe symptoms" (Bphs) and "Hispa" (His) classifications, where it misclassified a notable portion of instances. For instance, in the case of Bphs, it incorrectly classified 3 instances as "Hispa" and 2 instances as "Brown spot". Similarly, in the case of His, it misclassified 2 instances as "Healthy" and 3 instances as "Neck Blast". These misclassifications suggest a weakness in model-1's ability to discern subtle symptoms and distinguish

between closely related classes. The pie charts illustrating these performances are provided below the table, offering a visual representation of the distribution of correct and incorrect predictions for each class, shedding further light on the model's classification capabilities.

On the other hand, model-2 displayed improved performance in predicting early symptoms of "Brown plant hopper severe symptoms" (Bphs) and "Hispa" (His) compared to model-1. It correctly classified a higher number of instances for these classes, indicating its enhanced

sensitivity to subtle symptoms and finer distinctions. For example, in the case of Bphs, model-2 correctly classified 2 instances, which were misclassified by model-1. Similarly, for the His class, model-2 correctly classified all instances, while model-1 misclassified several instances. This suggests that model-2 may be more adept at

capturing nuanced patterns and making accurate predictions for these particular classes. Similarly, the pie charts depicting the class-wise performances of model-2 are located below the table, providing visual insights into its classification accuracy.

S.No	Disease/Class Name	Model-1 Performance	Model-2 Performance
1	Baterial leaf blight		
2	Brown plant hopper early symptoms		
3	Brown plant hopper severe symptoms		
4	Brown spot		
5	Flase smut		

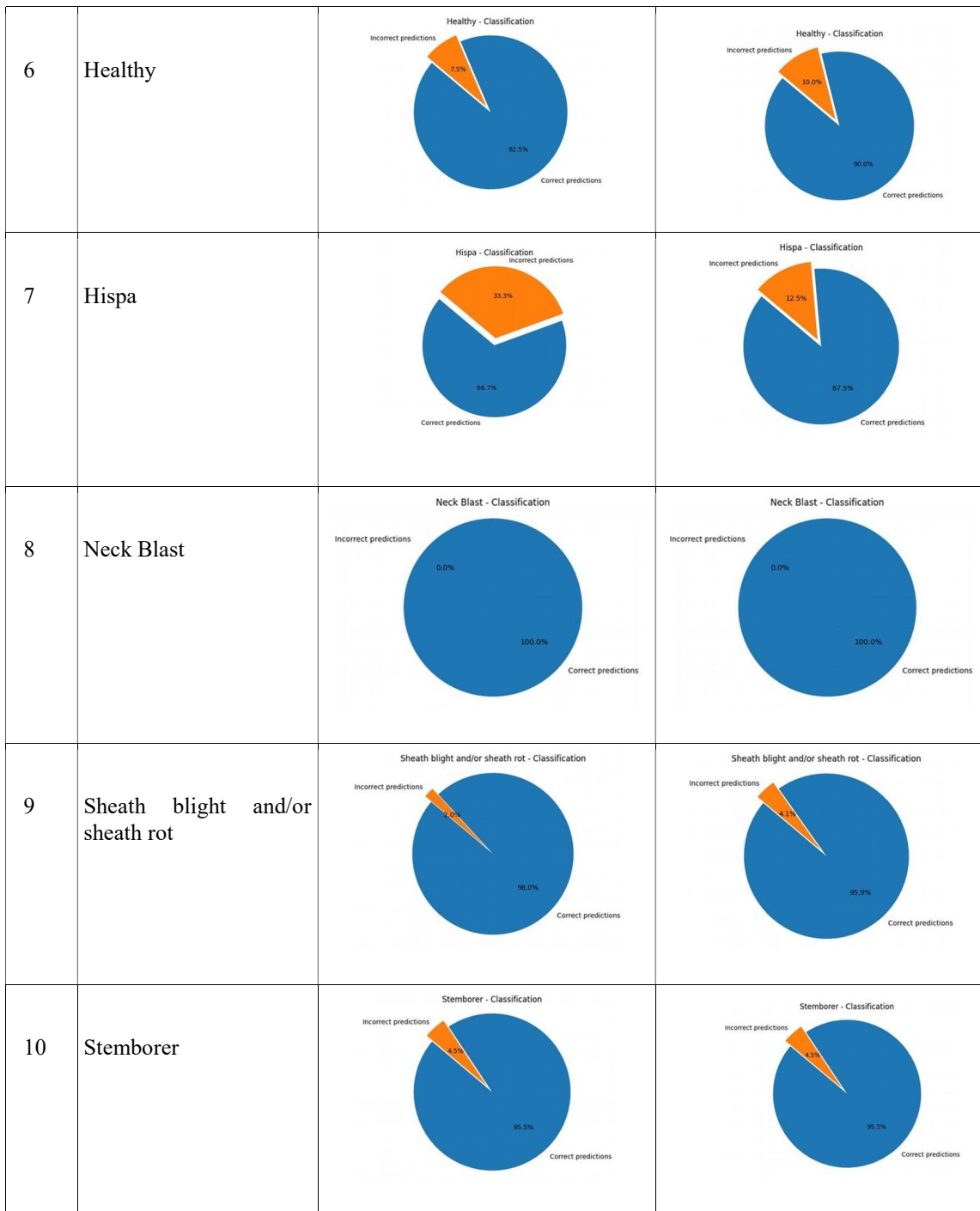


Figure 5 : Class Wise Distribution Of Predictions Of Model-1 (Left) And Model-2 (Right)

The comparison of the two models reveals that both exhibit high accuracy, with Model-2 outperforming Model-1. Model-2 demonstrates a lower rate of incorrect predictions (6.3%) compared to Model-1 (8.4%), resulting in a higher percentage of correct predictions (93.7% vs. 91.6%). This indicates that Model-2 offers enhanced precision in

classifying rice diseases, potentially leading to more reliable diagnoses. The pie chart visualizations underscore these findings, depicting a marginally improved distribution of predictions in Model-2, further highlighting its superior performance in disease classification tasks.

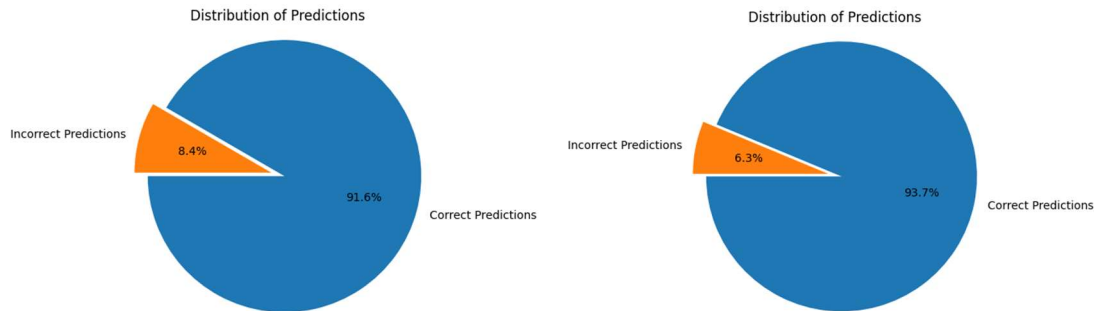


Figure 6 : Overall Distribution Of Predictions Of Model-1 (Left) And Model-2 (Right)

5. CONCLUSION

In conclusion, our study underscores the critical importance of accurate disease and pest identification in rice cultivation for effective agricultural management. Leveraging pretrained models, such as ResNet-50, in conjunction with advanced machine learning techniques like support vector machine (SVM) and principal component analysis (PCA), as demonstrated by Model-1 and Model-2, our research unveils promising avenues for enhancing classification accuracy in agricultural settings.

The utilization of state-of-the-art architecture like ResNet-50, already trained on large datasets, combined with machine learning methodologies, showcases the potential for enhanced disease and pest detection. Particularly noteworthy is the superior performance of Model-2, achieving an impressive accuracy of 93.7% while significantly reducing feature set size through PCA. These findings not only contribute to advancing agricultural technology but also address practical challenges such as model scalability and compatibility with resource-constrained environments.

By providing insights into optimizing model efficiency and performance, this research contributes to the development of accessible and effective tools for disease and pest management in rice cultivation, thereby promoting agricultural sustainability and economic resilience.

Looking ahead, there is substantial potential in further exploring pretrained models and integrating location, weather, and soil data with images of diseased plant parts to devise a comprehensive and automated plant disease detection system. Further

exploration into segmentation or object detection algorithms could amplify the efficacy of rice disease and pest classification, especially in environments with diverse backgrounds.

REFERENCES

- [1] X. E. Pantazi, D. Moshou and A. A. Tamouridou, "Automated leaf disease detection in different crop species through image features analysis and one class classifiers," *Computers and Electronics in Agriculture*, vol. 156, pp. 96–104, 2019.
- [2] D. Y. Kim, A. Kadam, S. Shinde, R. G. Saratale, J. Patra et al., "Recent developments in nanotechnology transforming the agricultural sector: A transition replete with opportunities," *Journal of the Science of Food and Agriculture*, vol. 98, no. 3, pp. 849–864, 2018.
- [3] A. K. Mahlein, "Plant disease detection by imaging sensors—Parallels and specific demands for precision agriculture and plant phenotyping," *Plant Disease*, vol. 100, no. 2, pp. 241–251, 2016.
- [4] F. T. Pinki, N. Khatun and S. M. M. Islam, "Content based paddy leaf disease recognition and remedy prediction using support vector machine," in *2017 20th Int. Conf. of Computer and Information Technology*, Dhaka, Bangladesh, IEEE, pp. 1–5, 2017.
- [5] M. M. Kamal, A. N. I. Masazhar and F. A. Rahman, "Classification of leaf disease from image processing technique," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 10, no. 1, pp. 191–200, 2018.
- [6] Loris Nanni, Stefano Ghidoni, Sheryl

- Brahnam, Handcrafted vs. non-handcrafted features for computer vision classification, *Pattern Recognition*, Volume 71, 2017, Pages 158-172, ISSN 0031-3203, <https://doi.org/10.1016/j.patcog.2017.05.025>.
- [7] Liakos, K.G.; Busato, P.; Moshou, D.; Pearson, S.; Bochtis, D. Machine Learning in Agriculture: A Review. *Sensors* 2018, 18, 2674. <https://doi.org/10.3390/s18082674>.
- [8] Mohanty, Sharada P., David P. Hughes, and Marcel Salathé. "Using deep learning for image-based plant disease detection." *Frontiers in plant science* 7 (2016): 1419.
- [9] Y. Lu, S. Yi, N. Zeng, Y. Liu and Y. Zhang, "Identification of rice diseases using deep convolutional neural networks," *Neurocomputing*, vol. 267, no. 1, pp. 378–384, 2017.
- [10] G. Dhingra, V. Kumar and H. D. Joshi, "A novel computer vision based neutrosophic approach for leaf disease identification and classification," *Measurement*, vol. 135, no. 1, pp. 782–794, 2019.
- [11] T. Islam, M. Sah, S. Baral and R. R. Choudhury, "A faster technique on rice disease detection using image processing of affected area in agro-field," in *2018 Second Int. Conf. on Inventive Communication and Computational Technologies*, Coimbatore, India, IEEE, pp. 62–66, 2018.
- [12] T. G. Devi and P. Neelamegam, "Image processing based rice plant leaves diseases in Thanjavur, Tamilnadu," *Cluster Computing*, vol. 22, pp. 13415–13428, 2019.
- [13] A. Kaya, A. S. Keceli, C. Catal, H. Y. Yalic, H. Temucin et al., "Analysis of transfer learning for deep neural network based plant classification models," *Computers and Electronics in Agriculture*, vol. 158, pp. 20–29, 2019.
- [14] A. S. Ferreira, D. M. Freitas, G. G. da Silva, H. Pistori and M. T. Folhes, "Weed detection in soybean crops using ConvNets," *Computers and Electronics in Agriculture*, vol. 143, pp. 314–324, 2017.
- [15] Chowdhury R. Rahman, Preetom S. Arko, Mohammed E. Ali, Mohammad A. Iqbal Khan, Sajid H. Apon, Farzana Nowrin, Abu Wasif, "Identification and recognition of rice diseases and pests using convolutional neural networks", *Biosystems Engineering*, Volume 194, 2020, Pages 112-120, ISSN 1537-5110, <https://doi.org/10.1016/j.biosystemseng.2020.03.020>.
- [16] Jayme Garcia Arnal Barbedo, "Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification", *Computers and Electronics in Agriculture*, Volume 153, 2018, Pages 46-53, ISSN 0168-1699, <https://doi.org/10.1016/j.compag.2018.08.013>.
- [17] Wang G, Sun Y, Wang J. Automatic image-based plant disease severity estimation using deep learning. *Computational intelligence and neuroscience*. 2017 Jul 5;2017.
- [18] Yang Lu, Shujuan Yi, Nianyin Zeng, Yurong Liu, Yong Zhang, "Identification of rice diseases using deep convolutional neural networks", *Neurocomputing*, Volume 267, 2017, Pages 378-384, ISSN 0925-2312, <https://doi.org/10.1016/j.neucom.2017.06.023>.
- [19] Prabira Kumar Sethy, Nalini Kanta Barpanda, Amiya Kumar Rath, Santi Kumari Behera, "Deep feature based rice leaf disease identification using support vector machine", *Computers and Electronics in Agriculture*, Volume 175, 2020, 105527, ISSN 0168-1699, <https://doi.org/10.1016/j.compag.2020.105527>.
- [20] K. Thenmozhi, U. Srinivasulu Reddy, "Crop pest classification based on deep convolutional neural network and transfer learning", *Computers and Electronics in Agriculture*, Volume 164, 2019, 104906, ISSN 0168-1699, <https://doi.org/10.1016/j.compag.2019.104906>.
- [21] Konstantinos P. Ferentinos, "Deep learning models for plant disease detection and diagnosis, *Computers and Electronics in Agriculture*", Volume 145, 2018, Pages 311-318, ISSN 0168-1699, <https://doi.org/10.1016/j.compag.2018.01.009>.
- [22] U. Shruthi, V. Nagaveni and B. K. Raghavendra, "A Review on Machine Learning Classification Techniques for Plant Disease Detection," *2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS)*, Coimbatore, India, 2019, pp. 281-284, doi: 10.1109/ICACCS.2019.8728415.
- [23] P. Sharma, P. Hans and S. C. Gupta, "Classification Of Plant Leaf Diseases Using Machine Learning And Image Preprocessing Techniques," *2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, Noida, India, 2020, pp. 480-484, doi: 10.1109/Confluence47617.2020.9057889.
- [24] Diseases-IRRI Rice Knowledge Bank. [4177](http://www.knowledgebank.irri.org/step-by-</p>
</div>
<div data-bbox=)

- step-production/growth/pests-and-diseases/diseases. Accessed 11 Nov 2021.
- [25] Udayananda G, Kumara P (2022), “ A comprehensive review on plant disease diagnosis and controlling using convolutional neural networks”, In: 2022 3rd international conference for emerging technology (INCET), pp 1–7. <https://doi.org/10.1109/INCET54531.2022.9824148>.
- [26] Singh S, Kumar I (2021) Rice plant infection recognition using deep neural network systems. Ceur-ws.org. [Online]. <http://ceur-ws.org/Vol-2786/Paper47.pdf>. Accessed 06 Sep 2022.
- [27] Bari BS et al (2021) A real-time approach of diagnosing rice leaf disease using deep learning-based faster R-CNN framework. PeerJ Comput Sci 7:e432. <https://doi.org/10.7717/peerj-cs.432>.
- [28] Purbasari IY, Rahmat B, Putra CS (2021) Detection of rice plant diseases using convolutional neural network. IOP Conf Ser Mater Sci Eng 1125(1):012021. <https://doi.org/10.1088/1757-899X/1125/1/012021>.
- [29] Deng R et al (2022),” Automatic diagnosis of rice diseases using deep learning”. Front. Plant Sci., vol. 12, 2021. [Online]. <https://www.frontiersin.org/article/10.3389/fpls.2021.701038>. Accessed 12 Feb 2022.
- [30] Wang Y, Wang H, Peng Z (2021) Rice diseases detection and classification using attention based neural network and Bayesian optimization. Expert Syst Appl 178:114770. <https://doi.org/10.1016/j.eswa.2021.114770>.
- [31] Anandhan K, Singh AS (2021) Detection of paddy crops diseases and early diagnosis using faster regional convolutional neural networks. In: 2021 international conference on advance computing and innovative technologies in engineering (ICACITE), pp 898–902. <https://doi.org/10.1109/ICACITE51222.2021.9404759>.
- [32] Krishnamoorthy N, Narasimha PLV, Pavan KCS, Subedi B, Abraha HB (2021) Rice leaf diseases prediction using deep neural networks with transfer learning. Environ Res 198:111275. <https://doi.org/10.1016/j.envres.2021.111275>.
- [33] Chander Nenavath, and Mummadi Upendra Kumar. "Metaheuristics with deep convolutional neural network for class imbalance handling with anomaly detection in industrial IOT environment" Journal of Theoretical and Applied Information Technology 101.10 (2023).
- [34] Mahalakshmi, C.V.S.S., Mridula, B., Shrivani, D, “Automatic Water Level Detection Using IoT”, Advances in Decision Sciences, Image Processing, Security and Computer Vision. Learning and Analytics in Intelligent Systems, 2020, vol 4. Springer, Cham.
- [35] D. Shrivani, Imtiyaz Khan, Amogh Deshmukh, Veeramalla Anitha, Masrath Saba, & Syed Shabbeer Ahmad. (2022). LISF: A Security Framework for Internet of Things (IoT) Integrated Distributed Applications. Journal of Advanced Zoology, 43(1), 584–594. <https://doi.org/10.53555/jaz.v43i1.1985>.
- [36] Dr. D. Shrivani, Anusha Padala, Image Processing: Human Facial Expression Identification using Convolutional Neural Networks, Turkish Online Journal of Qualitative Inquiry (TOJQI), Volume 11, June 2020: 823-830.
- [37] Ravi, Eslavath, and Mummadi Upendra Kumar. "Android malware detection with classification based on hybrid analysis and N-gram feature extraction." International Conference on Advancements in Smart Computing and Information Security. Cham: Springer Nature Switzerland, 2022. https://doi.org/10.1007/978-3-031-23095-0_13.
- [38] Eslavath, Ravi, and Upendra Kumar Mummadi. "ENSIC: Feature Selection on Android Malware Detection Attributes Using an Enhanced Non-Linear SVM Integrated with Cross Validator." International Journal of Intelligent Systems and Applications in Engineering 12.2 (2024): 495-504. <https://doi.org/10.5281/zenodo.10491837>.
- [39] Ravi, Eslavath, and Mummadi Upendra Kumar. " A Novel Mechanism for Tuning Neural Network for Malware Detection in Android Device." International Conference on Advancements in Smart Computing and Information Security. Cham: Springer Nature Switzerland, 2023.
- [40] Boddupally Janaiah, Dr. Suresh Pabboju, Artificial Intelligence Enabled Methods for Human Action Recognition using Surveillance Videos, International Journal on Recent and Innovation Trends in Computing and Communication, ISSN: 2321-8169 Volume: 11 Issue: 9 pp 3937-3945.
- [41] Needa Iffath, Upendra Kumar Mummadi, Fahmina Taranum, Syed Shabbeer Ahmad, Imtiyaz Khan, D. Shrivani; Phishing website detection using ensemble learning models. AIP Conf. Proc. 20 February 2024; 3007 (1): 090001. <https://doi.org/10.1063/5.0192754>.

- [42] Saba Noor Ayesha Khanum, Upendra Kumar Mummadi, Fahmina Taranum, Syed Shabbeer Ahmad, Imtiyaz Khan, D. Shravani; Emotion recognition using multi-modal features and CNN classification. AIP Conf. Proc. 20 February 2024; 3007 (1): 030001.
- [43] Surender Mogilicharla, Upendra Kumar Mummadi; Grain quality analysis from the image through the approaches of segmentation. AIP Conf. Proc. 20 February 2024; 3007 (1): 070001. <https://doi.org/10.1063/5.0192997>.
- [44] Surender Mogilicharla, Upendra Kumar Mummadi; The literature survey: Precision agriculture for crop yield optimization. AIP Conf. Proc. 20 February 2024; 3007 (1): 090005. <https://doi.org/10.1063/5.0192998>.
- [45] Imtiyaz Khan, A.Yashwanth Reddy, Maniza Hijab, Kotari Sridevi, Syed shabbeer Ahmad, D. Shravani (2024). "Secure and efficient data sharing scheme for multi-user and multi-owner scenario in federated cloud computing." Journal of Theoretical and Applied Information Technology, Vol. 102, No. 6, pp. [pages], 31st March 2024.