

# Synergizing Convolutional Neural Networks and Big Transfer Learning for Enhanced Rice Disease Identification



Tharani Pavithra P<sup>(D)</sup>, Baranidharan Balakrishnan<sup>\*(D)</sup>

Department of Computing Technologies, SRM Institute of Science and Technology, Kattankulathur 603203, India

Corresponding Author Email: baranidb@srmist.edu.in

Copyright: ©2024 The authors. This article is published by IIETA and is licensed under the CC BY 4.0 license (http://creativecommons.org/licenses/by/4.0/).

https://doi.org/10.18280/ts.410637

### Received: 13 January 2024 Revised: 21 April 2024 Accepted: 15 June 2024 Available online: 31 December 2024

### Keywords:

Convolutional Neural Network (CNN), rice leaf diseases, big transfer learning (BiT model), computational complexity

### ABSTRACT

This research addresses the pressing global challenge of securing rice production, a staple for over half of the world's population, amidst projections of a population surge to 9.7 billion by 2050 and a potential peak of 11 billion by 2100. Despite economic growth, food security remains precarious due to plant diseases impacting rice, intensified by evolving cultivation practices and climate shifts. Our innovative solution combines a Convolutional Neural Network (CNN) with Big Transfer (BiT) to enhance image classification for rice diseases. BiT, known for adaptability and outstanding performance on limited datasets, integrates seamlessly with a scalable CNN, yielding robust results across diverse tasks, even with sparse training data. Operating on a dataset with six classes, each containing only 20 training images, our hybrid CNN\_BiT model demonstrates remarkable efficacy. Achieving 100% accuracy, 93.75% precision, 93% recall, and a 94.2 F1 Score, this model surpasses recent counterparts in identifying rice leaf diseases. The integration of customized feature extraction from the CNN with BiT's advanced feature understanding results in a potent, resilient, and efficient model. These fusion holds promise for addressing image classification challenges in agriculture, showcasing its potential impact on global food security.

### **1. INTRODUCTION**

Over the next three decades, the global population is projected to increase by approximately 2 billion individuals, going from the current 7.7 billion to 9.7 billion by 2050. The projected population peak around 2100 is estimated to approach nearly 11 billion individuals [1]. Rice, being a staple grain for much of the world's population, serves as a vital source of calories for over half of the people on Earth [2]. The significant increase in rice productivity, attributed to the adoption of using enhanced cultivars, fertilization methods, and irrigation techniques during the Green Revolution, played a crucial role in boosting production and causing a sustained decrease in rice prices. This development has been a key driver in alleviating the reduction of poverty in Asia over the preceding centuries. In spite of these prior accomplishments, the ongoing growth of rice productivity remains critical for various reasons. The pace of rice harvest improvement has considerably experienced a deceleration in the recent years, failing to retain pace in tandem with the expansion of the population. This shortfall has resulted in shortages and elevated prices, particularly impacting impoverished communities. The 2008 food crisis and the spike in rice prices underscored the vulnerability of food security. Despite substantial economic growth in many regions, the stability of food security remains somewhat precarious [3]. Plant diseases constitute a significant hindrance and play a crucial role in realizing the maximum potential yield. Rice, being the most vital global food crop, faces threats from various diseases caused by fungi, bacteria, and viruses. Changes in agronomy techniques, diminished cultivar variation leading to a slender inherited foundation, and observable weather-related shifts have altered the changing patterns of rice ailments throughout the duration. The primary diseases have increasingly developed potent and expanded to novel regions. Numerous ailments previously deemed minor have now gained economic significance in various areas [4].

Timely disease detection is crucial for optimizing agricultural output. Innovation is imperative in addressing this need, and the application of advanced techniques is key to more accurate solutions. Globally, Artificial Intelligence (AI) methods are being employed in agriculture to enhance the efficiency of crop health monitoring. These AI techniques surpass human capabilities, offering precision in crop management.

Agriculturists are increasingly integrating artificial intelligence and machine learning techniques to elevate the effectiveness of crop management. This involves the identification and treatment of crops affected by various diseases and pest infestations. Emerging technologies such as machine learning, computer vision, satellite imaging, artificial intelligence, and data analysis play pivotal roles in disease management across a wide array of crops. Their adoption signifies a paradigm shift in agricultural practices, ensuring enhanced accuracy and efficacy in disease detection and mitigation [5]. It is evident that models based on machine learning hold significant promise in the agricultural sector, particularly in the realm of detecting plant and crop diseases. Nonetheless, several challenges persist, including the dilemma encountered during model training with extensive datasets. While utilizing large datasets can undoubtedly enhance the precision of predicting disease models which comes at the expense of increased computational demands. In particular, computational models or edge devices, which are typically smaller and less computationally intensive, struggle to achieve optimal performance under these circumstances. This challenge is compounded by the issue of covariate shift, stemming from disparities between the distributions of the training data used to develop the model and the data on which the model is subsequently applied. Addressing these challenges is essential to ensure the effective integration of machine learning in agriculture and to strike a balance between accuracy and computational efficiency [6].

This research paper tackles the aforementioned challenge by integrating a CNN as a more scalable approach to image classification, alongside BiT. BiT embodies a collection of image models that have been pre-learned to exhibit exceptional adaptability, delivering outstanding performance on new datasets, even when limited examples are available per class. The integrated approach not only yields significantly improved results across diverse tasks but also demonstrates robust transferability, proving effective even when confronted with datasets containing only a sparse number of images per class.

BiT, developed by Google Research, is a comprehensive transfer learning framework designed to offer a robust and versatile feature extractor. It can be fine-tuned for a range of downstream tasks. However, the computational demands of advanced image processing algorithms, particularly for tasks like leaf detection on large datasets or real-time monitoring applications, can be substantial. This could pose challenges in resource-limited environments or for tasks requiring swift processing. Thanks to its thorough pre-training, BiT demonstrates strong generalization capabilities to new tasks even with limited data. Even when provided with a small number of images, BiT maintains good classification performance. The fundamental concept of BiT involves pretraining a model on a comprehensive and varied dataset, such as ImageNet, employing self-supervised learning methods. Self-supervised learning entails training the model to predict specific data properties without relying on explicit human annotations. Through pre-training in this fashion, the model acquires complex representations of the input data, facilitating their transfer to other tasks with minimal fine-tuning. Therefore, in this study, we opt for BiT for rice disease detection with limited datasets. BiT does have certain drawbacks. However, when paired with CNN, it has the potential to alleviate or counteract some of these disadvantages.

- BiT focuses on learning high-level, abstract representations, while combining it with CNNs enhances the capture of fine-grained visual details. This fusion has the potential to produce more robust and comprehensive representations of the input data.
- ii) In contexts with limited labeled data, this integration could expedite fine-tuning processes, as CNNs excel at extracting task-specific features from restricted datasets.
- iii) CNNs are frequently employed in domain adaptation scenarios, where the target domain diverges substantially from the source domain. Integrating a CNN model with BiT can bolster the

model's capacity to adapt to new domains or datasets by harnessing the domain-specific features extracted by the CNN.

One of the drawbacks associated with CNNs is their reliance on the Spatial Invariance Assumption.

iv) The spatial invariance assumption within CNNs suggests that CNNs treat all regions of an image equally, regardless of their spatial position. In contrast, BiT captures extensive long-range dependencies and global context within the data, potentially mitigating the spatial invariance assumption inherent in CNNs.

So synergizing CNN with BiT innovatively for rice leaf disease detection with a limited dataset can lead to a more robust solution. By combining the strengths of both models, we can capitalize on the feature extraction capabilities of CNNs while harnessing the generalization power of BiT's extensive pre-training. This approach allows us to overcome the limitations of each model individually. Additionally, by leveraging the complementary aspects of CNNs and BiT, we can boost the model's capability to generalize to novel and data that has not been observed before, ultimately improving accuracy and reliability of rice leaf disease detection even with limited dataset sizes.

Incorporating a Big Transfer (BiT) model into a machine learning pipeline alongside a CNN also offers numerous advantages, particularly in the realm of image processing and classification tasks such as:

- 1. Utilizing Pre-Trained Features: BiT models undergo pre-training on extensive and diverse datasets, endowing them with a robust and versatile feature extraction capability. Through integration with a CNN, this combines the benefits of tailored feature extraction with the sophisticated, generalized feature comprehension inherent in the BiT model.
- 2. Improved Performance on Limited Datasets: In scenarios where training data is restricted, a model developed entirely from scratch may encounter challenges in acquiring effective features and achieving generalization. Nevertheless, the pre-trained elements of the BiT model offer a remedy by supplying a comprehensive, pre-learned set of features. Subsequently, the CNN layers can fine-tune these features tailored to the specific task, resulting in superior performance compared to a model trained solely on the constrained dataset.
- 3. Enhanced Generalization: The synergy between CNN and BiT enhances the model's skill to generalize effectively to novel, data that has not been observed before. The CNN layers adapt to the intricacies of the current task, and concurrently, the BiT model, with its comprehensive pre-training, contributes a broader understanding of image features. This collaborative approach results in heightened accuracy and robustness in real-world applications.
- 4. Streamlined Training Process: Employing a pretrained BiT model reduces the computational costs and time typically associated with training a model entirely from the ground up. Subsequently, the CNN layers demand comparatively less data and training time for fine-tuning, rendering the overall process

more efficient.

5. Mitigated Overfitting: In situations where the dataset is limited or lacks diversity, training a model from the ground up can frequently result in overfitting, wherein the model becomes excessively attuned to the training data but struggles with performance on new data. The incorporation of a pre-trained BiT model serves to alleviate this risk by infusing more generalized learning, enhancing the model's overall robustness.

The following sections of this document are structured as follows: Section 2 contains a review of relevant literature, exploring various research documents applying deep learningbased methods to improve the precision of identifying rice diseases. In Section 3, we outline the methodology, with subsection 3.1 detailing the dataset and 3.2 providing a comprehensive outline of the recommended approach for detecting crop ailments. Section 4 discusses experimental analyses, where thorough experiments are carried out and results are assessed through a comparative examination. This section includes an analysis of computational complexity and cost in 4.1, followed by a comparison of CNN\_BiT with other deep learning models in 4.2. To conclude, Section 5 presents the paper's conclusion.

### 2. LITERATURE REVIEW

Numerous research studies have used deep learning-based methods to boost the accurateness of identifying rice diseases. Ghazanfar Latif and colleagues have introduced a novel technique for precisely identifying and categorizing diseases that impact rice leaves. They employed Deep Convolutional Neural Networks (DCNN) and domain adaptation techniques to develop this innovative method. The method involves a modified knowledge transfer process built upon the VGG19 construction. Over this improved system, they successfully identify and diagnose six different disease modules that affect rice foliage [7].

Upadhyay and Kumar [8] introduced an efficient practice for sensing diseases in paddy plants by employing layers of Convolution. In order to enhance the accuracy of rice leaves ailment detection, the proposed model incorporates globalscale Otsu's thresholding for converting images into binary format, thereby efficiently removing incidental noise from them.

Chen et al. [9] conducted a thorough examination of deep learning-based methods, leading to the creation of a combination of convolutional networks premeditated to improve the model's capability to identify subtle characteristics in plant lesions. By smearing principles of ensemble learning, they combined three frivolous CNNs to create a unique network called "Es-MbNet".

Zhou et al. [10] introduced an original architecture termed as "residual-distilled transformer." Drawing inspiration from the initial achievements of employing transformers in computer vision tasks, they incorporated a distillation technique to extract and enhance weights and parameters from pre-trained vision transformer models. Afterwards, the refined features are input into a multi-layer perceptron (MLP) to generate predictions.

Sudhesh et al. [11] introduced a novel method for noticing paddy leaf ailments over the application of Decomposition of Dynamic Modes in conjunction with preprocessing driven by attention. They engrossed on four different groups of paddy leaf ailments, conducting four groups of trials to evaluate the success of ten pre-learned DCNN models.

Reference No.	Methodology	<b>Rice Diseases Detected</b>	Total No. of Images Used
[7]	Deep Convolutional Network VGG19	Bacterial leaf blight, Brown spot, Leaf blast, Leaf scald, Narrow brown,	1750
[8]	Otsu's global thresholding technique, CNN	spot. leaf smut, Brown spot, Bacterial, leaf blight. Rice blast,	4000
[9]	Lightweight CNNs-SE-MobileNet, Mobile-DANet, MobileNet V2	Brown spot, Leaf smut, Leaf scald, Stackburn, white tip	500
[10]	Vision Transformer, MLP	Bacterial blight, Brown spots, Blast and 2 Tungro.	805
[11]	DenseNet121, Dynamic Mode Decomposition, Random Forest, XceptionNet, SVM	Bacterial blight, Blast, Brown spot and Tungro.	3416
[12]	MobileNet, Augmented attention mechanism, Bayesian optimization method	Brown spot, Rice hispa damage, Rice leaf blast.	2370

Table 1. Analysis of literature review papers for rice leaf diseases detection

 Table 2. Analysis of BiT literature review papers across various applications

Reference No.	Methodology	Utilized for
[13]	big transfer learning	Skin Cancer Classification
[14]	big transfer learning	Fine Art Classification

Zhao et al. [14] delved into the effectiveness of convolutional layers for image classification tasks correlated to art. To assess how different hyperparameters affect model performance, diverse hyperparameter configurations were tested in the experiments. The researchers systematically compared the outcomes of five weight initializations across various tasks to understand the impact of transfer learning on the final results. Notably, refining the networks prelearned on a larger data collection demonstrated improved transferability. This observation underscores that the prior information acquired by implementations in practical scenarios is also applicable to the creative domain, a method referred to as BiT.

Arkah et al. [13] encounters challenges in supplying neural networks with ample data, primarily due to the costly annotation progression and the expertise essential. A commonly employed strategy to address this issue is big knowledge transfer, involving utilization of pre-trained models from ImageNet (such as VGG, GoogleNet, and ResNet50) on a substantial volume of unlabeled pictures of skin cancer initially. Subsequently, these models are adjusted on a smaller set of labeled skin images.

Wang et al. [12] introduced a new method named ADSNN-BO, this model is designed for identifying and categorizing rice diseases using images of rice leaves. Built upon the MobileNet architecture, the ADSNN-BO model integrates an attention mechanism and undergoes fine-tuning using Bayesian optimization. To ensure interpretability, the model employs feature analysis techniques like activation mapping and filter visualization. The results suggest that the attentionbased mechanism in the ADSNN-BO model improves the learning of pertinent features more effectively.

In the literature review, the research presented indicates that the image data size for classifying rice diseases is often large, as noted in Table 1. Our study, however, focuses on a smaller dataset consisting of only 20 images per class, totalling 120 images. This contrasts with the datasets typically used in the reviewed papers.

BiT, a recent innovation in transfer learning, presents promising potential for application in rice leaf disease detection, although further exploration is warranted. Recent studies [13, 14] as in Table 2 have showcased the efficacy of BiT in various applications beyond rice leaf disease detection. Our research endeavors involve integrating BiT with CNN, leading to the development of a robust model for accurately detecting rice leaf diseases, suggesting a fruitful avenue for future investigation.

### **3. METHODOLOGY**

### 3.1 Dataset

This paper addresses five significant rice diseases that pose considerable challenges to rice farmers worldwide, as they have the potential to substantially diminish both the yields and quality of rice crops.

1. Rice Blast: Rice blast, a fungal disease caused by Magnaporthe oryzae, poses a significant threat to rice cultivation globally. It can result in substantial yield losses and jeopardize food security in affected regions. Symptoms include lesions on leaves, stems, and panicles, ultimately leading to plant death. It is characterized by the infection of rice plants through spores, leading to the development of abnormalities or marks on various parts of plants, including leaves, leaf collars, panicles, stems, and stem nodes. The infectious agent retains the ability to generate spores for more than 20 days, presenting a significant menace to rice crops susceptible to this disease [15].

2. Brown Spot: Bipolaris oryzae, the fungus responsible for brown spot, flourishes in warm and humid conditions, which heightens the vulnerability of regions with tropical and subtropical climates to outbreaks of this disease. During the winter, the fungus persists in plant debris and soil. Its spores disperse via wind, water, or human actions. It impacts rice plants, affecting different plant components including leaves, glume, Seedlings, leaf sheaths, stems, and mature plant grains. Particularly, dark coffee-colored spots emerge on the panicle, and in severe cases, spot formation may occur. on the grains, ultimately leading to diminished yield and compromised milling quality [16].

3. Bacterial Sheath Blight: Bacterial sheath blight, triggered by Burkholderia gladioli bacteria, impacts various crops like rice and maize. It induces distinct symptoms such as watersoaked lesions on leaf sheaths, with rapid spread often culminating in the formation of a mucilaginous bacterial sheath. It usually becomes evident during the heading stage of the rice plant. In mature plants, water-soaked and translucent lesions develop near the edges of the leaves. Over time, these lesions increase in size and display a wavy margin. Eventually, the affected areas turn a straw-yellow color, spreading to cover the entire leaf. If left uncontrolled, this disease can lead to decreased crop yields and compromised quality [17].

4. Rice Tungro: Rice tungro is characterized by stunted growth, yellowing of leaves, and reduced tillering in infected plants. The disease stems from the collaborative impact of two viruses transmitted by leafhoppers. The viruses are Rice tungro bacilliform virus (RTBV) and Rice tungro spherical virus (RTSV). RTBV induces symptoms of stunting and chlorosis, while RTSV enhances the transmission efficiency of RTBV by the insect vectors. The manifestation of this disease includes symptoms such as leaf discoloration, stunted growth, reduced tiller numbers, and grains that are only partially filled. Tungro disease also affects certain wild rice relatives and other grassy weeds commonly found in rice fields. The resulting yield losses can be substantial, particularly in regions where the disease is widespread [18].

5. False Smut: It is caused by the pathogen Ustilaginoidea virens infiltrates the rice spikelet over a minor aperture prior to heading. It's characterized by the formation of fungal spore masses within the rice panicle, resembling grains but actually composed of fungal structures. This disease is more prevalent in warm, humid climates, fostering ideal conditions for fungal proliferation and spore generation. Elevated nitrogen fertilizer usage and waterlogged environments can further elevate the likelihood of disease onset. In the vegetative phase of rice development, the fungus takes hold by infiltrating the tissue at the rising tips of the tillers. The effects of rice false smut predominantly influence the quality, altering the visual

presentation of the crop. False smut can indeed lead to substantial reductions in yield and the quality of grains [19].



**Figure 1.** Comprehensive dataset: Six classes featuring healthy rice crops and five ailments affecting rice plants

In addition to these five categories of rice diseases, the dataset also incorporates images of healthy rice crops. In total, the dataset comprises six classes, with each class containing only 20 images for training purposes which is depicted in Figure 1. Current data comprising images of both diseased and healthy rice crops was gathered in the open fields of Melmaruvathur, Kavaraipettai, and Gummidipondi regions in South India during December 2021, the pictures were captured using Xiaomi and Redmi phones of 48 megapixels resolution. However, they were susceptible to noise and distortion due to the unregulated environmental conditions in the open field, along with fluctuations in lighting and backgrounds. Preserving reliability in image proportions poses a challenge, particularly when the data collection of imageries of fluctuating dimensions. To solve this issue, the training hired data expansion methods to change the size of all pictures to the specified input dimension of 224×224 for the CNN. Furthermore, image normalization was implemented to alleviate issues related to gradient propagation. Additionally, image enhancement methods such as Erosion, Dilation, Opening, and Closing were performed to augment areas of images with changing levels of intensity.

## **3.2** Synergizing BiT model within CNN framework for enhanced transfer learning in computer vision

The block diagram in Figure 2 illustrates the overarching process outlined in this research article. Elaboration on the procedural specifics will be provided in the subsequent subsections.

### 3.2.1 Convolutional Neural Network

Convolution represents a mathematical operation performed on two functions with real-valued arguments. This operation is commonly represented by an asterisk (\*). It is essential that one of the functions, denoted as w, be a valid probability density function to ensure that the output represents a weighted average. Additionally, w must be zero for all negative arguments to prevent looking into the future, a capability beyond our practical reach.

In the context of CNNs, the first function x involved in the convolution is often called the input, while the second function (in this case, w) is referred to as the kernel. The result of the convolution is termed the feature map. It is represented as

$$s(t) = (x * w)(t)$$
 (1)

Convolutional networks have the capacity to produce outputs that extend beyond traditional class labels in classification tasks or numerical values in regression tasks. Instead, they can generate high-dimensional, structured objects, typically in the form of tensors produced by standard convolutional layers. As an illustration, the model might generate a tensor S, where  $S_{i,j,k}$  represents the probability that the pixel at coordinates (j, k) in the input to the network belongs to class i. This capability enables the model to meticulously label each pixel in an image and create precise masks that accurately delineate the contours of individual objects [20].



Figure 2. Block diagram of proposed hybrid model framework of BiT with CNN

Fully connected layers, as their name suggests, establish extensive connections with preceding layers. Typically, these layers utilize activation functions like "sigmoid" or "softmax" in the final layer to generate predictions related to classes. Essentially, convolutional layers discern features extracted from input data, which are then condensed by successive aggregation layers. Utilizing the high-level features obtained, fully connected layers typically conduct the classification of input data into predefined categories during the final stages. Moreover, the classification layer not only categorizes the data but also extracts features crucial for both classification and detection tasks. Figure 3 illustrates the components of a typical CNN [21].



Figure 3. Components of a standard CNN layers

### 3.2.2 BiT-Big transfer model

Several recent advancements have been made in enhancing the training of deep neural networks. The objective is not to incorporate additional components or complexities, but rather to present a methodology that relies on the minimal number of techniques while achieving outstanding performance across various tasks. This approach is referred to as "Big Transfer" (BiT).

They conduct training on networks using three distinct sizes of data collections. The major, BiT-L, undergoes training concerning JFT-300M data collections, consisting of 300 million images with noisy labels. BiT is subsequently transferred to a variety of tasks, encompassing dimensions of data collections used for learning fluctuates from one instance per class to an over-all of one million examples. These assignments encompass a variety of datasets, including ImageNet's ILSVRC-2012, CIFAR-10/100, Oxford-IIIT Pet, Oxford Flowers-102 (including few-shot variants), and the VTAB-1k benchmark comprise 19 different datasets. Remarkably, BiT-L achieves effectiveness even when limited downstream data is available (see Table 3).

When BiT-M utilizes ILSVRC-2012 for pre-training (needed only once) and is additionally skilled on the public ImageNet-21k data collections, it delivers notable enhancements, providing cost-effective fine-tuning for subsequent tasks. BiT not only demands a concise refining practice for each freshchore nevertheless obviates the necessity for wide-ranging model fine-tuning on novel tasks. As an alternative, it employs an exploratory for configuring parameter tuning during transfer, proven effective across our diverse evaluation suite, is presented. It is crucial to highlight that the most significant aspect of BiT is its effectiveness, providing insights into the intricate relationship among dimensions, structure and learning parameters [22].

The essential elements identified for constructing a proficient network for transfer can be categorized into two groups: upstream components employed in pre-training at the upstream and transmission of downstream components for optimizing the process for an innovative task.

Pre-training at the upstream stage. The initial crucial factor is scale. In the realm of deep learning, it is widely acknowledged that larger networks exhibit superior performance in their designated tasks. Additionally, the relationship between larger datasets and the necessity for correspondingly expansive architectures is well-established. The efficacy of scale, particularly during the pre-training phase, is meticulously examined within the framework of transfer learning, extending to tasks characterized by a scarcity of data points. This analysis explores the intricate balance among the computational resources allocated (training time), the dimensions of the architecture, and the magnitude of the dataset. To explore this further, three BiT models undergo training on three significant data collections: ILSVRC-2012, which includes 1.3 million pictures (BiT-S); ImageNet-21k, comprising 14 million images (BiT-M); and JFT, boasting an impressive 300 million imageries (BiT-L).

Another crucial module involves Group Normalization (GN) and Weight Standardization (WS). Although Batch Normalization (BN) is a standard feature in the majority of advanced image models for stabilizing learning, it has been found that BN presents challenges for Big Transfer aimed at two specific motives. Firstly, during the learning of extensive models having batches of small size per device, BN performance demonstrates suboptimal or imposes synchronization costs across devices. Secondly, because of the need to update statistics in succession, BN is disadvantageous for transfer learning. In contrast, the pairing of GN and WS has shown enhancements in performance, particularly in situations where batches of small sizes are trained for ImageNet and COCO. This study showcases GN and WS proves beneficial for conducting training with sizable batch dimensions which also significantly influences the landscape of transfer learning.

Table 3. BiT: Cutting-edge transfer learning for computer vision

Datasets	ILSVRC-2012	CIFAR-10	CIFAR-100	Pets	Flowers
Models	Accuracy (in Percent)				
BiT-L	95	98	98	99	100
Generalist SOTA	90	97	95	96	96
Baseline (ILSVRC-2012)	75	95	80	95	90

<u>Transmission to task that downstream</u>. They suggest an economical refinement procedure that is applicable to a varied range of downstream jobs. Significantly, they have sidestepped the need for a costly exploration of tuning parameters for every original task and data volume. Instead, they experimented with a singular hyperparameter per task referred to as "BiT-HyperRule." This rule intelligently selects crucial tuning parameters based on a straightforward process with basic image resolution and the quantity of data instances.

Remarkably, during the downstream tuning phase, none of the following regularization forms were utilized: reducing weight decay to zero, reverting weight decay to initial parameters, or omitting dropout are all viable options. Remarkably, despite the significant size of the network, with BiT boasting 928 million parameters, its performance remains robust even without these techniques and their corresponding hyperparameters. This resilience persists even when transitioning to extremely small datasets. BiT adopts a streamlined training and fine-tuning framework, incorporating a select few components chosen with precision to strike a balance between complexity and performance. Leveraging extensive pre-training on a large scale is instrumental in achieving commendable performance [23].

3.2.3 Optimizinning in rice crop disease detection: integrating the BiT model into a CNN framework

Implementation utilizes TensorFlow, employing Keras as its high-level API for constructing and training deep learning models. After the datasets are stored, images undergo resizing to 384×384 pixels followed by cropping to 224×224.The dataset is partitioned within the realms of training and validation sets, with an 85% allocation for training and 15% for validation. As depicted in Figure 4, the information stream of the Hybrid CNN BiT model is depicted. is presented where various preprocessing functions are applied to images, encompassing resizing, cropping, and normalization. Training and validation datasets are meticulously prepared, incorporating diverse data augmentation techniques such as random flipping, along with batching. This meticulous preparation optimizes the data for seamless integration into the neural network during the training process.



The images undergo processing through a CNN block of layers. The BiT (Big Transfer) Model, denoted as BiT\_model, is employed as a pre-trained model sourced from TensorFlow Hub. This model takes the output of a custom CNN as its input, with the CNN serving as a feature extractor. The resulting outputs are then channeled into the BiT model, offering the flexibility for further fine-tuning as necessary. This model features the integration of the BiT model along with a Dense layer tailored for the classification task into NUM\_CLASSES categories. This deliberate integration strategy seeks to leverage the synergies between CNNs and transfer learning, capitalizing on the distinctive strengths inherent in both approaches.

The fusion of a BiT model with a CNN represents a potent alliance in tackling image classification tasks. This integration harnesses the benefits of both deep, pre-trained feature representations and tailored, task-specific feature extraction. Consequently, this methodology frequently yields models that exhibit heightened accuracy, resilience, and efficiency when confronted with diverse challenges in image classification.

### 4. EMPIRICAL INVESTIGATION AND FINDINGS

In the empirical setup, the proffered model utilized a data collection consists of six categories, covering five varieties of rice diseases (Blast, Brown spot, Bacterial Sheath Blight, Falsemut, Tungro), with the sixth category specifically for identifying healthy rice leaves. Each category contained only 20 images, resulting in a total of 120 images. These images underwent preprocessing, involving resizing to 384×384 pixels and subsequent cropping to 224×224.

The dataset was partitioned into training and validation sets, with an 85% allocation for training and 15% for validation purposes. The experimentation was directed utilizing functions from TensorFlow library and Keras on Google Collab Pro. The model experienced a 25-epoch training regimen, retaining the Stochastic Gradient Descent (SGD) optimizer. For the multi-class classification nature of the problem, Sparse Categorical Cross entropy was utilized.

To enhance the training process, the experiment incorporated the use of callbacks, specifically employing Early Stopping. This mechanism facilitated the cessation of training when the validation accuracy ceased to show improvement, contributing to a more efficient and optimized training procedure.

The research outcomes are assessed through evaluation metrics such as Accuracy, Precision, Recall, and F1 Score. Accuracy assesses the classifier's capacity to accurately categorize the complete dataset, considering both positive and negative instances. Precision, on the other hand, measures the ratio of correctly identified positive samples among all instances classified as positive, providing insight into the classifier's precision in identifying positive cases.

In contrast, recall measures the model's performance concerning the actual observations of a specific class. F1 Score, a machine learning evaluation metric, quantifies a model's overall accuracy by combining the precision and recall scores. This comprehensive approach provides a more nuanced understanding of the model's effectiveness in handling both positive and negative instances [24].

Figure 4. Information stream of the hybrid CNN BiT model

$$Accuracy = \frac{TN + TP}{TN + FN + TP + FP}$$
(2)

$$Precision = \frac{TP}{TP + FP}$$
(3)

$$\operatorname{Recall} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}$$
(4)

F1 Score = 
$$\frac{2^* \text{ (precision*recall)}}{(\text{precision+recall})}$$
 (5)

1. TP means True Positive: The instances correctly identified as positive by the classifier.

2. TN means True Negative: The instances correctly identified as negative by the classifier.

3. FN means False Negative: The instances that are genuinely positive but were incorrectly classified as negative by the classifier.

4. FP means False Positive: The instances that are genuinely negative but were incorrectly classified as positive by the classifier.

Leveraging the innovative hybrid CNN\_BiT model construction, the model achieves exceptional performance, boasting a flawless accuracy of 100%, as depicted in Figure 5. Additionally, the precision stands at an impressive 93.75%, with a recall rate of 93%, and an F1 Score reaching 94.2. Despite having only limited dataset comprising a minimal number of images (only 20 images for each class), these compelling results demonstrate the efficacy of the hybrid model. Notably, it surpasses the efficiency of the most recent models in precisely recognizing rice foliage ailments.



Figure 5. The accuracy chart for the hybrid CNN\_BiT model

Figure 6 presents the statistical analysis conducted for the detection of rice leaf disease, wherein confidence intervals (CIs) are utilized to estimate the range of values likely to contain the population parameter with a specified level of confidence. These intervals provide valuable insights into the precision and reliability of our findings. By incorporating CIs into our analysis, we gain a clearer understanding of the variability inherent in our data and the robustness of our conclusions. Moreover, they serve as a valuable tool for communicating the uncertainty associated with our estimates

to stakeholders and fellow researchers. Thus, the inclusion of confidence intervals enhances the comprehensiveness and credibility of our study's statistical assessment of rice leaf disease detection.



Figure 6. Performance metric utilization through confidence intervals

#### 4.1 Computational complexity and cost analysis

The utilization of BiT big for transfer learning alongside a CNN model displays promising outcomes, as evidenced by the provided metrics. Here's a breakdown of the comparison and feedback:

With a batch size of 64 and training over 25 epochs with 45 steps per epoch, the time taken for each epoch ranged from 996 to 1114 seconds. The total number of parameters amounted to 23,512,646, resulting in a computational complexity of 52,903,453,500 for the entire training process, based on the provided computations per epoch.

Given a TPU cost of \$10 per hour, the estimated cost for training the model for approximately 7.107 hours totals \$71.072. This demonstrates efficient training time, which can be advantageous, especially for scenarios involving large datasets or limited computational resources. Considering the substantial computational complexity of BiT big transfer learning, significant computational resources may be required. However, the estimated computational cost of \$71.072 appears reasonable, given the performance and efficiency benefits it offers.

In inference, BiT big transfer learning, when integrated with CNNs, showcases satisfactory performance metrics, positioning it as a strong contender against standard models.

### 4.2 Comparative analysis of CNN\_BiT and other DL models

The proposed method is compared to traditional neural network models such as DenseNet121, ResNet50, MobileNet, and BiT individually. The results, shown in Table 4, demonstrate the superior performance of the proposed approach. Figure 7 gives Comparative Performance Evaluation of various Deep learning Models versus Hybrid CNN\_BiT Model. When employing a consistent minimal set of 20 images for the execution of various deep learning models, namely ResNet50, MobileNet, and DenseNet121, the achieved accuracies are notably low at 60.52%, 25%, and 75%, respectively. However, the utilization of big transfer learning produces significantly improved results. Particularly, when Big Transfer is integrated with convolutional layers, it excels in feature extraction on both global and local scales.



 
 Table 4. Comparative analysis between the CNN\_BiT model and other deep learning models

Figure 7. Comparative performance evaluation of various deep learning models versus hybrid CNN\_BiT model

The fundamental concept behind BiT lies in transfer learning, wherein a model initially developed for a particular task is repurposed as the foundation for a second task. With BiT, one can take a pre-trained model and refine it on a specific dataset, even if that dataset is significantly smaller than the original training set. Integrating the BiT model into our pipeline harnesses the potency of pre-trained features, notably enhancing overall performance, especially for tasks with limited or insufficiently diverse training data that may impede the training of a robust model from scratch.

Through its integration with a CNN, this approach effectively amalgamates the strengths of tailored feature extraction (tailored to the specific task) with the BiT model's advanced and generalized understanding of features. This synergistic strategy often culminates in models that exhibit heightened accuracy, resilience, and efficiency when confronted with diverse challenges in image classification.

### 5. CONCLUSIONS

In conclusion, this research paper addresses the pressing challenge of timely disease detection in rice, a critical global food crop facing threats from various pathogens. The integration of a CNN with BiT in a hybrid model proves to be a highly effective and innovative solution, particularly in the domain of image processing and classification tasks. The advantages of this integrated approach are manifold. First and foremost, leveraging pre-trained features from the BiT model, which undergoes extensive and diverse pre-training, enhances the hybrid model's ability for robust and versatile feature extraction. This, when combined with the CNN's tailored feature extraction capabilities, results in a comprehensive understanding of image features that contributes to improved performance. It also demonstrates the practical benefits of this hybrid model, especially in scenarios with limited datasets.

The experimental results are compelling, with the hybrid model achieving a flawless accuracy of 100% and impressive precision, recall, and F1 Score metrics. Despite the challenging scenario of a dataset comprising only 20 images per class, the hybrid model surpasses the efficiency of the most recent models in precisely categorizing the ailments of rice crops. In practical terms, the streamlined training process, reduced computational costs, and mitigation of overfitting risks make the hybrid CNN\_BiT model a valuable tool for optimizing agricultural output. The research underscores the importance of incorporating advanced AI techniques in agriculture, particularly in the context of disease detection, to address evolving challenges and ensure global food security.

Future work entails validating the robustness and generalizability of the model through extensive testing on diverse datasets and under various environmental conditions. This validation process will provide insights into the model's performance across different domains and its ability to adapt to unseen scenarios. Additionally, exploring transfer learning techniques to fine-tune the model on specific tasks and evaluating its effectiveness in real-world applications will be crucial steps in further enhancing the model's utility and reliability.

### REFERENCES

- [1] United Nations. United Nations Official Website. http://www.un.org/, accessed on September 04 2024.
- [2] Chen, J., Zhang, D., Zeb, A., Nanehkaran, Y.A. (2021). Identification of rice plant diseases using lightweight attention networks. Expert Systems with Applications, 169: 114514. https://doi.org/10.1016/j.eswa.2020.114514
- [3] Pandey, S., Byerlee, D., Dawe, D., Dobermann, A., Mohanty, S., Rozelle, S., Hardy, B. (2010). Rice and global climate change. In Rice in the Global Economy: Strategic Research and Policy Issues for Food Security. https://www.researchgate.net/publication/258883000.
- [4] Laha, G.S., Singh, R., Ladhalakshmi, D., Sunder, S., Prasad, M.S., Dagar, C.S., Babu, V.R. (2017). Importance and management of rice diseases: A global perspective. Rice Production Worldwide, Springer, Cham, pp. 303-360. https://doi.org/10.1007/978-3-319-47516-5 13
- [5] Aggarwal, S., Suchithra, M., Chandramouli, N., Sarada, M., Verma, A., Vetrithangam, D., Pant, B., Ambachew Adugna, B. (2022). Rice disease detection using artificial intelligence and machine learning techniques to improvise agro-business. Scientific Programming, 2022(1): 1757888.

https://doi.org/10.1155/2022/1757888

- [6] Wani, J.A., Sharma, S., Muzamil, M., Ahmed, S., Sharma, S., Singh, S. (2022). Machine learning and deep learning based computational techniques in automatic agricultural diseases detection: Methodologies, applications, and challenges. Archives of Computational methods in Engineering, 29(1): 641-677. https://doi.org/10.1007/s11831-021-09588-5
- [7] Latif, G., Abdelhamid, S.E., Mallouhy, R.E., Alghazo, J., Kazimi, Z.A. (2022). Deep learning utilization in agriculture: Detection of rice plant diseases using an improved CNN model. Plants, 11(17): 2230.

https://doi.org/10.3390/plants11172230

- [8] Upadhyay, S.K., Kumar, A. (2022). A novel approach for rice plant diseases classification with deep convolutional neural network. International Journal of Information Technology, 14(1): 185-199. https://doi.org/10.1007/s41870-021-00817-5
- [9] Chen, J., Zeb, A., Nanehkaran, Y.A., Zhang, D. (2023). Stacking ensemble model of deep learning for plant disease recognition. Journal of Ambient Intelligence and Humanized Computing, 14(9): 12359-12372. https://doi.org/10.1007/s12652-022-04334-6
- [10] Zhou, C., Zhong, Y., Zhou, S., Song, J., Xiang, W. (2023). Rice leaf disease identification by residualdistilled transformer. Engineering Applications of Artificial Intelligence, 121: 106020. https://doi.org/10.1016/j.engappai.2023.106020
- [11] Sudhesh, K.M., Sowmya, V., Kurian, S., Sikha, O.K. (2023). AI based rice leaf disease identification enhanced by dynamic mode decomposition. Engineering Applications of Artificial Intelligence, 120: 105836. https://doi.org/10.1016/j.engappai.2023.105836
- [12] Wang, Y., Wang, H., Peng, Z. (2021). Rice diseases detection and classification using attention based neural network and Bayesian optimization. Expert Systems with Applications, 178: 114770. https://doi.org/10.1016/j.eswa.2021.114770
- [13] Arkah, Z.M., Al-Dulaimi, D.S., Khekan, A.R. (2021). Big transfer learning for automated skin cancer classification. Indonesian Journal Electrical Engineering Computer Science, 23: 1611-1619. http://doi.org/10.11591/ijeecs.v23.i3.pp1611-1619
- [14] Zhao, W., Jiang, W., Qiu, X. (2022). Big transfer learning for fine art classification. Computational Intelligence and Neuroscience, 2022(1): 1764606. https://doi.org/10.1155/2022/1764606
- [15] Liu, X., Zhang, Z. (2022). A double-edged sword: Reactive oxygen species (ROS) during the rice blast fungus and host interaction. The FEBS Journal, 289(18): 5505-5515. https://doi.org/10.1111/febs.16171
- [16] Terensan, S., Fernando, H.N.S., Silva, J.N., Perera,

S.C.N., Kottearachchi, N.S., Weerasena, O.J. (2022). Morphological and molecular analysis of fungal species associated with blast and brown spot diseases of Oryza sativa. Plant Disease, 106(6): 1617-1625. https://doi.org/10.1094/PDIS-04-21-0864-RE

- [17] Singh, P., Mazumdar, P., Harikrishna, J.A., Babu, S. (2019). Sheath blight of rice: A review and identification of priorities for future research. Planta, 250: 1387-1407. https://doi.org/10.1007/s00425-019-03246-8
- Singh, S.P., Pritamdas, K., Devi, K.J., Devi, S.D. (2023). Custom convolutional neural network for detection and classification of rice plant diseases. Procedia Computer Science, 218: 2026-2040. https://doi.org/10.1016/j.procs.2023.01.179
- [19] Wikipedia. Ustilaginoidea virens. https://en.wikipedia.org/wiki/Ustilaginoidea\_virens, accessed on September 04 2024.
- [20] Goodfellow, I., Bengio, Y., Courville, A. (2017). Deep Learning. The MIT Press: Cambridge, Massachusetts, London, England.
- [21] Tugrul, B., Elfatimi, E., Eryigit, R. (2022). Convolutional neural networks in detection of plant leaf diseases: A review. Agriculture, 12(8): 1192. https://doi.org/10.3390/agriculture12081192
- [22] TensorFlow. BigTransfer (BiT): A State-of-the-Art Transfer Learning Model for Computer Vision. TensorFlow Blog. https://blog.tensorflow.org/2020/05/bigtransfer-bitstate-of-art-transfer-learning-computer-vision.html, accessed on May. 28, 2020.
- [23] Kolesnikov, A., Beyer, L., Zhai, X., Puigcerver, J., Yung, J., Gelly, S., Houlsby, N. (2020). Big transfer (BiT): General visual representation learning. In Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, pp. 491-507. https://doi.org/10.48550/arXiv.1912.11370
- [24] Labelf AI. What is Accuracy, Precision, Recall, and F1score. https://www.labelf.ai/blog/what-is-accuracyprecision-recall-and-f1-score, accessed on September 4 2024.