International Journal of Computing and Digital Systems
2026, VOL. 19, NO. 1, 1–11

http://dx.doi.org/10.12785/ijcds/1571160366

Transformative Approaches to Wheat Disease Detection: An Efficient Deep Learning Framework for Enhanced Crop Resilience

Seifeldin A. Ismail[1], Yahia M. Anas[1] and Nermin K. Negied[1]

[1]*Faculty of Computer Science, University for Modern Sciences and Arts (MSA), Giza, Egypt*

Received 1 June 2025, Revised 27 June 2026, Accepted 30 June 2026

Abstract: Wheat is the third most consumed grain globally and a critical staple food for billions of people; yet its production is severely impacted by fungal diseases that cause annual yield losses of 15% to 20%, posing a direct threat to global food security and agricultural economics. Traditional methods of disease detection, reliant on manual field inspections, are inefficient, labour-intensive, require expertise that is not always available, and are prone to subjectivity—motivating the need for automated, reliable, and scalable solutions for early and accurate wheat disease identification. This study presents a controlled comparative evaluation of four CNN-based pre-trained architectures—DenseNet121, ResNet50, VGG19, and InceptionV3—for wheat disease classification using the Large Wheat Disease Classification Dataset (LWDCD2020), employing targeted data augmentation strategies to improve model robustness and generalizability without overfitting. The results reveal that InceptionV3 achieved the highest performance with a Macro F1-score of 96.75%, outperforming all other proposed models and existing approaches in the literature, while requiring only ten training epochs compared to 50–200 epochs in prior studies. The introduction of the Revolutions Per Inference (RPI) metric further quantifies the efficiency–accuracy trade-off, demonstrating InceptionV3's practical suitability for deployment in resource-constrained agricultural environments.

Keywords: Wheat disease classification, CNNs, transfer learning, precision agriculture, InceptionV3, ResNet50, VGG19

1. INTRODUCTION

Wheat ranks as the third most consumed grain worldwide, following maize and rice. The United Nations Food and Agriculture Organization (UNFAO) reports that one in ten people globally experience severe malnutrition because of food scarcity. In numerous developing nations, the yield of grains per hectare is significantly lower than that in developed countries [1]. This is attributed to several factors, but primarily to diseases, which can account for annual yield losses of 15% to 20%. Major wheat diseases responsible for these losses consist of rusts, blotches, and head blight/scab, all of which are caused by fungal pathogens [2].

Conventional detection relies on trained personnel conducting manual field inspections—a process that is highly inefficient, time-consuming, and inherently subjective, while being critically dependent on the availability of domain expertise that is often scarce in rural settings [3]. Traditional machine learning approaches, while partially automating this process, require hand-crafted feature engineering and demonstrate limited generalizability across diverse field conditions. These limitations define a clear research gap: the need for automated, scalable, and computationally efficient deep learning frameworks capable of accurately classifying wheat diseases directly from images, without reliance on manual intervention or domain-specific feature design. An automated detection system allows users to quickly identify crop diseases without specialized knowledge, while supporting the implementation of cost-effective fungicide solutions.

The primary objectives of this study are: (i) to evaluate and compare the performance of four widely used pre-trained CNN architectures—InceptionV3, DenseNet121, ResNet50, and VGG19—on the LWDCD2020 benchmark dataset; (ii) to apply and assess targeted data augmentation strategies to enhance model generalizability and robustness; and (iii) to benchmark the proposed models against state-of-the-art approaches using both classification performance and computational efficiency metrics. The key contributions of this work are: (a) the first rigorous, controlled comparative evaluation of these four architectures on LWDCD2020

E-mail address: seifeldin.amr3@msa.edu.eg, yahia.mohamed7@msa.edu.eg, nnegied@msa.edu.eg

under identical training conditions; (b) demonstration that InceptionV3 achieves state-of-the-art classification performance (Macro F1-score of 96.75%) with only ten training epochs—significantly fewer than any prior work on the same dataset; and (c) the introduction of the Revolutions Per Inference (RPI) metric as a practical efficiency–accuracy benchmark for agricultural AI deployment contexts.

Deep learning methods have proven effective in diagnosing diseases across various crops through image analysis [4], [5], [6]. These methods include classification, in which CNNs process raw pixel data represented as three-channel matrices to learn hierarchical feature representations, and object detection, which additionally localises disease instances within an image. In this study, we deploy and compare several CNN-based classification models to identify the most effective and efficient architecture for wheat disease identification.

The remainder of this paper is organised as follows: Section 2 surveys the state-of-the-art approaches in this area. Section 3 describes the dataset, data preparation, and proposed methodology. Section 4 presents and discusses the experimental results. Section 5 concludes the paper and outlines future research directions.

2. RELATED WORK

Intelligent systems and automation have been widely used in the field of agriculture since the early emergence of AI and Machine Learning [4], [5], [6], [7], [8], [9], [10]. In this section, the focus of the literature review will be on the detection and the classification of wheat diseases. An automated disease detection system would enable users to easily identify diseases in their crops without requiring specialists or experts. Deep learning techniques have demonstrated significant potential in identifying diseases across different crops through image analysis [11], [12], [13]. There are different deep learning techniques that could be used including classification and object detection, each having their own approaches and methodologies. In the following paragraphs we will summarize the previous work done in the field of automatic detection and classification of wheat diseases.

M. Long et al.[14] in 2022 deployed deep neural networks to classify wheat diseases among four foliar diseases, namely yellow rust, brown rust, powdery mildew and Septoria leaf blotch. The authors trained a deep learning model, and they confirmed that their model succeeded in classifying the wheat disease with an accuracy that reached 97%.

Many researchers studied the pretrained models' capability in addressing the classification of wheat diseases. Goyal et al.[1] built their own CNN model and trained it from scratch and compared the results of their proposed CNN model and famous CNN architectures like ResNet50 and VGG16. The proposed CNN model had the highest scores having a testing accuracy of 97.8%. Abdalla et al.[15] applied different pretrained models and compared them. They collected their dataset by combining multiple sources and trained various deep learning models, including EfficientNetB0, EfficientNetB1, DenseNet161, DenseNet169, and ResNet152, to evaluate and compare their performance. DenseNet161 had the best scores with 98.42% precision and 98.48% recall; however, this approach required substantially higher computational resources and was not evaluated on a standardized benchmark. Samundiswary et al.[16] proposed a deep learning-based method for classifying wheat diseases using ResNet152, achieving an F1-score of 85.5% but requiring 200 training epochs, indicating significant computational overhead. Islam et al.[17] presented a lightweight framework that optimally utilizes features, combining a pretrained CNN with a spatial attention (SA) mechanism to effectively distinguish between healthy and unhealthy patterns. This proposed framework leverages its capacity to identify critical areas impacted by wheat diseases through a streamlined, attentive, and optimized network. They utilized their approach on two distinct datasets, being the LWDCD2020 dataset and their custom-made WD5CC dataset. They evaluated various pretrained models with and without SA across both datasets, with EfficientNet-B3 featuring SA attaining an F1 score of 92.87% on the WD5CC dataset and 93.52% on the LWDCD2020 dataset, which was the highest compared to all other CNN models. Nawaz et al.[18] proposed a multi-stage Ensemble Learning (EL) strategy based on Convolutional Neural Networks (CNN), which employs multiple parallel CNN models alongside a bagged EL method to identify wheat leaf diseases. The different stages of the EL strategy utilized a multi-tiered structure to improve feature extraction and understand intricate data patterns. The suggested approach attained an accuracy of 86.78%, 98.28%, and 99.16% across the three publicly available datasets. Maqsood et al.[19] introduced a new approach that utilizes Transfer Generative Adversarial Networks to enhance current datasets, addressing issues related to class imbalance and limited data availability. The authors presented a tailored architecture of Vision Transformers (ViT), in which the feature vector is generated by combining features derived from the custom ViT and Graph Neural Networks. Additionally, this paper introduces an ensemble classifier based on Model Agnostic Meta Learning (MAML) to achieve precise classification, achieving an F1-score of 97.9% on the public dataset.

Some researchers tried applying different techniques like using segmentation to pinpoint specific diseases. Anwar et al.[3] applied a UNet semantic segmentation model to detect the leaf rust. The authors collected their own dataset with masks of the infected leaves achieving an F1 score of 55.7%. Zhang et al.[20] segmented the Wheat Fusarium Head disease. They constructed their own dataset with masks for the wheat ear. The image was first segmented using a segmentation model to get the wheat ear, then an IABC-K-PCNN was used to segment the image again to only output the parts of the wheat ear that are diseased. Their method achieved a segmentation accuracy of 98%.

Some other researchers used object detection to localize the diseased plant. Etaati et al.[21] applied two different techniques on the LWDCD2020 dataset and the Global Wheat Head Detection Dataset. They trained the Yolo-v4 on the Global Wheat Head Detection Dataset using various pre-trained weights, such as the COCO and GWHD weights, to locate the wheat heads. They then classified the localized disease using the best pre-trained model with the best result which was the VGG19. The object detection stage achieved a mean Average Precision of 91%, while the VGG19 classifier reached an F1 score of 95%. Hong et al.[22] used Yolo-V4 for detection on a dataset with images of Fusarium head blight disease. They changed the backbone of the Yolo-V4 from Darknet53 to MobileNetv2 to be more lightweight for better usage in real-time applications like UAV monitoring. Despite the reduction in model complexity, the modified YOLOv4-MobileNetV2 achieved a detection accuracy of 93.62%, while also reducing the average detection time by 0.014 second. Yao et al.[23] included a total of 3,622 original photos for the following classes: wheat stem rust, wheat smut, brown rust, yellow rust, and healthy wheat, which were captured by photographing wheat in an agricultural setting over different time frames. The proposed detection method, YOLO-Wheat, merges the essential core components of the original YOLOv8s with a new small target detection layer, which enhances the feature extraction process and boosts the precision of small target defect identification. To enhance the extraction of disease characteristics and broaden the sensory field for the input features, the algorithm utilizes the new C2f-DCN module along with the SCNet attention mechanism. As a result, YOLO-Wheat achieved a mean Average Precision of 93.23%, representing a 12% improvement over the original YOLOv8s. Hassan et al.[24] proposed an alternative approach for classifying the disease. The authors utilized object detection to find the location of the leaf before applying augmentation to position the leaf in a way that allows a rectangular bounding box to minimize the background as much as possible. Then, they segmented the image so that only the leaf is visible and used it as input for a classification model to classify the disease. They compared several segmentation models, with U-Net achieving the best performance, obtaining an Intersection over Union (IoU) of 0.9563. For classification, their method was evaluated against various CNN architectures, with the Swin Transformer attaining the highest accuracy of 95.8%.

The surveyed literature reveals several recurring limitations. First, many high-performing models require 50–200 training epochs, making them computationally prohibitive for development on affordable agricultural hardware. Second, few studies simultaneously optimize for both accuracy and inference efficiency, despite the practical importance of this trade-off in field settings. Third, direct comparisons using the same benchmark dataset are rare, complicating objective performance assessment [25], [26], [27], [28], [29], [30], [31], [32], [33], [34]. The present study addresses these gaps by conducting a controlled comparative evaluation on LWDCD2020, training all models for only ten epochs, and introducing the RPI metric to quantify the efficiency–accuracy balance.

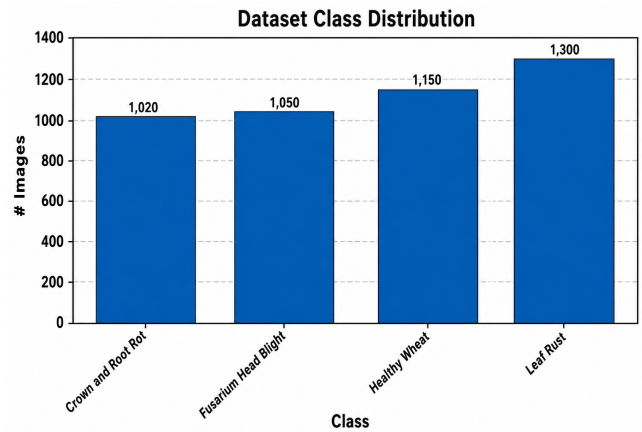


Figure 1. LWDCD2020 Training set Class Distribution

3. MATERIALS & METHODS

This work proposes several pre-trained CNN models for wheat disease classification and compares their performance. The Large Wheat Disease Classification Dataset (LWDCD2020) [1] was used in this work to train and test the proposed models. The following sub-sections describe the dataset, the data preprocessing and augmentation steps, and the architectures of the proposed models.

A. Dataset

LWDCD2020 was used for training the models proposed in this work. A data augmentation step was held to add variety to the model and avoid overfitting. The LWDCD2020 dataset originally contains 12,000 images for 10 different classes, but the authors only published 4504 images for four classes from the original dataset. The four classes are: Healthy, Leaf Rust, Crown and Root Rot, and Fusarium Head Blight. The dataset was divided into 70% training, 15% validation, and 15% testing. The class distribution is balanced for training, testing and validation data. The balanced class distribution guarantees no bias that may be caused by class imbalance. The class distribution for training, testing and validation of the dataset are illustrated in Figure 1, Figure 2, and Figure 3 respectively. Sample images for the dataset are shown in Figure 4.

B. Data Preprocessing

To ensure the best model performance, one method to avoid overfitting through regularization is data augmentation [35]. Geometric transformation was selected in this work as it involves moving pixels from their original locations to new ones while maintaining the pixel values. These transformations are used to modify the training data in ways that better represent changes in appearance that occur in the actual world [36]. Geometric augmentations (rotation, shear) preserve class labels if the transformation T_Θ is

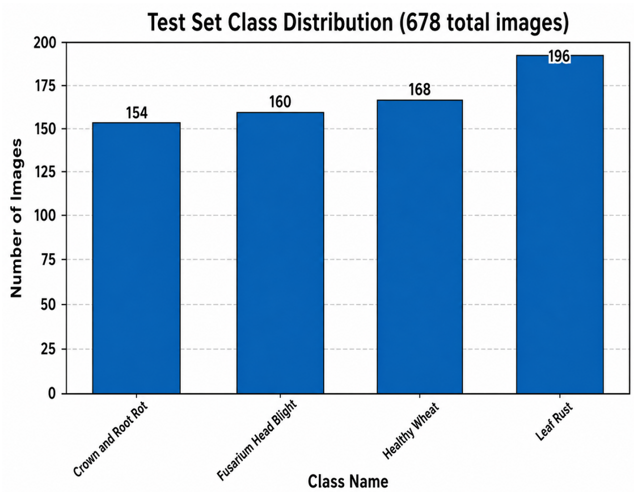


Figure 2. LWDCD2020 Test set Class Distribution

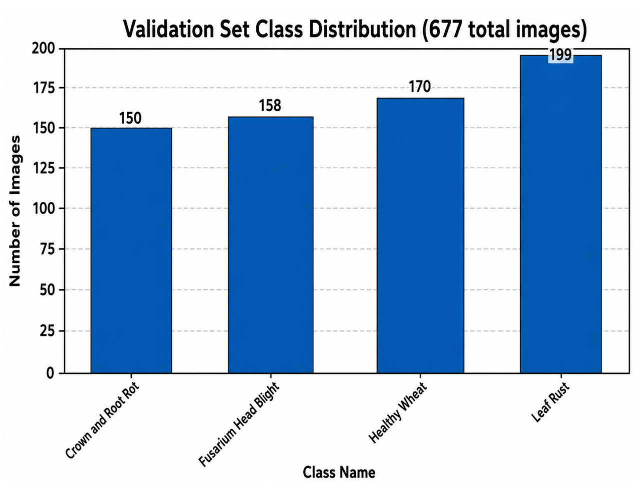
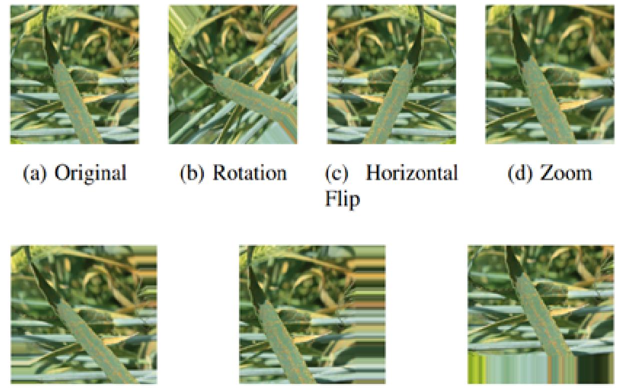


Figure 3. LWDCD2020 Validation set Class Distribution

(a) Healthy (b) Leaf Rust

(c) Crown & Root Rot (d) Fusarium Head Blight

Figure 4. Example images from each class in the LWDCD2020



(a) Original (b) Rotation (c) Horizontal Flip (d) Zoom

(e) Shear (f) Width Shift (g) Height Shift

Figure 5. Example augmentations applied to an image from the LWDCD2020 Data

diffeomorphic, i.e. for a transformation T with parameters θ the augmented data satisfies:

$$P(y|\tilde{x}) = P(y|x) \quad \text{if } |J_{T_\theta} - I|_F < \varepsilon \tag{1}$$

In the case of the LWDCD2020 data, only for training data, we applied different transformations including horizontal/vertical shifts (±20%), rotations (±40°), horizontal flipping, shear mapping (0.2 radians), and zoom (±20%). For training, testing and validation, the pixel values were normalized to $[0, 1]$ and all images were resized to 224×224. Figure 5 shows an example of these augmentations in a sample image.

C. Methodology

This study employed a transfer learning approach to evaluate and compare the performance of multiple pre-trained convolutional neural network (CNN) models for wheat disease classification. Transfer learning was selected as the primary methodology due to its ability to leverage knowledge gained from large-scale image recognition tasks (typically trained on datasets like ImageNet) and adapt it to our specific domain of agricultural disease detection. This approach offers significant advantages, including reduced training time and improved performance compared to training models from scratch, particularly when working with limited annotated data.

Four well-established CNN architectures were carefully selected for this comparative analysis based on their proven performance in various computer vision tasks:

1) **DenseNet121**: This architecture utilizes dense connectivity patterns between layers, where each layer receives additional inputs from all preceding layers and passes its own feature maps to all subsequent layers. This design promotes feature reuse and has demonstrated strong performance in medical image analysis and fine-grained classification tasks.
2) **InceptionV3**: Characterized by its innovative use of

parallel convolutional operations at multiple scales within inception modules, this model efficiently captures features at varying levels of abstraction. Its design is particularly effective for recognizing patterns that may appear at different scales in input images.

3) **VGG19**: As a classical CNN architecture, VGG19 employs a uniform structure of small (3×3) convolutional filters stacked in increasing depth. While computationally intensive, its simplicity and depth have made it a reliable baseline for many image classification problems.
4) **ResNet50**: This model introduced the concept of residual learning through skip connections, which help mitigate the vanishing gradient problem in deep networks. Its 50-layer depth provides substantial representational capacity while remaining computationally feasible.

The complete system architecture, illustrated in Figure 6, follows a standardized pipeline for transfer learning applications. Each pre-trained model was adapted for our specific classification task by replacing the original fully connected layers with new layers tailored to our four-class wheat disease problem. The models were initialized with weights pre-trained on ImageNet, allowing them to leverage low-level feature detectors (like edge and texture filters) that are generally useful across most visual recognition tasks.

To ensure fair comparison, all models were subjected to identical training conditions:

- Identical input preprocessing pipelines
- The same data augmentation strategies
- Consistent batch sizes and learning rate schedules
- Matching hardware and software environments

This controlled experimental setup allows us to make meaningful comparisons between the architectures and identify which model offers the best balance of accuracy and computational efficiency for our specific application in agricultural disease detection. The training process monitored multiple performance metrics, including not just overall accuracy but also per-class precision and recall, to thoroughly evaluate each model's capabilities.

1) DenseNet121

DenseNet121 is part of the Dense Convolutional Network family introduced in 2017 [37]. The main intuition behind DenseNet is to link each layer to all preceding layers in a feed-forward manner, which improves feature reuse and minimizes the total number of parameters. Comprising 121 layers, DenseNet121 is structured into dense blocks interconnected by transition layers. The model has approximately 8 million parameters, making it very efficient in comparison to other deep networks that achieve similar levels of accuracy. Figure 7 demonstrates the high-level architecture of the DenseNet121 model.

2) ResNet50

ResNet50 is a member of the Residual Networks (ResNet) family that was introduced by Microsoft Research in 2015 [38]. It tackles the degradation problem found in deep networks by implementing identity shortcut connections, referred to as residual connections or skip connections that enhance gradient flow during training. Comprising 50 layers, ResNet50 manages about 25 million parameters and delivers high performance while mitigating vanishing gradient issues. The residual block in ResNet50 and the architecture of ResNet50 are shown in Figures 8 and 9 respectively.

3) VGG19

The VGG19 architecture developed by the Visual Geometry Group (VGG) in 2014 increases the depth of traditional CNNs by utilizing a series of small 3×3 convolutional filters. VGG19 incorporates 16 convolutional layers, 5 max-pooling layers, and 3 fully connected layers. All convolutional layers employ ReLU activation, while a SoftMax function is used in the output layer for classification purposes. The overall count of parameters in VGG19 is around 144 million [39]. Figure 10 shows the high-level architecture of VGG19.

4) InceptionV3

InceptionV3 is a deep convolutional neural network built upon the original Inception models by incorporating factorized convolutions, auxiliary classifiers, and label smoothing to boost performance and lower computational costs. The structure consists of a series of Inception modules and has around 23 million parameters and is recognized for its efficiency and precision in large-scale image classification tasks [40]. InceptionV3's efficiency stems from its parallel filter pathways that allow managing a relatively small number of parameters. For an input image x the output y can be represented as follows:

$$y = \mathrm{Concat}[\mathrm{Conv}_{1\times1}(x), \mathrm{Conv}_{3\times3}(x), \mathrm{Conv}_{5\times5}(x), \mathrm{Pool}(x)] \quad (2)$$

Where *Concat* merges features at different scales.

This architecture efficiently manages the diversity-parameters ratio, as it maximizes information diversity while minimizing parameters, proven by the Feature Diversity Bound:

$$\mathcal{D}(y) \geq \sum_{i=1}^{4} \mathcal{D}(y_i) - \mathcal{O}(\text{Receptive Field Size}) \quad (3)$$

Where $\mathcal{D}$ denotes feature diversity. This allows InceptionV3 to obtain a higher F1-score with fewer epochs. Figure 11 demonstrates the high-level architecture of InceptionV3.

4. RESULTS AND DISCUSSION

This paper deployed and compared several pre-trained models to identify the best one for wheat disease classification. The training hyperparameters were as follows: 10

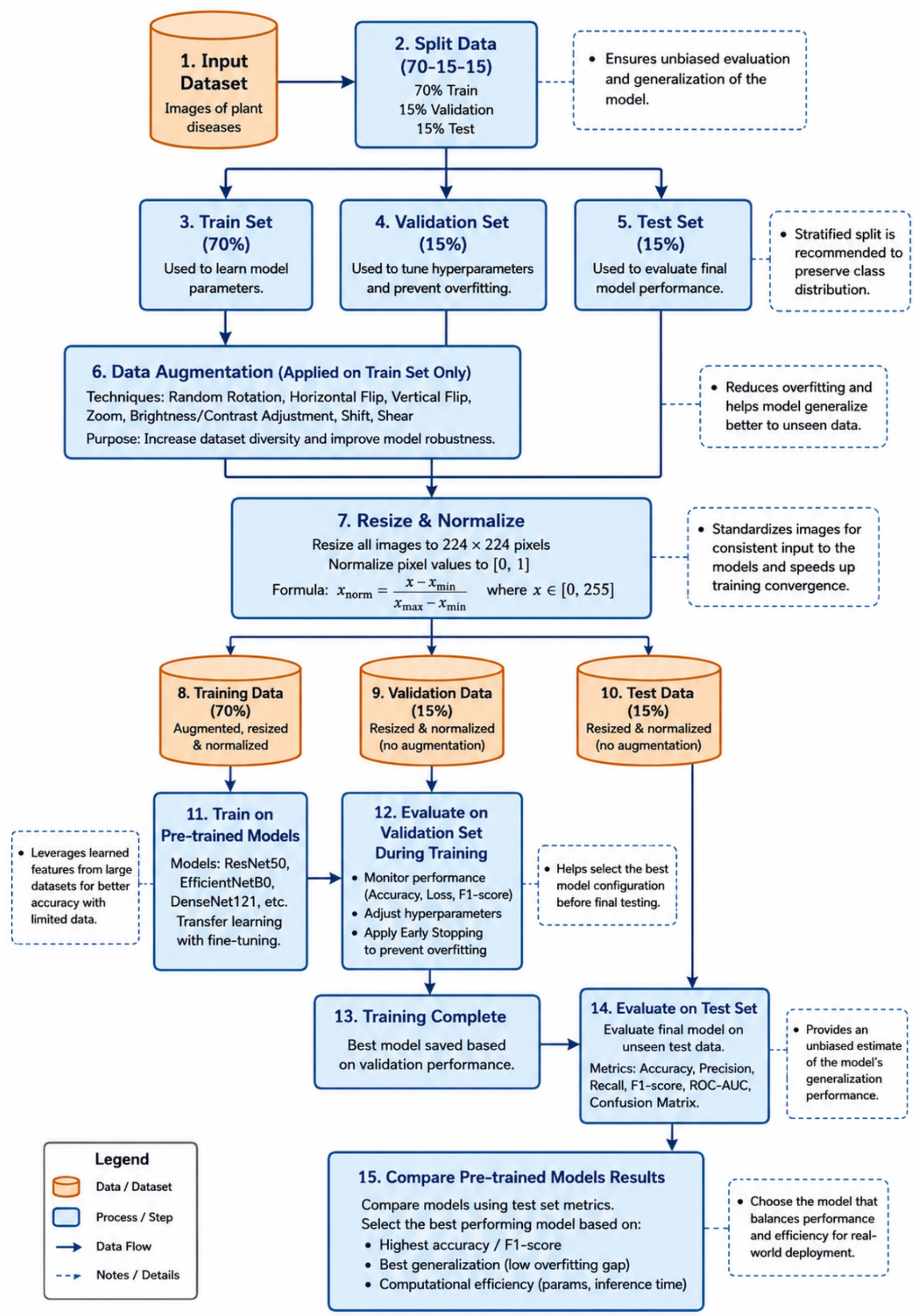


Figure 6. System Overview for Classification

epochs, batch size 32, and learning rate 0.0001. Figure 12 provides a visual comparison of all four models across Accuracy, Precision, Recall, and F1-score. Figures 13, 14, 15, and 16 show the confusion matrix of each model.

InceptionV3 outperformed the other proposed models with a Macro F1-score of 96.75%, compared to 95.75% for DenseNet121 in second place, 90.50% for ResNet50, and 82.25% for VGG19. Table I provides a full comparison of all four models across all reported metrics.

The champion model (InceptionV3) was also compared to the state-of-the-art attempts found in the literature that used the same dataset (LWDCD2020) to evaluate their approaches addressing the same problem. InceptionV3 achieved a Macro F1 score of 96.75% compared to 95% achieved by the highest performance in the literature. Table II shows an apple-to-apple comparison between the

TABLE I. Performance Comparison of Pretrained CNN Models proposed in this work

Model	Avg Precision	Avg Recall	Macro F1-Score
VGG19	82.25%	81.75%	82.25%
ResNet50	90.25%	90.50%	90.50%
DenseNet121	96.25%	96.00%	95.75%
InceptionV3	96.75%	96.50%	96.75%

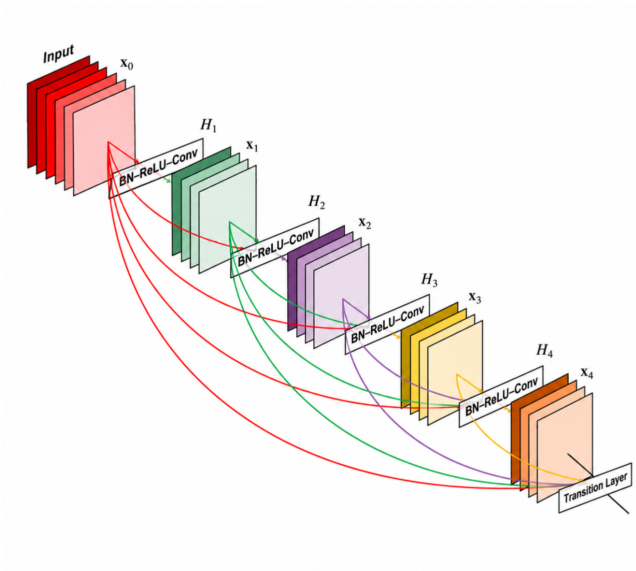


Figure 7. Architecture of DenseNet121 [37]

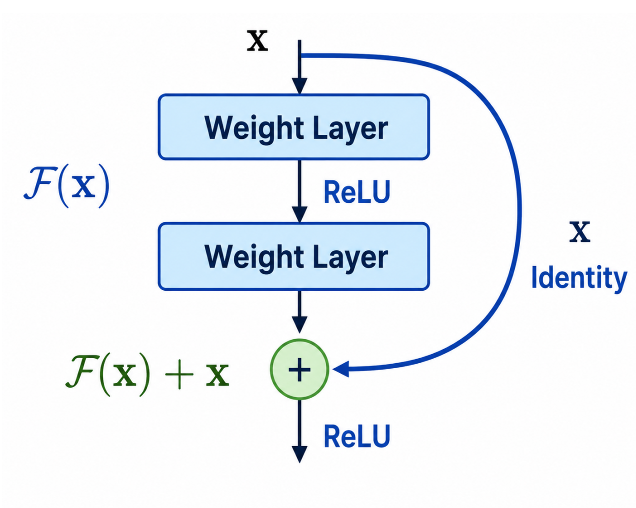


Figure 8. Residual Block in ResNet50 [38]

proposed work and other attempts that used the same data for the same purpose in the literature.

The superiority of InceptionV3 can be attributed to its architectural design, which efficiently manages the diversity–parameters ratio through parallel multi-scale filter pathways. As formalised by the Feature Diversity Bound (Equation 3), InceptionV3 maximises information diversity while minimising parameter count, enabling it to learn discriminative disease features with fewer gradient update cycles. In contrast, VGG19's uniform stack of 3×3 filters lacks this multi-scale capacity, leading to the weakest performance; ResNet50's residual connections aid gradient flow but do not provide explicit multi-scale feature fusion; and DenseNet121's dense connectivity promotes feature

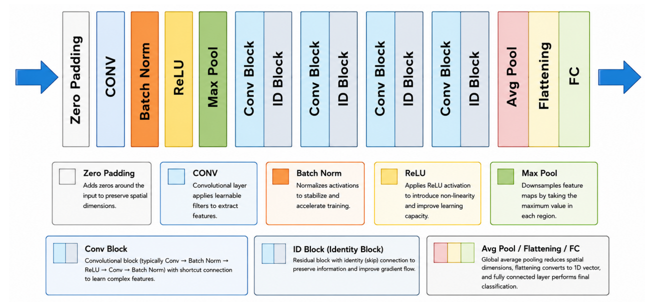


Figure 9. Architecture of ResNet50 [38]

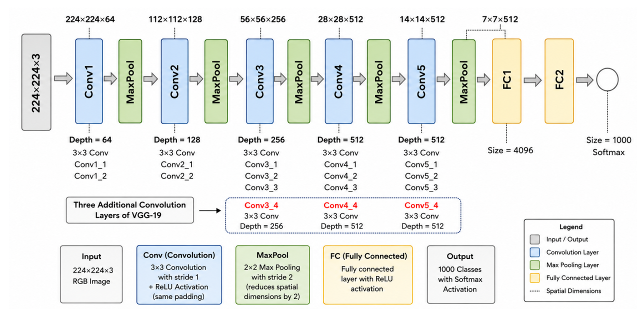


Figure 10. Architecture of VGG19 [39]

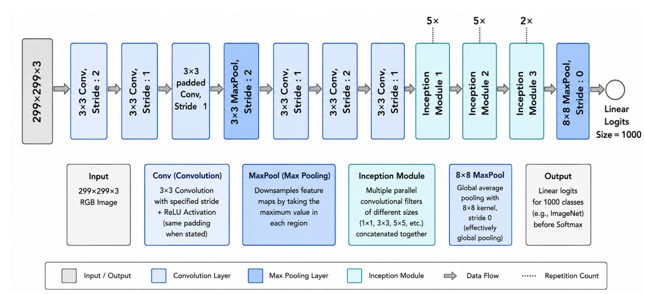


Figure 11. Architecture of InceptionV3 [40]

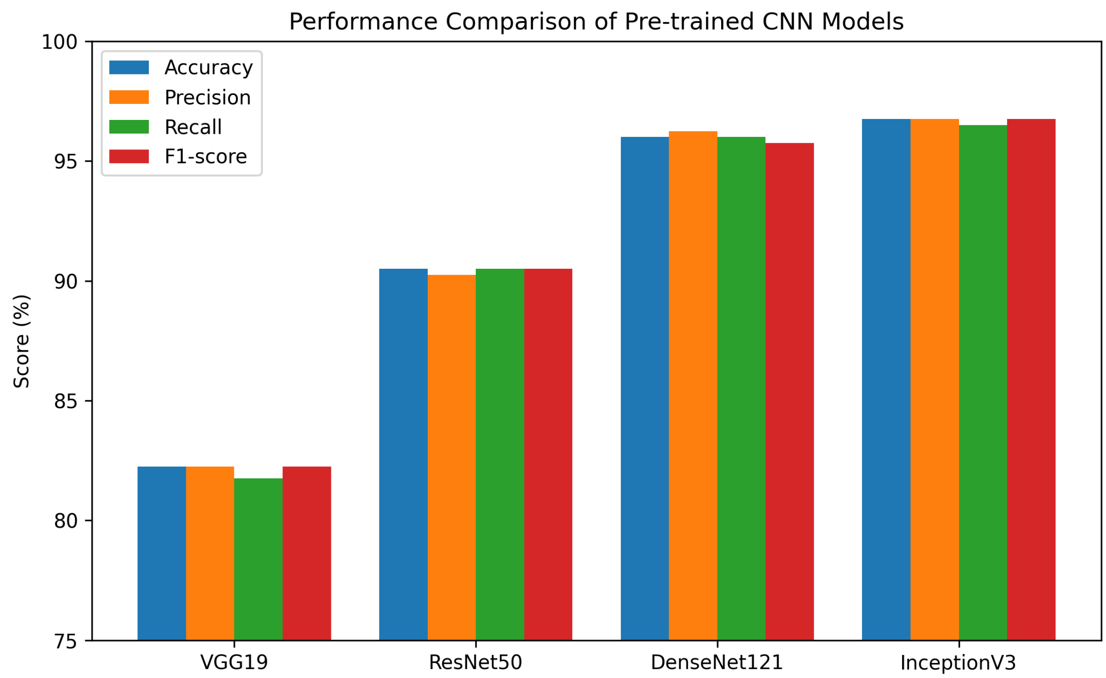


Figure 12. Performance Comparison of Pre-trained CNN Models across Accuracy, Precision, Recall, and F1-score

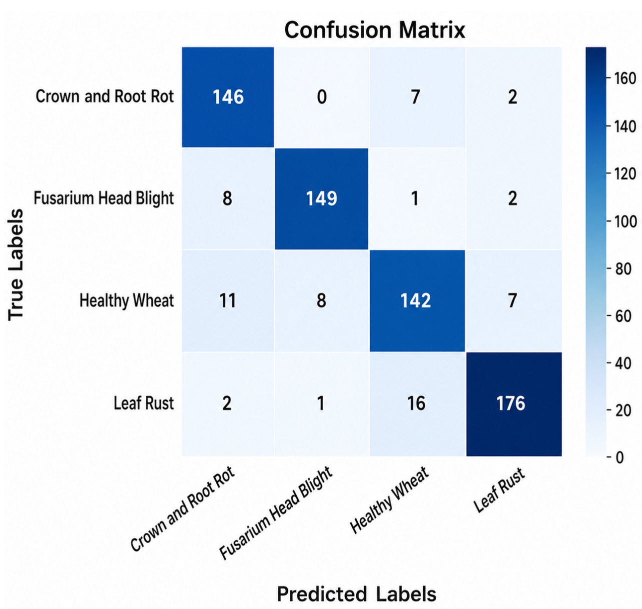


Figure 13. Confusion Matrix of ResNet50

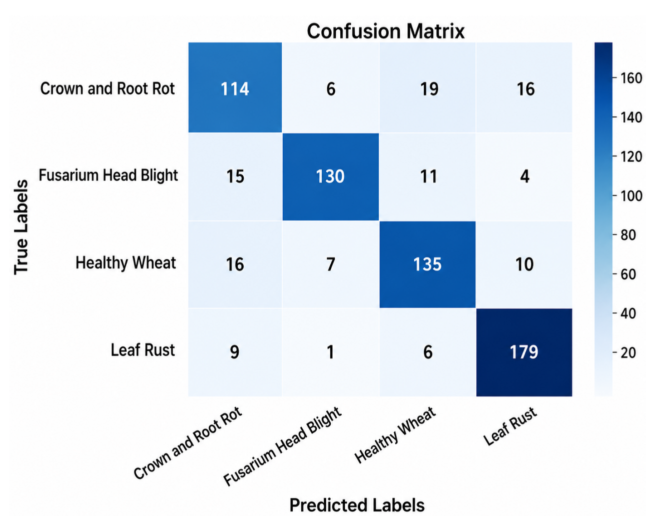


Figure 14. Confusion Matrix of VGG19

reuse but at the cost of fixed receptive fields. For wheat disease imagery—where discriminative patterns such as rust spores, blotch discolouration, and blight lesions appear at varying scales—InceptionV3's architecture provides a natural match to the visual characteristics of the classification task.

Another strength factor of our models is that we addressed the Efficiency–Accuracy Trade-off, which is very important to farmers who usually need faster, cheaper, and portable applications, as our light models can run on smartphones [41], [42], [43], [44]. The proposed models were trained for just 10 epochs compared to 50, 100, and 200 training epochs in state-of-the-art attempts in the literature. This suggests that the proposed approaches succeeded in achieving higher performance with a lower number of training epochs, which underscores their high

TABLE II. Performance Comparison of the proposed work and other state-of-the-art attempts

Paper	Model	Epochs	F1-Score	RPI
Samundiswary et al.[16]	ResNet 152	200	85.5%	0.428
Islam et al.[17]	Efficient Net+SA	100	92.8%	0.928
Etaati et al.[21]	VGG19	50	95%	1.900
This Work	**InceptionV3**	**10**	**96.75%**	**9.675**

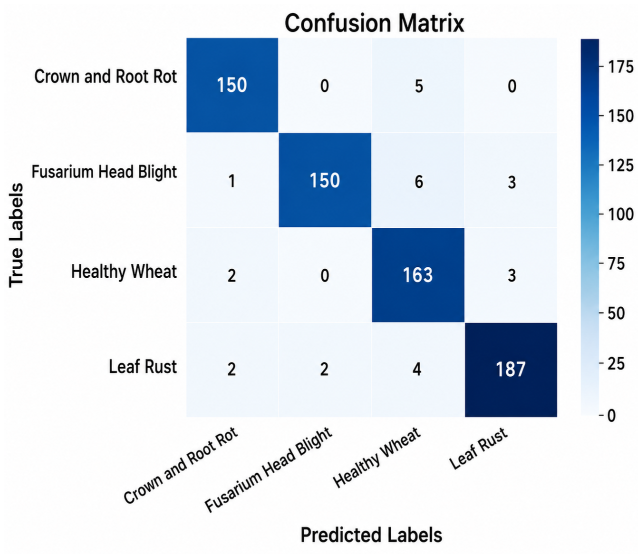


Figure 15. Confusion Matrix of DenseNet121

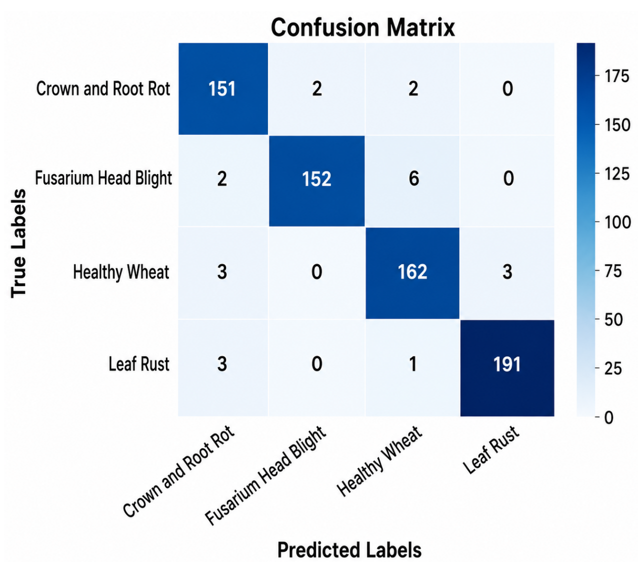


Figure 16. Confusion Matrix of InceptionV3

efficiency compared to other models proposed in the literature. Revolutions Per Inference (RPI) is introduced here as a relative training efficiency index to compare methods evaluated on the same dataset [45]. It quantifies how much classification performance is achieved per unit of training computation (epochs), making it a meaningful proxy for training cost under equivalent data conditions. It is important to note that RPI is not a hardware deployment metric; real-world deployment efficiency should additionally be measured by FLOPs, inference latency (ms), model size (MB), and frames per second (FPS), which we identify as an important direction for future work. The following equation measures the RPI for the best performing model in this paper (InceptionV3):

$$\text{RPI} = \frac{F1_{\text{proposed}}}{\text{Epochs}_{\text{proposed}}} \times 100 = 9.675 \tag{4}$$

Where RPI is the Relative Training Efficiency Index (F1-score per training epoch × 100).

A. Limitations

Despite the strong results achieved, several limitations should be acknowledged. First, the evaluation is confined to a single dataset (LWDCD2020), and generalizability to other wheat disease datasets or real-world field conditions has not been formally assessed. Second, only four disease categories are considered; performance on the full ten-class LWDCD2020 or other crop disease datasets remains unexplored. Third, while training for ten epochs prioritizes efficiency, performance could potentially be improved with additional training or more advanced learning rate scheduling. Fourth, the current framework addresses image classification only and does not include a disease localization component needed for a fully automated end-to-end pipeline. Fifth, hardware configuration details (GPU model, memory, per-model training time) were not formally logged in the original experimental runs and are therefore not reported. Finally, real-world deployment efficiency metrics—including FLOPs, inference latency (ms), model size (MB), and FPS—were not benchmarked on target hardware platforms, which would be required to fully validate the portability claims made in this study.

5. CONCLUSIONS AND FUTURE WORK

This study presents a comprehensive comparative evaluation of four CNN-based transfer learning models—InceptionV3, DenseNet121, ResNet50, and VGG19—for the classification of wheat diseases using the LWDCD2020 benchmark dataset. All models were trained under uniform and computationally efficient conditions (10 epochs, batch size 32, learning rate 0.0001), with targeted data augmentation applied to improve generalizability and prevent overfitting. InceptionV3 proved to be the champion model, achieving a Macro F1-score of 96.75%—surpassing all other proposed models and all previously reported results on LWDCD2020. ResNet50 came in second place, followed by DenseNet121, then VGG19 which produced the weakest performance.

From a technical standpoint, InceptionV3's superiority is explained by its Feature Diversity Bound, which enables it to maximise information diversity while minimising parameter count through parallel multi-scale convolutional pathways. This property allows it to converge in only 10 training epochs, compared to 50–200 epochs in competing approaches, yielding an RPI of 9.675—more than five times higher than the next best approach. From an agricultural standpoint, these results demonstrate the feasibility of lightweight, smartphone-deployable wheat disease detection systems that could significantly reduce crop losses in resource-constrained farming environments across the developing world.

Several directions for future work are identified. First, the framework should be extended with an automatic object detection module to create a fully automated end-to-end disease localisation and classification pipeline. Second, an ablation study evaluating the individual contribution of each augmentation strategy and fine-tuning decision would provide valuable insights into the design choices made in this work. Third, future studies should report hardware benchmarks (FLOPs, inference latency, model size, FPS) on edge devices to formally validate portability claims. Finally, cross-dataset validation on the full ten-class LWDCD2020 and additional wheat disease benchmarks is needed to assess generalizability.

References

[1] L. Goyal, C. M. Sharma, A. Singh, and P. K. Singh, "Leaf and spike wheat disease detection & classification using an improved deep convolutional architecture," *Informatics in Medicine Unlocked*, vol. 25, p. 100642, 2021.

[2] M. Figueroa, K. E. Hammond-Kosack, and P. S. Solomon, "A review of wheat diseases—a field perspective," *Molecular Plant Pathology*, vol. 19, no. 6, pp. 1523–1536, 2018.

[3] H. Anwar *et al.*, "The nwrd dataset: An open-source annotated segmentation dataset of diseased wheat crop," *Sensors*, vol. 23, no. 15, p. 6942, 2023.

[4] V. Balafas, E. Karantoumanis, M. Louta, and N. Ploskas, "Machine learning and deep learning for plant disease classification and detection," *IEEE Access*, pp. 1–1, 2023.

[5] M. Loey, A. Elsawy, and M. Afify, "Deep learning in plant diseases detection for agricultural crops: A survey," *International Journal of Service Science, Management, Engineering, and Technology*, vol. 11, p. 18, 2020.

[6] S. A. A. Qadri, N.-F. Huang, T. M. Wani, and S. A. Bhat, "Advances and challenges in computer vision for image-based plant disease detection: A comprehensive survey of machine and deep learning approaches," *IEEE Transactions on Automation Science and Engineering*, vol. 22, pp. 2639–2670, 2025.

[7] P. Broklyn, "AI in agriculture: Revolutionizing crop monitoring and disease management through precision technology," *Machine Learning Research*, 2024.

[8] N. K. Negied, "Expert system for wheat yields protection in Egypt (ESWYP)," *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, vol. 3, no. 11, pp. 2278–3075, 2014.

[9] A. Thengne, K. Indurkar, S. Nagdeote, and V. Nagrale, "Disease detection and expert system for farmers," *IJSR*, vol. 7, pp. 873–876, 2020.

[10] J. Barbedo, "Expert systems applied to plant disease diagnosis: Survey and critical view," *IEEE Latin America Transactions*, vol. 14, pp. 1910–1922, 2016.

[11] A. Chaudhary, M. Gupta, and U. Tiwari, "Crop disease detection using deep learning models," *International Journal of Innovative Science and Research Technology*, vol. 8, 2023.

[12] A. Trisal, K. Saini, and D. Mandloi, "Disease identification in crops using deep learning models," 2023.

[13] I. Bouacida, B. Farou, L. Djakhdjakha, H. Seridi, and M. Kurulay, "Innovative deep learning approach for cross-crop plant disease detection: A generalized method for identifying unhealthy leaves," *Information Processing in Agriculture*, vol. 12, no. 1, pp. 54–67, 2025.

[14] M. Long, M. Hartley, R. J. Morris, and J. K. M. Brown, "Deep learning for wheat disease classification by using deep learning networks with field and glasshouse images," *Plant Pathology*, 2022.

[15] M. Abdalla, O. Mohamed, and E. M. Azmi, "Adaptive learning model for detecting wheat diseases," *International Journal of Advanced Computer Science and Applications*, vol. 15, no. 5, 2024.

[16] P. Samundiswary, "Wheat disease detection and classification using ResNet152," *IJCRT*, vol. 11, pp. 2320–2882, 2023.

[17] M. Islam, M. Aloraini, S. Habib, M. D. Alanazi, I. Khan, and Khan, "Optimal features driven attention network with medium-scale benchmark for wheat diseases recognition," *IEEE Access*, vol. 12, pp. 150 739–150 753, 2024.

[18] S. N. Yousafzai *et al.*, "Multi-stage neural network-based ensemble learning approach for wheat leaf disease classification," *IEEE Access*, vol. 13, pp. 30 101–30 116, 2025.

[19] Y. Maqsood, S. M. Usman, M. Alhussein, K. Aurangzeb, S. Khalid, and M. Zubair, "Model agnostic meta-learning (MAML)-based ensemble model for accurate detection of wheat diseases using vision transformer and graph neural networks," *Computers, Materials & Continua*, pp. 1–10, 2024.

[20] D. Zhang *et al.*, "Using neural network to identify the severity of wheat fusarium head blight in the field environment," *Remote Sensing*, vol. 11, no. 20, p. 2375, 2019.

[21] S. Etaati, J. Khoramdel, and E. Najafi, "Automated wheat disease detection using deep learning: an object detection and classification approach," in *2022 10th RSI International Conference on Robotics and Mechatronics (ICRoM)*, 2023, pp. 116–121.

[22] Q. Hong *et al.*, "A lightweight model for wheat ear fusarium head

blight detection based on RGB images," *Remote Sensing*, vol. 14, no. 14, p. 3481, 2022.

[23] X. Yao, F. Yang, and J. Yao, "YOLO-Wheat: A wheat disease detection algorithm improved by YOLOv8s," *IEEE Access*, 2024.

[24] A. Hassan, R. Mumtaz, Z. Mahmood, M. Fayyaz, and M. K. Naeem, "Wheat leaf localization and segmentation for yellow rust disease detection in complex natural backgrounds," *Alexandria Engineering Journal*, vol. 107, pp. 786–798, 2024.

[25] C. Li *et al.*, "A comprehensive review of deep learning-based real-world image restoration," *IEEE Transactions on Artificial Intelligence*, vol. 4, no. 5, pp. 1109–1132, 2023.

[26] J. Chen, D. Zhang, Y. A. Nanehkaran, and D. Li, "Wheat disease recognition based on multi-scale feature fusion and attention mechanism," *Frontiers in Plant Science*, vol. 15, p. 1320177, 2024.

[27] X. Wang, J. Liu *et al.*, "An efficient convolutional neural network for crop disease and pest detection," *Agriculture*, vol. 14, no. 2, p. 312, 2024.

[28] A. Kumar, A. K. Singh, and P. Gaur, "Deep learning for plant disease detection and diagnosis: Current and future scenarios," *Electronics*, vol. 12, no. 4, p. 849, 2023.

[29] Y. Zhao, H. Li, M. Wang, and K. Xu, "Attention-guided multi-scale convolutional neural network for wheat disease classification," *Computers and Electronics in Agriculture*, vol. 218, p. 108691, 2024.

[30] J. Sun, Y. Yang, X. He, and X. Wu, "Multi-scale feature extraction and classification of wheat leaf diseases using deep learning," *Frontiers in Plant Science*, vol. 14, p. 1010540, 2023.

[31] B. Liu, Y. Zhang, J. Wang, and C. Zhao, "A lightweight deep learning model for real-time plant disease detection on edge devices," *Computers and Electronics in Agriculture*, vol. 220, p. 108914, 2024.

[32] Y. Peng, S. Zhao, and J. Liu, "Transfer learning for plant disease recognition: A survey," *Plant Phenomics*, vol. 2023, p. 9855486, 2023.

[33] S. Ali, T. Abuhmed, S. El-Sappagh *et al.*, "Deep learning techniques for crop disease detection: A systematic review and meta-analysis," *Artificial Intelligence Review*, vol. 57, p. 58, 2024.

[34] Z. Xu, X. Guo, A. Zhu, X. He, and X. Zhao, "A hybrid deep learning approach for simultaneous detection and classification of wheat diseases," *Sensors*, vol. 23, no. 3, p. 1242, 2023.

[35] K. Maharana, S. Mondal, and B. Nemade, "A review: Data preprocessing and data augmentation techniques," *Global Transitions Proceedings*, vol. 3, no. 1, pp. 91–99, 2022.

[36] A. Mumuni and F. Mumuni, "Data augmentation: A comprehensive survey of modern approaches," *Array*, vol. 16, p. 100258, 2022.

[37] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 4700–4708.

[38] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.

[39] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[40] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 2818–2826.

[41] K. Shirini, M. Balaneshin Kordan, and S. Samadi Gharehveran, "Impact of learning rate and epochs on LSTM model performance: A study of chlorophyll-a concentrations in the Marmara Sea," *The Journal of Supercomputing*, vol. 81, 2024.

[42] D. Wilson and T. Martinez, "The need for small learning rates on large problems," in *Proceedings of the International Joint Conference on Neural Networks*, vol. 1, 2001, pp. 115–119.

[43] S. AlMuhaideb, L. AlAbdulkarim, D. AlShahrani, H. AlDhubaib, and D. AlSadoun, "Achieving more with less: A lightweight deep learning solution for advanced human activity recognition (HAR)," *Sensors*, vol. 24, no. 16, p. 5436, 2024.

[44] Y. S. Taspinar, "Light weight convolutional neural network and low-dimensional images transformation approach for classification of thermal images," *Case Studies in Thermal Engineering*, vol. 41, p. 102670, 2023.

[45] S. Mathe, H. Kondaveeti, S. Vappangi, S. Vanambathina, and N. Kumaravelu, "A comprehensive review on applications of Raspberry Pi," *Computer Science Review*, vol. 52, p. 100636, 2024.