scientific reports

Check for updates

OPEN

Novel dual-input stream-based hybrid approach for wheat leaf disease classification using edge-aware features

Ayesha Razaq[1,5], Shabana Ramzan[1], Basharat Ali[2,5], Ali Raza[3], Yeonghyeon Gu[4]✉, Muhammad Syafrudin[4]✉ & Norma Latif Fitriyani[4]

The prevalence of diseases in wheat crops poses a significant threat to global food security, as it reduces yield and quality. Addressing these challenges is critical for sustainable agriculture. This study proposes and evaluates a hybrid deep learning (DL) model, EffiXB3, which combines Xception and EfficientNetB3 architectures, enhanced with edge-aware features, to improve disease classification in wheat crops. EffiXB3 employs a dual-input stream architecture, where one stream processes structural features, while the other incorporates textural features through Canny edge detection. The performance of individual models, Xception and EfficientNetB3, was assessed alongside the hybrid EffiXB3 model in a multi-class classification task involving five wheat leaf categories: Blast, Brown Rust, Healthy, Leaf Blight, and Septoria. Xception and EfficientNetB3 achieved classification accuracies of 95% and 93%, respectively. The proposed EffiXB3 model outperformed both, achieving an accuracy of 98.5%. The integration of edge-aware features substantially improved robustness and classification performance, particularly in differentiating visually similar disease patterns. The findings demonstrate the effectiveness of hybrid DL models with edge feature integration in diagnosing agricultural diseases. EffiXB3 offers a promising approach for enhancing disease detection in wheat, contributing to improved crop management and food security.

Wheat (Triticum aestivum L.) is one of the most essential crops grown and a basic need in the world's food supply[1]. Ensuring its healthy production is crucial for food safety and economic stability. However, its yield and quality are continually at risk due to numerous foliar diseases. Such diseases not only lower production but also lead to significant economic losses and pose a threat to food security, especially in regions that heavily depend on wheat cultivation[2]. Traditionally, crop disease identification has relied on manual field inspections conducted by experts. Although effective in some scenarios, these methods are labor-intensive, slow, and prone to human error and subjective judgment, particularly when distinguishing between diseases with similar visible symptoms. As a result, there has been growing interest in automated, reliable approaches for plant disease diagnosis. Advances in computer vision and deep learning (DL) have enabled models to automatically learn informative features using leaf images[3]. These techniques have become a cornerstone in modern agricultural informatics, offering scalable solutions for crop health monitoring[4]. Convolutional Neural Networks (CNNs), particularly pre-trained models, are utilized to diagnose crop-specific diseases with greater accuracy. Evaluating these models on real-world field images helps farmers maintain healthier, more productive crops. Crop productivity depends on many factors, including environmental conditions, soil quality, and proper nutrition. Making informed decisions about field selection and disease management plays a key role in sustaining yields. Early detection of diseases is critical because it can reduce the rate of infection and limit damage. Integrating accurate predictive models into IoT-enabled sensing devices can further improve automated crop monitoring. These systems can help identify suitable soils for planting and detect foliar diseases in real-time, supporting more effective management[5–7].

CNNs and DL models have achieved high accuracy in recognizing visual symptoms across many crops[8]. Combining advanced DL architectures with hybrid optimization techniques can substantially enhance disease classification accuracy in crops[9]. Wheat crops are affected by a range of conditions that give rise to different

[1]Department of Computer Science and IT, Government Sadiq College Women University, Bahawalpur, Pakistan. [2]Agronomic Research Station, Bahawalpur, Pakistan. [3]Department of Precision Medicine, Sungkyunkwan University School of Medicine, Suwon 16419, Republic of Korea. [4]Department of Artificial Intelligence and Data Science, Sejong University, Seoul 05006, Republic of Korea. [5]These contributed equally to this work: Ayesha Razaq and Basharat Ali. ✉email: yhgu@sejong.ac.kr; udin@sejong.ac.kr

categories of diseases. These diseases can often be identified by examining the health and appearance of wheat leaf samples. Several common foliar diseases that significantly impact wheat production include leaf rusting[10], wheat blast[11], and Septoria[12]. Features of wheat leaves extracted using DL models trained on datasets containing both diseased and healthy leaf samples.

EfficientNetB3 is one of the CNNs that scales model dimensions in a principled way, balancing accuracy on evaluation by extracting features from sample datasets. This model is well-suited for transfer learning in image classification tasks[13]. Xception is a deep convolutional neural network architecture that enhances the original Inception design. It integrates depth-wise separable convolutions, which separate spatial relationships from cross-channel correlations and a linear sequence of separable convolution layers combined with residual connections, making the network more efficient and effective. Xception achieves higher accuracy on the ImageNet benchmark while staying computationally efficient, making it a strong backbone for both image classification and feature extraction tasks[14].

Despite these advances, most existing systems focus solely on RGB image features, often overlooking additional structural or texture information, such as leaf edge patterns, that can help distinguish diseases with similar appearances. Moreover, while robust CNN architectures like Xception and EfficientNet have set new performance benchmarks, effectively combining their strengths typically requires novel design strategies. Hybrid models that merge multiple feature extractors and incorporate preprocessing steps, such as edge detection, offer a promising improvement in accuracy. To overcome these challenges, this study introduces EffiXB3, a hybrid DL model combining Xception and EfficientNetB3 with a dual-input architecture. Our model processes regular RGB images in one branch, while the other analyzes Canny edge-enhanced images to capture fine structural details of wheat leaf samples. By integrating these complementary data streams, the model aims to improve classification accuracy and resilience under varying visual conditions. This research evaluates EffiXB3 on a balanced dataset covering five wheat disease categories, including healthy samples. The proposed model demonstrates clear improvements over single-stream models, compares its performance with leading existing methods, and shows that incorporating edge-aware features substantially boosts recognition accuracy. The results highlight a scalable and dependable approach to early disease detection, supporting better crop management and informed decision-making in agriculture.

Study contribution

These are the main contributions of this study:

- We propose EffiXB3, a hybrid model combining Xception for Canny edge-enhanced and EfficientNetB3 for RGB image streams. This dual-path setup allows simultaneous learning of texture and structural features.
- By incorporating Canny edge detection as a separate input channel, enabling the model to capture fine-grained structural cues.
- The model is assessed on a balanced, multi-class dataset of five crucial wheat class categories–Blast, Brown Rust, Healthy, Leaf Blight, and Septoria–with thorough evaluation using accuracy, precision, recall, and F1-score metrics.

This study introduces EffiXB3, a novel hybrid DL model. The novelty of our research lies in its architecture and approach. We propose a novel dual-input stream architecture that processes two distinct input modalities in parallel: one stream analyzes RGB textures (using EfficientNetB3). In contrast, a separate, dedicated stream processes Canny edge-enhanced images (using Xception) to explicitly capture structural contours and lesion boundaries. The model's innovation is not merely the concatenation of features, but the strategic fusion of complementary feature types, color or texture from RGB, and shape or structure from edges. This approach provides a richer and more discriminative feature set than either modality alone. We rigorously validate this approach on a comprehensive fifteen-class wheat disease dataset to demonstrate significant performance improvement and enhanced generalizability over state-of-the-art methods.

Research hypotheses

To thoroughly assess the performance of the proposed EffiXB3 model, we defined the following testable hypotheses to shape our experimental design.

- **H1:** Integrating Canny edge-enhanced inputs with RGB image streams in a hybrid dual-branch architecture improves classification accuracy over single-stream models.
- **H2:** The combination of Xception and EfficientNetB3 backbones achieves better feature representation and disease discrimination.
- **H3:** The inclusion of structural edge information reduces false positives, particularly in disease classes with visually similar symptoms, such as Septoria and Brown Rust.

The remainder of this paper is organized as follows. Section 2 reviews related work and summarizes recent advances in wheat disease detection and hybrid DL approaches. Section 3 provides a detailed description of the proposed methodology, including dataset preparation, preprocessing, model architecture, training strategies, and evaluation metrics. Section 4 presents the experimental results, state-of-the-art comparison, and discusses the findings, practical implications, and limitations of this study. Section 5 concludes the paper and outlines directions for future research.

Literature review

Advancements in artificial intelligence have significantly improved plant disease detection systems by leveraging DL models for classifying healthy and diseased crop leaves[15]. DL methods have become the dominant approach for automatic plant disease classification due to their strong feature representation capabilities and scalability. Multiple studies have demonstrated that convolutional neural networks (CNNs) outperform traditional image processing techniques in identifying crop diseases under varied conditions. Recent research has focused on enhancing classification accuracy through the use of advanced architectures, data augmentation, and hybrid feature extraction strategies.

The research in[16] trained an EfficientNet-B0 model covering five fungal diseases: leaf rust, stem rust, yellow rust, powdery mildew, and septoria, as well as seedling and healthy states. To reduce data redundancy, they applied image-hashing during dataset curation. The model achieved 94.2% accuracy, with high performance in single-disease classification; however, multi-disease images proved more challenging, yet were still partially recognizable. The study[17] proposed a DAE-Mask model that integrated the DenseNet architecture with an edge-aware segmentation mechanism, demonstrating that the model accurately detected diseases in wheat fields with a rate of 96.02%. The researchers in[18] developed and evaluated two CNN models designed to automatically detect and classify diseases in wheat plants. They trained the models using a diverse dataset containing images of healthy and infected wheat leaves. In testing, one model achieved a validation accuracy of 77.23%, while the other, employing a deeper architecture, achieved a higher accuracy of 86.96%.

The researchers in[19]classified two models, Xception and ResNet-50, for detecting rust disease in wheat leaves. Their findings showed that the CNN ResNet-50 architecture performed more accurately, yielding an accuracy of 93%. For the detection of wheat leaf diseases[20], used a pre-trained EfficientNetB3 model. The model was trained on three categories of wheat leaf diseases and achieved an accuracy of 95%. In[21], they investigated multiple pre-trained models for identifying diseases in wheat leaves. Their study demonstrated that the MobileNetV2 model achieved a remarkable 96 percent accuracy. In[22], they introduced a hybrid DL model that integrates MobileNet and XGBoost to enhance the predictive capability of wheat diseases. The approach incorporated techniques such as Local Binary Patterns (LBP), K-means clustering, and Gaussian filtering to minimize noise. Their experimental results showed that the hybrid approach reached an accuracy of 94.67%.

The hybrid approach of integrating an EfficientNet-based CNN architecture to detect illness in wheat leaves demonstrates the model's performance, achieving 95.12% accuracy[23]. The study in[24]proposed a hybrid approach for detecting wheat diseases. Their SC-ConvNeXt model is based on a pre-trained architecture integrated with a CBAM attention mechanism. By evaluating the model, their results demonstrated an accuracy of 88.05%. By integrating EfficientNet-B3 with a feature-driven attention mechanism[25], introduced an approach to recognize diseases in wheat. Their proposed model was used to identify the difference between the structure of healthy and unhealthy leaves. They evaluated the model using two different datasets and presented the results for both datasets. The model achieved accuracies of 96.71% and 97.3% on the WD5CC and LWDCD2020 datasets, respectively. Researchers in[26] utilized the EfficientNetV2 DL model for detecting rust diseases in wheat leaves. They demonstrated that the model achieved a high accuracy of 95.2% as compared to the CNN and Vision Transformer (ViT) architectures. Another study in[27] introduced an approach to detect wheat rust and measure infection severity in a single-stage pipeline based on the Xception DL model. They used the GrabCut algorithm to segment the diseased area from the rest of the leaf and then analyzed the segmented region in the CIELAB to separate rust stripes from healthy tissue. Overall, the approach combines visual classification and precise quantification within a single, integrated framework. The Xception model proposed by[28] for the detection of diseases in wheat leaves. Their study showed that the proposed approach on the CNN model achieved 92.5% accuracy.

Research gap

While feature fusion techniques are well-established in broader computer vision literature[29]. Most existing approaches rely exclusively on RGB image information without explicitly incorporating structural or edge-based features. Furthermore, while attention mechanisms and multi-scale feature extraction have been widely adopted in other domains, their combination with edge enhancement techniques for wheat disease classification remains largely unexplored[30]. The limited integration of hybrid input streams, which combine both texture and edge information, represents a significant opportunity for improving model robustness and accuracy. Additionally, prior studies often focus on single backbone architectures, lacking comparative analysis of dual-branch frameworks that can simultaneously process complementary modalities. This research aims to address these gaps by introducing EffiXB3. This hybrid dual-stream network fuses Xception and EfficientNetB3 architectures with Canny edge-enhanced representations, thereby enriching the feature space for more precise disease identification.

Table 1 shows the summary of the methodology used and the comparative results of the related literature review. Despite these advancements, limited research has addressed the integration of dual-input architectures that combine RGB and edge-enhanced representations within a single framework. This gap motivated the development of the present study, which proposes EffiXB3, a hybrid EfficientNetB3–Xception model leveraging Canny edge information alongside conventional RGB features for more reliable wheat disease identification.

Proposed methodology

This study is based on EffiXB3, a hybrid model trained on a publicly available dataset. The required pre-processing techniques were applied to the dataset. The trained hybrid model is evaluated on the given test set based on various metrics, including accuracy, precision, recall, F1-score, and a confusion matrix. The methodology of the proposed model is outlined in Fig. 1.

Ref.	Year	Methodology used	Dataset	Results
16	2021	EfficientNet-B0 applied to wheat leaf diseases	WFD2020	Achieved 94.2% accuracy, demonstrating EfficientNet-B0 suitability
17	2023	DAE-Mask Edge-aware segmentation with DenseNet	MSWDD2022 and PlantDoc	Edge-aware segmentation with DenseNet model achieved 96.02% accuracy
18	2023	Proposed two models, based on the CNN and EfficientNetB7	PDDP	Model 1 reached an accuracy of 77.23%, while Model 2 employed a deeper architecture, achieved a higher accuracy of 86.96%.
19	2023	Xception and ResNet-50 model	Utilized region-specific dataset	Achieved 96% accuracy by ResNet-50 model
20	2023	EfficientNetB3 model	-	Potential algorithms achieved 95% accuracy
21	2024	Analyzed different pre-trained model	-	An impressive 96% accuracy achieved on MobileNetV2 model
22	2024	A hybrid MobileNet and XGBoost based model	-	Achieved 94.67% accuracy by merging lightweight neural network with gradient boosting classifier
23	2024	Efficient Net-based CNN architecture	Utilized publicly available datasets	Achieved an accuracy of 95.12%
24	2024	SC-ConvNeXt + CBAM attention mechanism	'Smart Agriculture' platform	Model identified the illness of wheat leaf with 88.05% accuracy
25	2024	EfficientNet-B3 with a spatial attention (SA) mechanism	WD5CC and LWDCD2020	Trained model achieved 96.71% accuracy on LWDCD2020 dataset
26	2025	EfficientNetV2 model	Utilized publicly available datasets	Model identified rusty wheat leaf with 95.2% accuracy
27	2025	Xception model	Yellow-Rust-19	Trained model tested on three test sets and achieved 63.4% to 90% accuracy
28	2025	Xception model	Utilized publicly available datasets	Trained model achieved 92.5% accuracy

Table 1. Summary of literature review of wheat leaf diseases detection.

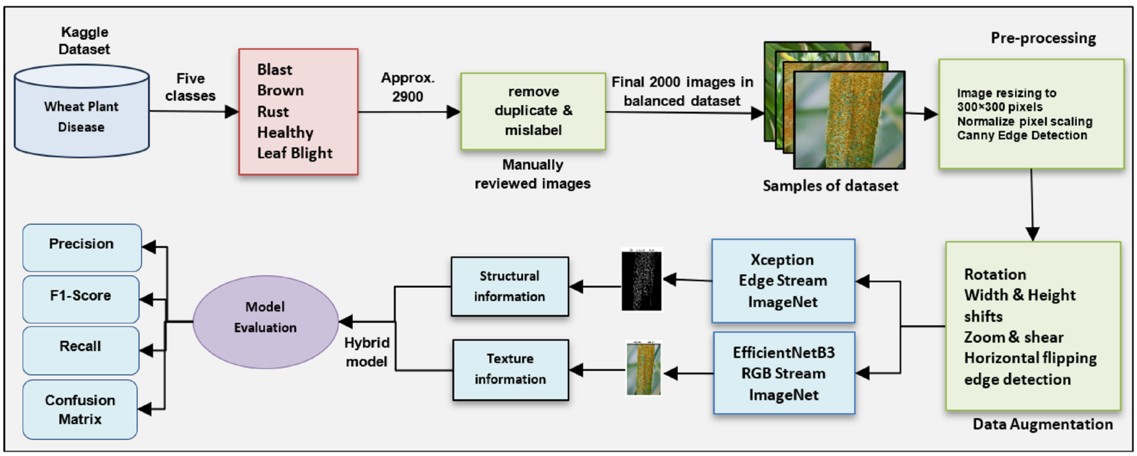

Fig. 1. Methodology diagram of proposed EffiXB3 model.

Dataset collection

A publicly available dataset on Wheat Plant Diseases[31] from Kaggle is utilized in this study. The dataset comprises wheat leaf images affected by common diseases. To prepare the data for modeling, all images were manually reviewed to remove duplicates and mislabels. In this study, a total of 2,000 (400 in each class) images of five distinct categories: Blast, Brown Rust, Healthy, Leaf Blight, and Septoria are used. The final dataset was partitioned into training (80%), validation (10%), and testing (10%) sets to ensure an unbiased evaluation of the proposed model. The random images from the balanced train set are shown in Fig. 2.

Reason for choosing five categories: The original dataset contains a comprehensive set of fifteen disease categories. These classes are chosen based on two primary criteria:

- Economic Significance: These diseases are among the most impactful on global wheat yields.
- High Visual Similarity: These diseases, particularly Brown Rust and Septoria, often present with similar visual symptoms (e.g., leaf spots, discolouration). This creates a rigorous testbed for evaluating our core hypothesis that edge-aware features can disambiguate subtle inter-class differences that are challenging for models relying solely on RGB data.

Data preprocessing

Before training the model, all images undergo a series of preprocessing steps to standardize them, thereby improving consistency and enhancing feature extraction. Each image is resized to 300 × 300 pixels to match

Fig. 2. Sample images of the wheat dataset.

the input requirements of the Xception and EfficientNetB3 architectures. Pixel values are normalized to a [0, 1] range by dividing by 255.0, as shown in Equation (1), ensuring stable convergence during training.

$$x_{\text{normalized}} = \frac{x}{255} \tag{1}$$

Where x is the original pixel value (ranging from 0 to 255) and $x_{\text{normalized}} \in [0, 1]$. In parallel, canny edge detection is applied to all input images. Each RGB image is converted to grayscale, and Canny edge detection is performed with threshold values of 100 and 200. The resulting edge maps are then converted back to RGB format to maintain compatibility with pre-trained CNN input specifications. Normalization is applied to both the RGB images and the Canny edge-enhanced images to maintain consistency across inputs. This process emphasized the structural boundaries of wheat leaf lesions, providing complementary feature representations that improve classification accuracy.

Data augmentation

To enhance the generalization of the proposed model and reduce the risk of overfitting caused by a limited dataset, various data augmentation techniques are applied during training. Data augmentation enhances dataset diversity by introducing random modifications to input images, enabling the model to learn features that remain consistent across different conditions. Augmentation is applied to the training dataset to provide a reliable measure of real-world performance. For different leaf orientations, a random rotation of 15 degrees is applied to the images, shifting them horizontally and vertically by up to 10% of their size to simulate positional changes. We also use random zoom adjustments to represent scale variations. To create perspective distortions, a shear transformation is applied.

Additionally, horizontal flipping is applied by randomly flipping images to capture mirror-image appearances of wheat leaves. These augmentations are applied dynamically during training using the ImageDataGenerator module, ensuring that each epoch receives fresh and varied data. For the Canny edge-enhanced inputs, the same augmentation settings are applied after edge detection to keep the RGB and edge streams properly aligned.

Proposed model architecture

The proposed model employs a hybrid dual-CNN architecture that integrates both Xception and EfficientNetB3 networks as parallel feature extractors. An input tensor of shape $300 \times 300 \times 3$ is first processed independently through each pretrained backbone, which is initialized with ImageNet weights to leverage prior knowledge of visual features. The convolutional layers of both base models are frozen during training, ensuring that only the classification head learns task-specific patterns. After passing through the convolutional stages, each branch applies Global Average Pooling, reducing the spatial dimensions to compact feature vectors. These two vectors, one from EfficientNetB3 and one from Xception, are concatenated to form a unified representation that captures both texture and structural information. This merged feature vector is normalized via Batch Normalization and regularized using Dropout layers (with dropout rates of 0.5 and 0.2). The categorical cross-entropy loss function equation (2) is used to measure the dissimilarity between the predicted probability distribution and the actual distribution.

$$L = -\sum_{i=1}^{C} y_i \log(\hat{y}_i) \tag{2}$$

where C is number of classes, y_i is ground-truth label and $\hat{y}_i$ is predicted probability for class *i*. A fully connected Dense layer further refines the representation, followed by a final Softmax output layer to generate class probabilities across the five target disease categories. The softmax activation function, as in equation (3), normalizes outputs to a probability distribution across all categories.

$$\hat{y}_i = \frac{\exp(z_i)}{\sum_{j=1}^{C} \exp(z_j)} \tag{3}$$

Where $\hat{y}_i$ is the predicted probability for class *i*, z_i is the logit output for class *i*, z_j is the logit (the raw output) from the last dense layer of the model for class *j*, and *C* is the number of classes. Internally, the dense layer of softmax computes z_j, the function as in equation (4).

$$z_j = W_j^T x + b_j \tag{4}$$

Where *W* is the weights vector for class *j*, *x* is input feature vector (after Dropout and BatchNorm), b_j bias term for class *j*. The architecture strikes a balance between rich feature extraction, computational efficiency, and robust convergence, making it well-suited for accurate wheat disease classification. Table 2 shows the hyperparameters applied in the proposed EffiXB3 model.

Proposed model training details

The model is trained using the Adam optimizer with an initial learning rate of 0.0001 and the categorical cross-entropy loss function. Early stopping is employed to prevent overfitting by monitoring validation loss with patience. Training continued for a maximum of 25 epochs with a batch size of 32 and 50 steps per epoch. The model training is performed in two phases to optimize both generalization and fine-tuned performance.

Phase 1: feature extraction

In the first stage of the model pipeline, features are independently extracted from two complementary input streams to capture both textural and structural details of wheat leaves. Original RGB images are input into a pretrained EfficientNetB3 network. As the network processes the images through its convolutional layers, it generates rich representations that encode color, texture, and contextual features important for disease identification. The resulting convolutional feature maps are then passed through a Global Average Pooling (GAP) layer to produce a compact feature vector summarizing the spatial information. At the same time, grayscale Canny edge-enhanced images are fed into a pretrained Xception model. This stream focuses on capturing the leaf's edge contours and structural boundaries, which are particularly helpful for distinguishing diseases that share similar color characteristics. The output feature maps from this branch are also condensed using Global Average Pooling, creating a concise descriptor of edge-related features. By extracting these complementary features in parallel, the model generates a richer and more discriminative feature space, ultimately enhancing classification accuracy and robustness. Initially, all layers of the Xception and EfficientNetB3 base models freeze, allowing only the classification head to learn task-specific representations. The model is trained with the Adam optimizer and categorical cross-entropy loss for 10 epochs. To prevent overfitting and improve generalization,

Component	Details
Pretrained Backbones	EfficientNetB3 and Xception
Pooling	Global Average Pooling on each backbone output
Concatenation	Fused features before dense layers
Dense Layer 1	256 units, ReLU activation
Dense Layer 2	5 units, Softmax activation
Batch Normalization	Yes
Dropout Rates	0.5 (after concatenation), 0.2 (after Dense)
Optimizer	Adam
Loss Function	Categorical Crossentropy
Batch Size	32
Input Shape	(300, 300, 3)
Learning rate	1e-4 - 1e-5
Callbacks	ModelCheckpoint, EarlyStopping, ReduceLROnPlateau
Epoch	25
Augmentation	rotation_range=15, width_shift_range=0.1, height_shift_range=0.1, zoom_range=0.1, shear_range=0.15, horizontal_flip=True, fill_mode='nearest'

Table 2. Hyperparameters configuration of EffiXB3 model.

the training process incorporates early stopping, model checkpointing, and learning rate reduction on a plateau, which dynamically lowers the learning rate when the validation loss plateaus.

Phase 2: fine-tuning

In the second phase, the model moves beyond basic feature extraction to fine-tune and learns task-specific representations tailored to wheat disease classification. After initializing Xception and EfficientNetB3 with pretrained ImageNet weights, the model undergoes selective fine-tuning to adapt these learned features to the unique visual characteristics of the wheat leaf dataset. Initially, all layers of both backbone networks are frozen to preserve the general visual features acquired during pre-training. As training progresses, selected upper layers are incrementally unfrozen in later epochs. This approach enables the model to learn dataset-specific details while minimizing the risk of catastrophic forgetting. A low learning rate (1e-4) is used during fine-tuning to ensure stable convergence and avoid abrupt updates that could disrupt the pretrained weights. Fine-tuning enables the model to distinguish more accurately among disease classes that are not adequately represented in broad datasets, such as ImageNet. Overall, this phase helps both the RGB and edge-based streams effectively learn relevant disease features while maintaining model stability and reducing the risk of overfitting. The second training cycle is performed for an additional 15 epochs using the same early stopping, model checkpointing, and learning rate reduction callbacks to monitor validation performance and prevent overfitting. All training and validation batches are generated using the custom CannyEdgeDataGenerator, with augmentation applied only to the training data.

Evaluation metrics

To comprehensively assess the performance of the EffiXB3 model, several quantitative metrics are employed. In this section, we discussed each metric in detail and visually represented the result values. These metrics are computed on the test set after the model has been trained. We analyze them collectively to evaluate the model's strengths and weaknesses in recognizing each disease category.

Accuracy: as in equation (5), is the proportion of correctly classified images among all predictions, providing an overall measure of correctness.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

Where, in terms of multi-class categories of diseases in wheat leaf, means:

- **TP (True Positive)**: Number of wheat leaf correctly predicted as the target class
- **TN (True Negative)**: Number of wheat leaves correctly predicted as NOT the target class
- **FP (False Positive)**: Wheat leaf incorrectly predicted as the target class (leaf belongs to another class).
- **FN (False Negative)**: Wheat leaf incorrectly predicted as NOT the target class (leaf actually belongs to the target class).**Precision:** is the ratio of true positive predictions (accurate detection of wheat disease class) to the total number of positive predictions, indicating how effectively the model avoids false positives for each class, as shown in equation (6).

$$\text{Precision} = \frac{TP}{TP + FP} \tag{6}$$

As in equation (7), **Recall** is the ratio of true positive predictions to all actual positive classes, reflecting the model's ability to detect disease instances without omission.

$$\text{Recall} = \frac{TP}{TP + FN} \tag{7}$$

F1-score: as in equation (8), is the harmonic mean of precision and recall, offering a balanced indicator especially useful in cases of class imbalance.

$$\text{F1-Score} = 2 \times \left(\frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \right) \tag{8}$$

Confusion Matrix: is a tabular representation showing the distribution of correct and incorrect predictions across all classes, used to identify specific misclassification patterns.

Receiver Operating Characteristic (ROC) Curve: The area under the curve (AUC)-ROC is another metric used to evaluate the performance of a classification model. Receiver Operating Characteristic (ROC) Curve plots the True Positive Rate (Sensitivity) against the False Positive Rate at different classification thresholds. The AUC summarizes the curve into a single value ranging from 0 to 1.

Results and discussion

The proposed EffiXB3 model and its performance in comparison to existing approaches are highlighted in this section. Further, we analyze the impact of integrating edge-based features and discuss the findings in the context of wheat disease detection.

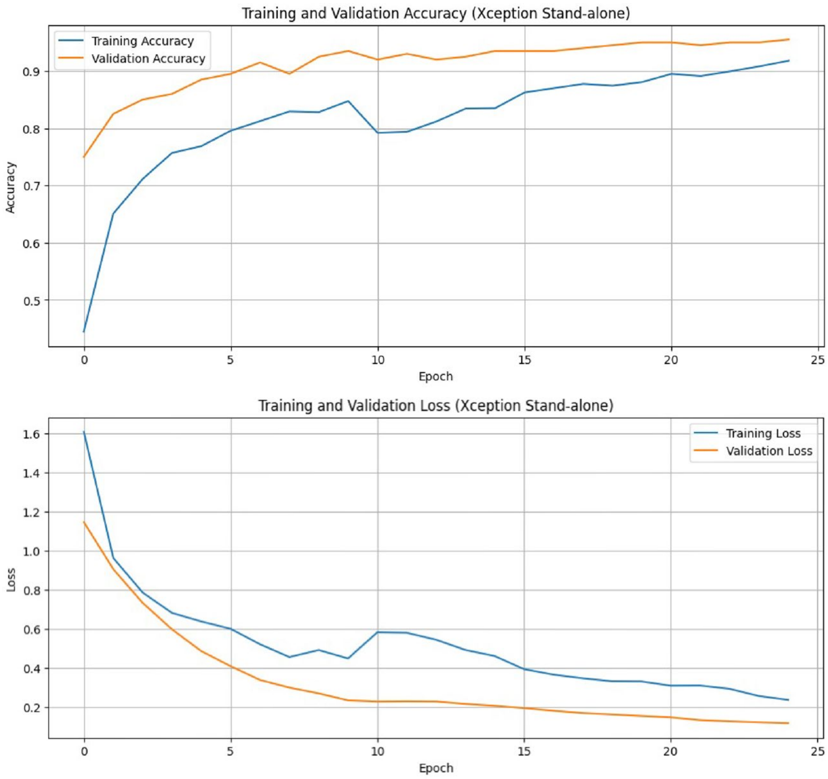

Fig. 3. The accuracy and loss of the Xception model.

Class	Precision	Recall	F1-Score
Blast	0.97	0.97	0.97
Brown Rust	0.97	0.93	0.95
Healthy	0.91	0.97	0.94
Leaf Blight	0.97	0.97	0.97
Septoria	0.95	0.93	0.94
Macro Avg	0.96	0.95	0.96

Table 3. Classification report of Xception model.

Experimental setup

The experiments for training and evaluating the proposed EffiXB3 model were conducted using Google Colab Pro, which provides a high-performance GPU environment suitable for DL workloads. The hardware configuration used during training consisted of an NVIDIA Tesla T4 GPU with 15 GB of GPU memory and approximately 51 GB of system RAM. The storage environment offered 235 GB of disk space allocated for datasets, model checkpoints, and logs. Due to the limited local computational resources, model development, preliminary code implementation, and all training were completed on the Google Compute Engine to ensure faster processing and higher batch throughput. The software environment included Python 3.11.13, TensorFlow 2.18.0, OpenCV for image preprocessing, and supporting libraries such as NumPy and scikit-learn. Performance was evaluated to assess the model's generalization ability, using metrics such as accuracy, precision, recall, F1-score, confusion matrix, and AUC-ROC.

Performance evaluation of Xception model

The standalone Xception model is evaluated for performance, achieving an accuracy of 0.95, as shown in Fig. 3, which demonstrates the model's improvement and generalization ability.

Table 3 presents the overall performance in classifying wheat leaf diseases. Despite the high scores, specific minor differences can be observed. For instance, the Brown Rust and Septoria classes show slightly lower recall values (0.93), suggesting occasional false negatives. These misclassifications may arise from visual similarities between early-stage symptoms of different diseases or overlapping leaf textures. The Healthy class also showed a slightly lower precision (0.91), indicating that a few diseased samples were incorrectly predicted as healthy.

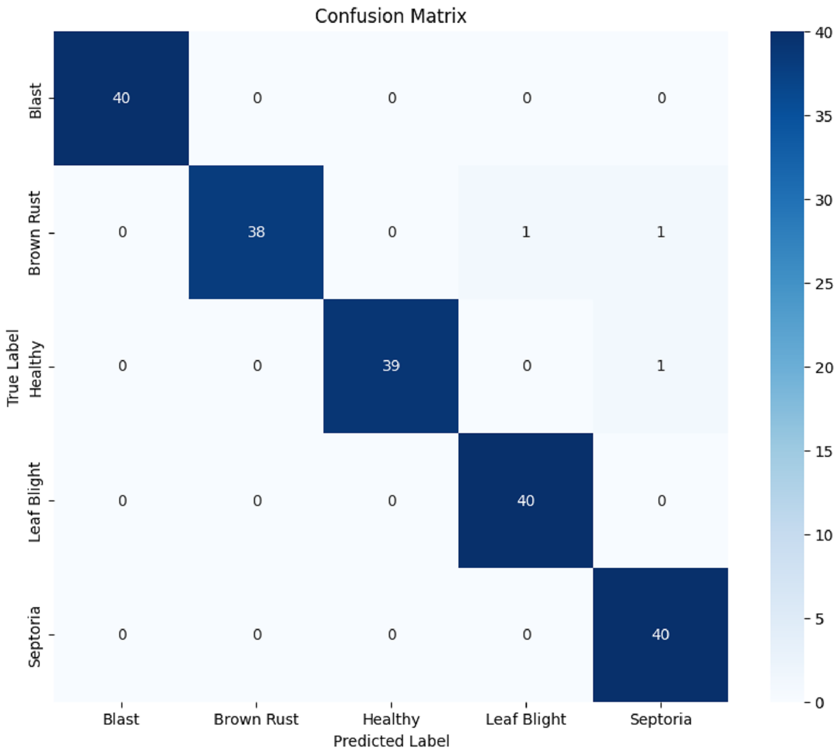

Fig. 4. Confusion matrix of Xception model.

Class	Precision	Recall	F1-Score
Blast	0.95	0.97	0.96
Brown Rust	0.90	0.90	0.90
Healthy	0.90	0.93	0.91
Leaf Blight	0.97	0.93	0.95
Septoria	0.95	0.95	0.95
Macro Avg	0.94	0.93	0.94

Table 4. Classification report of EfficientNetB3 model.

The confusion matrix of the Xception model, as shown in Fig. 4, demonstrates the true and predicted labels of each class.

Performance evaluation of efficientNetB3 model

Table 4 presents the performance of the EfficientNetB3 model, which exhibited comparatively lower generalization capability than the Xception model. Some uncertainty in convergence was also noted, likely due to the larger parameter space and higher complexity of EfficientNetB3. The model demonstrates strong performance in detecting Blast and Leaf Blight, with precision and F1-score above 95%. However, the Brown Rust and Healthy classes demonstrate relatively lower recall, indicating some misclassification among visually similar leaf symptoms. Despite these issues, the model still achieved promising results.

Figure 5 shows the accuracy and loss graphs of the standalone EfficientNetB3 model, which achieved an accuracy of 0.93. The confusion matrix of the EfficientNetB3 model, demonstrating the true and predicted labels of each class, is shown in Fig. 6.

Performance of proposed hybrid model

The classification report showed per-class F1 scores, with Septoria at 0.98, Brown Rust at 0.97, and Leaf Blight and Blast at 0.99. The proposed EffiXB3 model achieved an overall classification accuracy of 98.5%, demonstrating robust performance across all five disease categories. The macro-average F1-score is 0.98, indicating consistently high performance across all categories. The per-class evaluation metrics, including precision, recall, and F1-scores, are summarized in Table 5.

The confusion matrix in Fig. 7 shows that the majority of samples are correctly classified, with very few misclassifications between visually similar diseases such as Brown Rust and Septoria.

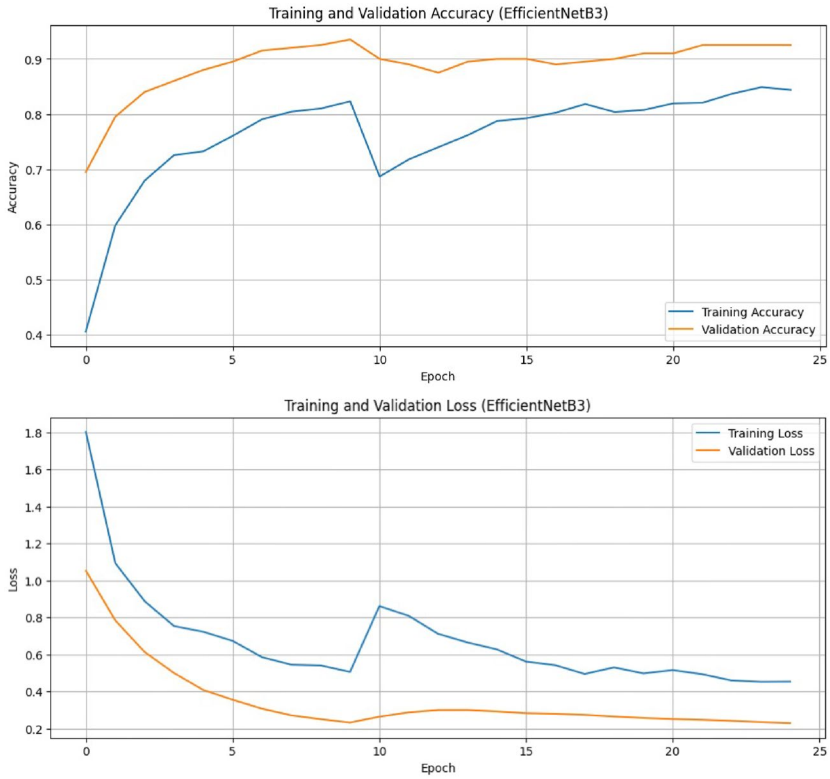

Fig. 5. Accuracy and loss of EfficientNetB3 model.

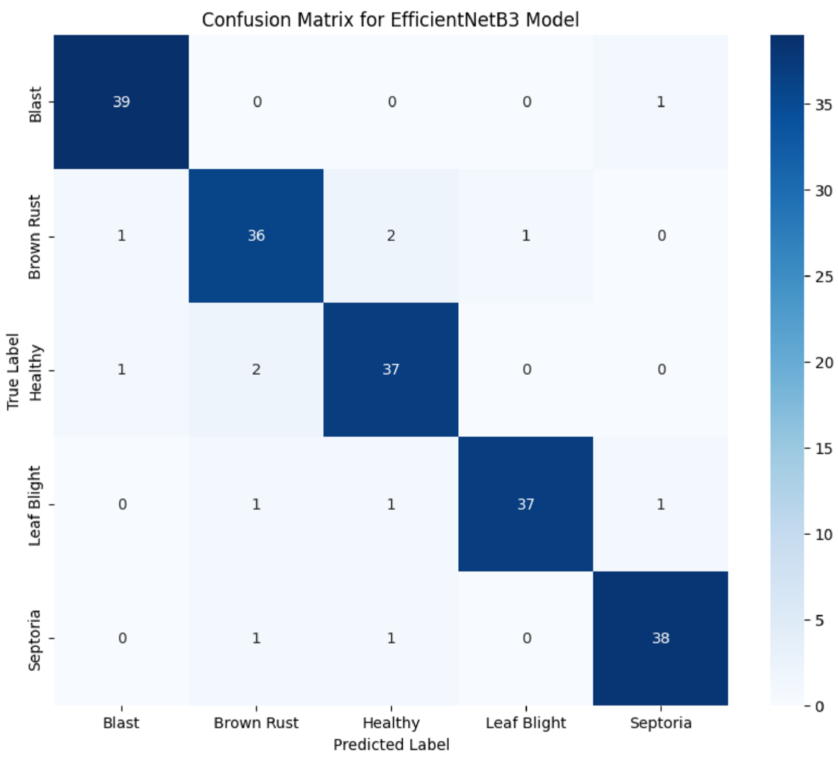

Fig. 6. Confusion matrix of EfficientNetB3 model.

Class	Precision	Recall	F1-Score
Blast	1.00	1.00	1.00
Brown Rust	1.00	0.95	0.97
Healthy	1.00	0.97	0.99
Leaf Blight	0.98	1.00	0.99
Septoria	0.95	1.00	0.98
Macro Avg	0.99	0.99	0.98

Table 5. Classification report of EffiXB3 model.

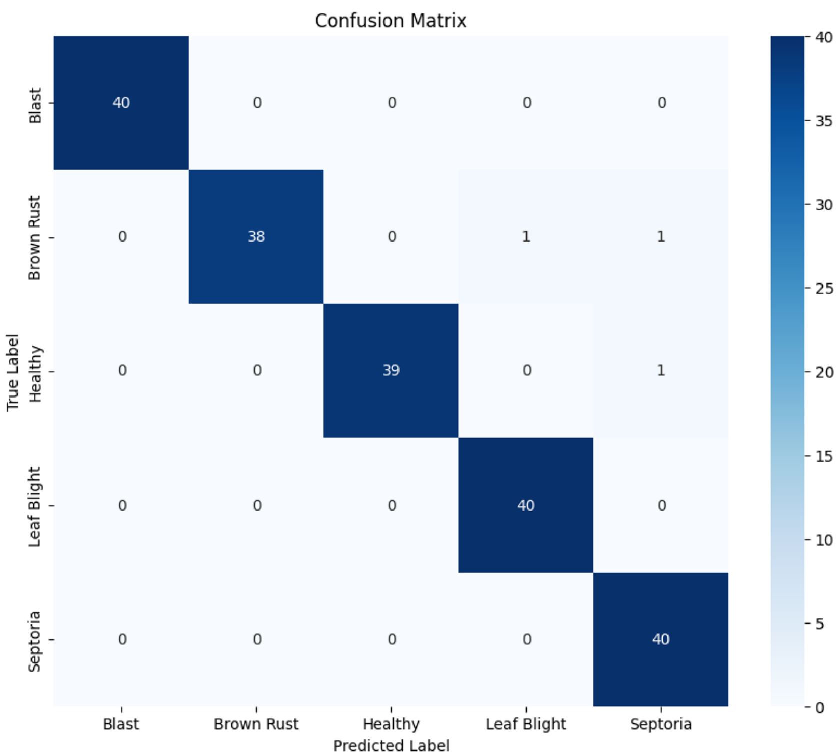

Fig. 7. Confusion Matrix of proposed model.

Notably, the inclusion of the Canny edge-enhanced stream contributed to improved detection of Septoria and Leaf Blight, supporting the hypothesis that structural edge features help disambiguate fine-grained disease characteristics. The accuracy and loss graphs per epoch of the trained model are shown in Fig. 8.

A random sample of predictions was examined to illustrate the model's confidence across different classes. Figure 9 shows example outputs where the model produced high-probability predictions for each disease type. In most cases, the predicted class probability exceeded 95%, highlighting the model's ability to distinguish categories with high certainty even when leaf textures were similar.

ROC-AUC

To further validate performance, ROC curves were computed for each disease class. The AUC values exceeded 0.98 across all categories, indicating excellent discriminative capability. Figure 10 shows the ROC curves demonstrating near-perfect separation between positive and negative samples for each disease label.

Comparison of Xception, efficientNetB3 and proposed hybrid EffiXB3 model

Figure 11 presents a comparison of the overall accuracy for the standalone Xception, which achieves 0.95 accuracy, EfficientNetB3, which obtains 0.93 accuracy, and the proposed hybrid model, which achieves remarkable accuracy.

Generalizability results with complete fifteen classes

We have rigorously assessed the generalizability of the proposed EffiXB3 model, and we conducted an additional evaluation on the complete fifteen-class dataset. This comprehensive dataset includes a broader spectrum of wheat diseases, such as Aphid, Black Rust, Common Root Rot, Fusarium Head Blight, Mildew, Mite, Smut, Stem Fly, Tan Spot, and Yellow Rust, alongside the originally studied five classes (Blast, Brown Rust, Healthy, Leaf Blight,

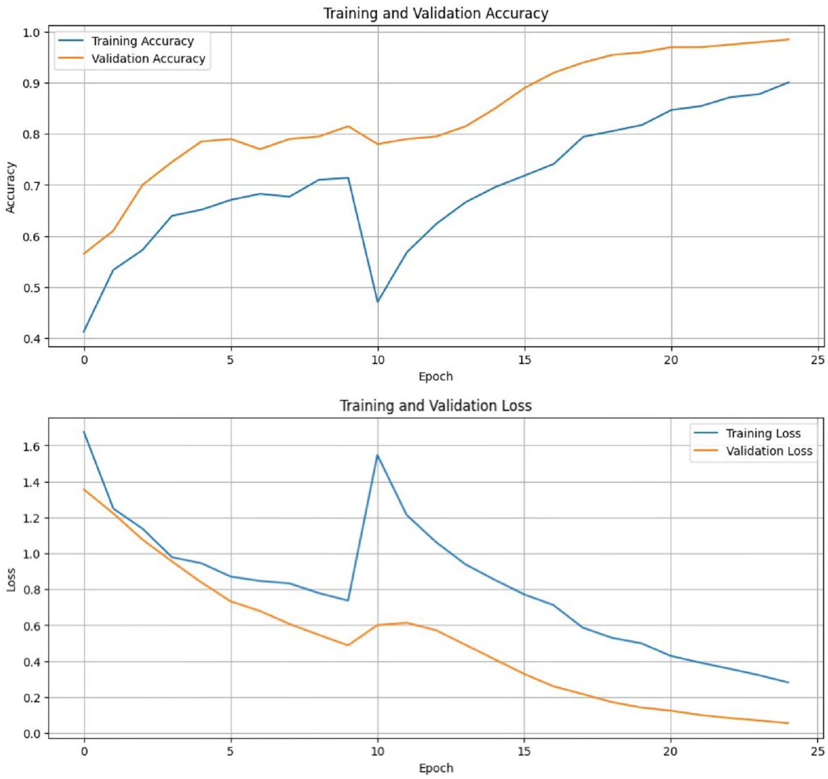

Fig. 8. Accuracy and loss of proposed model per epoch.

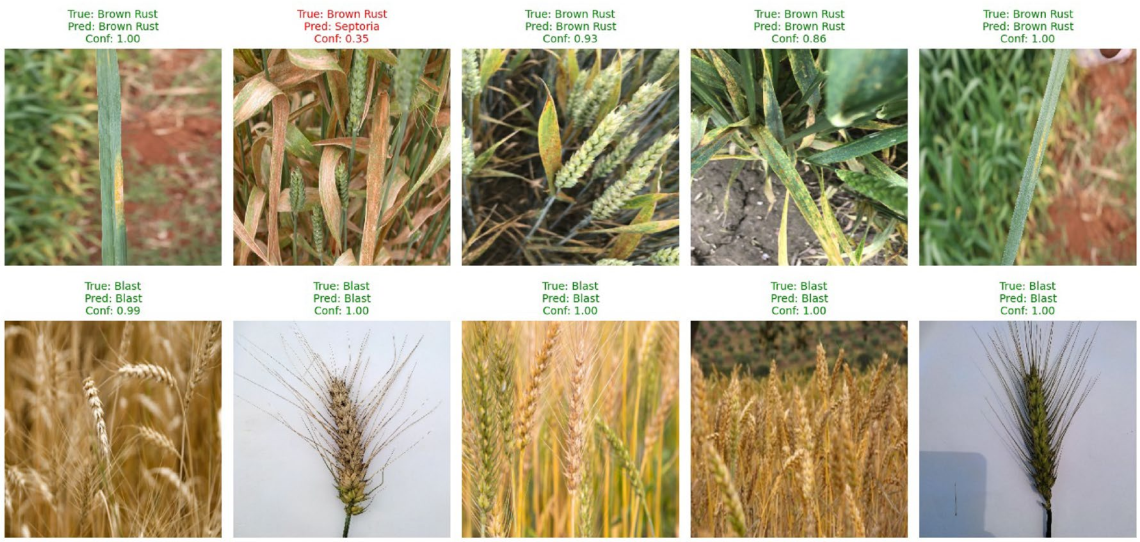

Fig. 9. Actual and predicted labels of each class.

Septoria. The hybrid EffiXB3 model is trained and evaluated using the same methodology, hyperparameters, and training strategy detailed in the methodology section. Remarkably, the model achieved a validation accuracy of 0.9917 and test accuracy of 98.50% on this more complex fifteen-class problem, with a macro-average F1-score of 0.98. The detailed classification report is provided in Table 6.

These findings suggest that the proposed hybrid architecture achieves consistent performance across all categories of wheat diseases. This generalisation ability highlights the value of combining RGB and edge-enhanced representations and confirms the potential of the EffiXB3 model. The confusion matrix in Fig. 12 shows the confusion matrix of the EffiXB3 model across 15 disease categories. The diagonal dominance indicates

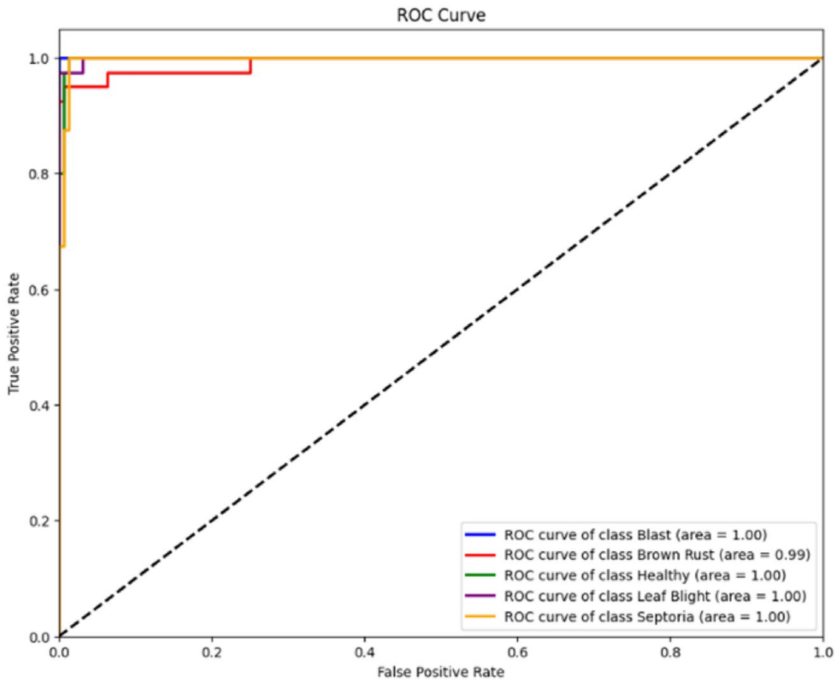

Fig. 10. ROC curves of each class showing model performance.

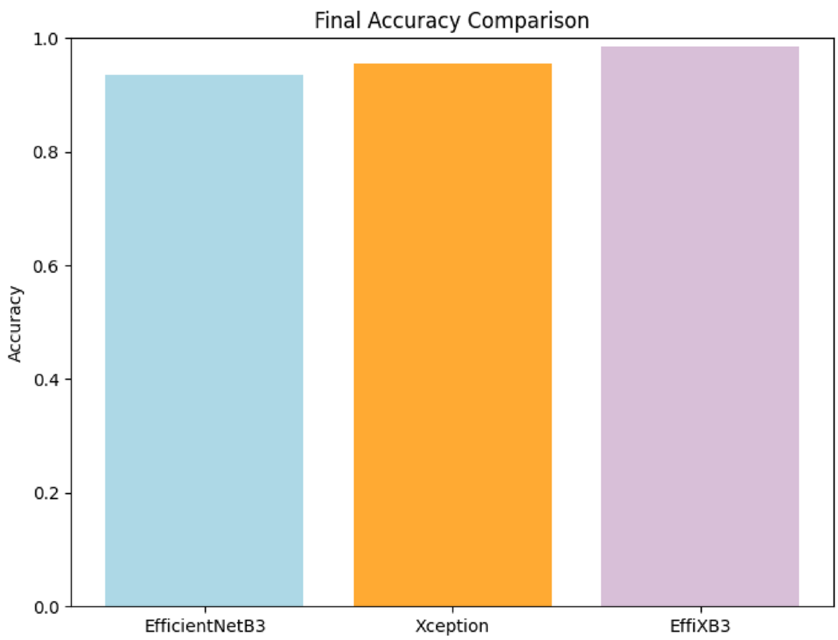

Fig. 11. Accuracy comparison: Xception vs. EfficientNetB3 vs. proposed hybrid EffiXB3 model.

that most samples were correctly classified, with only a few minor misclassifications, such as those involving Fusarium Head Blight and Common Root Rot.

State-of-the-art comparison

To compare the results of EffiXB3 with selected state-of-the-art models in the literature, we summarize the related work studies. The comparison in Table 7 demonstrates the efficient performance of the proposed model in comparison to other related DL architectures. Furthermore, by integrating the edge enhancement into the model, it can produce a comparatively good result.

Class	Precision	Recall	F1-Score
Aphid	1.00	1.00	1.00
Black Rust	0.98	1.00	0.99
Blast	0.93	1.00	0.96
Brown Rust	1.00	1.00	1.00
Common Root Rot	0.97	0.95	0.96
Fusarium Head Blight	1.00	0.93	0.96
Healthy	1.00	1.00	1.00
Leaf Blight	0.97	0.97	0.97
Mildew	1.00	1.00	1.00
Mite	0.95	0.97	0.96
Septoria	1.00	1.00	1.00
Smut	1.00	0.97	0.99
Stem fly	1.00	1.00	1.00
Tan Spot	0.97	0.97	0.97
Yellow Rust	1.00	1.00	1.00

Table 6. Classification report of the proposed model on the 15-class dataset.

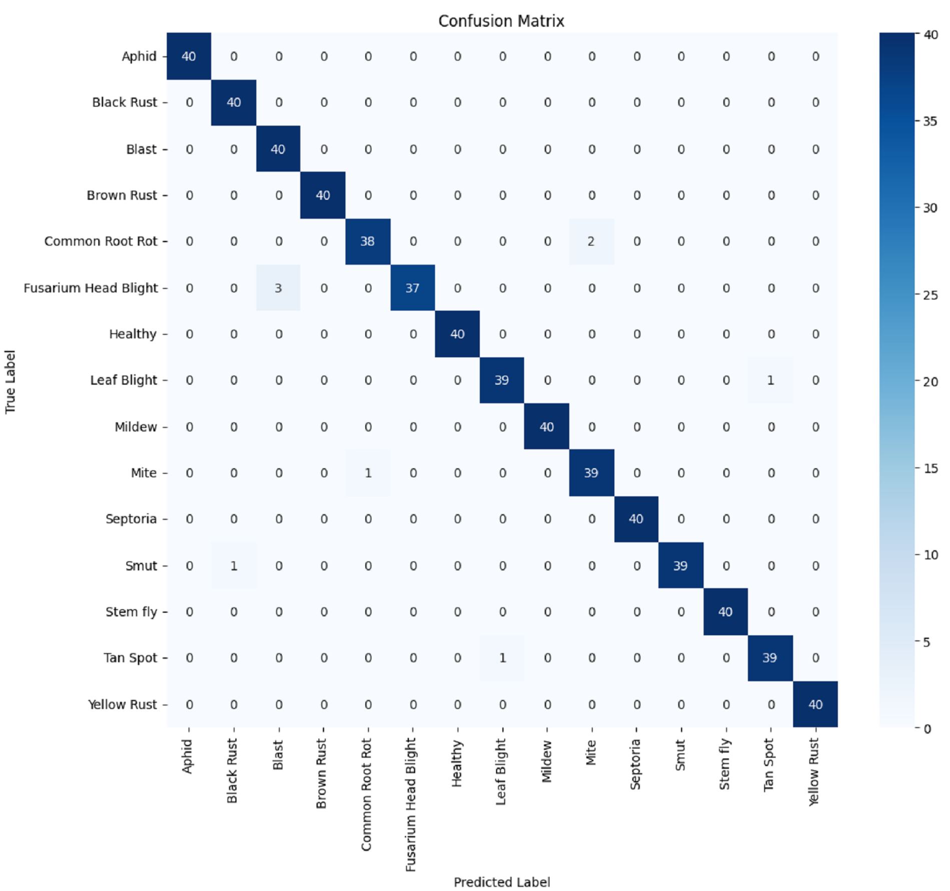

Fig. 12. Confusion matrix of the proposed EffiXB3 model evaluated on the comprehensive 15-class wheat plant disease dataset.

Discussions

The results from the proposed EffiXB3 model validate the hypothesis that integrating multiple feature extraction backbones with edge-enhanced inputs leads to significant improvements in wheat disease classification accuracy. However, the performance of both Xception and EfficientNetB3 standalone models is evaluated to assess their effectiveness in classifying wheat leaf diseases. Xception Model achieved an accuracy of 95%. It demonstrated robust generalization but exhibited some minor misclassifications in the Brown Rust and Septoria classes,

Study	Year	Model	Accuracy
[16]	2021	EfficientNet-B0	94.2%
[19]	2023	Xception and ResNet-50 model	96%
[32]	2023	DAE-Mask Edge-aware segmentation with DenseNet	96.02%
[33]	2024	Efficient Net-based CNN architecture	95.12%
[24]	2024	SC-ConvNeXt + CBAM attention mechanism	97.3%
[25]	2024	EfficientNet-B3 with a spatial attention (SA) mechanism	96.71%
[26]	2025	EfficientNetV2 model	95.2%
[27]	2025	Xception model	90%
[28]	2025	Xception model	92.5%
This	**2025**	**Xception + EfficientNetB3 hybrid model and Canny edge enhancement**	**98.50%**

Table 7. Comparison table of proposed model with related studies.

primarily due to slightly lower recall values. The Healthy class showed slightly reduced precision, suggesting some diseased samples were incorrectly labeled as healthy. Overall, it performed strongly across all classes. The EfficientNetB3 Model demonstrated good but comparatively lower performance with 93% accuracy. While it excelled in identifying Blast and Leaf Blight, it showed weaker recall for Brown Rust and Healthy classes, likely due to overlapping visual symptoms. The model also showed some signs of convergence difficulty, possibly due to its increased complexity. The classification reports and confusion matrices of both models highlight their respective strengths and weaknesses, providing a foundation for evaluating the performance of the proposed hybrid model.

The proposed EffiXB3 model achieved an overall accuracy of 98.5%, representing a notable approach for detecting multi-class diseases in wheat crops. These results highlight the benefit of combining RGB texture information with explicit structural features from edge detection. The dual-input approach, including Canny edge-enhanced images, remarkably reduced confusion between Healthy and Septoria classes, which often share overlapping textures. This suggests that edge information provides complementary details that some RGB-based models may overlook.

Comparisons with existing state-of-the-art approaches further underscore the strength of this hybrid design. While EfficientNet-based and attention-augmented models have reported accuracies around 94–97% in similar studies, EffiXB3 demonstrates that combining two high-capacity CNNs in parallel with thoughtful preprocessing can achieve superior performance without introducing unnecessary complexity. The results show that each component of the model contributes incrementally to the overall accuracy. Supporting the idea that blending spectral and structural features improves robustness.

The consistent precision and recall scores across all disease classes in the dataset indicate that the model generalizes well rather than overfits to specific categories. Analysis of the confusion matrix showed that false positives in these categories decreased significantly, demonstrating improved classification. Additionally, the consistently high AUC scores from the ROC analysis across all disease classes highlighted the model's robustness and reliability in distinguishing between disease types and healthy leaves. Overall, the experimental results suggest that EffiXB3 is an effective solution for automated wheat disease detection systems.

To thoroughly address concerns regarding generalizability and scope, we expanded our evaluation to the full fifteen-class dataset. On this comprehensive task, the proposed EffiXB3 model achieved validation accuracy of 0.9917 and test accuracy of 0.9850, representing a notable and highly generalizable approach for detecting multi-class diseases in wheat crops. The model demonstrated exceptional performance across nearly all categories, achieving perfect F1-scores (1.00) for Aphid, Brown Rust, Healthy, Mildew, Septoria, Stem fly, and Yellow Rust. High performance was also maintained for more challenging classes, with F1-scores of 0.99 for Black Rust and Smut, and 0.97–0.98.97.98 for Leaf Blight and Tan Spot. The most significant misclassifications occurred in Blast (F1-score: 0.96) and Fusarium Head Blight (F1-score: 0.96), likely due to their more variable visual symptoms. These results highlight the benefit of combining RGB texture information with explicit structural features from edge detection.

While EffiXB3 achieved excellent accuracy, some limitations remain. Although the dataset was balanced, its overall size is still quite limited compared to the broad variability encountered in real-world field conditions. The practical deployment of the proposed model in real-world agricultural settings presents specific challenges. Field-acquired images often suffer from variable lighting conditions, occlusions such as by soil or other plants, and complex background clutter, which are not prominent in our curated dataset. This challenge of transitioning from controlled lab conditions to variable field environments is a well-recognized hurdle in the broader domain of agricultural AI and big data.

Furthermore, the Canny edge detection algorithm, although effective in this study, can be sensitive to image noise and may perform inconsistently under suboptimal lighting conditions. These factors could impact the model's robustness and prediction reliability in practice. Additionally, the use of Canny edge detection is effective in this study; however, it can be sensitive to image noise and poor lighting, which may impact its performance in practice. The computational requirements of running dual deep CNN backbones may make it challenging to deploy the model on low-resource devices without further optimization.

Conclusions and future work

This study introduces EffiXB3, a hybrid dual-input DL model designed for the automated classification of wheat plant diseases. Xception and EfficientNetB3 models were individually assessed to compare their effectiveness in multi-class classification. The results showed that the Xception model achieved a classification accuracy of 0.95, while the EfficientNetB3 model reached 0.93 in accuracy. By combining Xception and EfficientNetB3 architectures in a parallel configuration, EffiXB3 is capable of learning both texture RGB features and structural edge-based patterns derived from Canny transformations. Experimental results demonstrated that the model achieved a classification accuracy of 98.50%, outperforming several recent DL approaches while maintaining moderate computational demands. The study further confirmed that integration of edge-enhanced representations significantly contributes to improved predictive performance, particularly in distinguishing diseases with similar visual appearances. Overall, EffiXB3 provides a practical framework for detecting wheat leaf diseases and lays the groundwork for further enhancements. Despite the strong performance on a controlled dataset, there are practical considerations for real-world deployment. Field images often contain variable lighting, occlusions, and background clutter that could impact prediction reliability.

Future research could explore domain adaptation methods or further data augmentation to improve resilience under these conditions. Incorporating attention mechanisms or explainability techniques may also enhance model interpretability. Additionally, domain adaptation and style transfer techniques will be explored to enhance the model's ability to generalize across diverse environmental conditions, including variations in lighting, weather, and crop growth stages.

Data availability

A publicly available dataset on Wheat Plant Diseases[31] from Kaggle is utilized in this study.

Received: 21 July 2025; Accepted: 27 October 2025
Published online: 26 November 2025

References

1. Kukreja, D. K. & Kukreja, V. Image segmentation, classification, and recognition methods for wheat diseases: Two decades' systematic literature review. *Computers and Electronics in Agriculture* 109005 (2024).
2. Figueroa M, H.-K. K. S. P. A review of wheat diseases-a field perspective. *Mol Plant Pathol* 1523–1536 (2018).
3. Sharma, K. P. & N. Crop health detection using image processing and machine learning for better yield production. *Cureus Journal Of Computer Science* (2025).
4. Chouhan, S. S., Singh, U. P., Saxena, A. & Jain, S. Assessing the importance and need of artificial intelligence for precision agriculture. In *Artificial intelligence techniques in smart agriculture*, 1–6 (Springer, 2024).
5. Ramzan, S., Ghadi, Y. Y., Aljuaid, H., Mahmood, A. & Ali, B. An ingenious iot based crop prediction system using ml and el. *Computers, Materials and Continua* **79**, 183–199 (2024).
6. Belkher, S. et al. Deficit irrigation and an iot-based system for improving yield and water use efficiency of winter wheat grown under semi-arid conditions. *Agronomy* **15** (2025).
7. Ramzan, S. et al. An innovative artificial neural network model for smart crop prediction using sensory network based soil data. *PeerJ Computer Science* (2024).
8. Leite, D., Brito, A. & Faccioli, G. Advancements and outlooks in utilizing convolutional neural networks for plant disease severity assessment: A comprehensive review. *Smart Agricultural Technology* **9** (2024).
9. Nouman, B. et al. Citrus diseases detection using innovative deep learning approach and hybrid meta-heuristic. *PLoS ONE* **20** (2025).
10. Rehman, S. et al. Exploring the frontier of wheat rust resistance: Latest approaches, mechanisms, and novel insights. *Plants* **13** (2024).
11. Vicentini, S. et al. Aerobiology of the wheat blast pathogen: Inoculum monitoring and detection of fungicide resistance alleles. *Agronomy* **13** (2023).
12. Kokhmetova, A. et al. Identification of wheat septoria tritici resistance genes in wheat germplasm using molecular markers. *Plants* **13** (2024).
13. Tan, M. & Le, Q. V. Efficientnet: Rethinking model scaling for convolutional neural networks. *International conference on machine learning* (2020).
14. Chollet, F. Xception: Deep learning with depthwise separable convolutions. *Proceedings of the IEEE conference on computer vision and pattern recognition* (2017).
15. Chouhan, S. S., Singh, U. P. & Jain, S. Performance evaluation of different deep learning models used for the purpose of healthy and diseased leaves classification of cherimoya (annona cherimola) plant. *Neural Computing and Applications* **37**, 4531–4544 (2025).
16. Genaev, M. et al. Image-based wheat fungi diseases identification by deep learning. *Plants* **10** (2021).
17. Mao, R. et al. Dae-mask: a novel deep-learning-based automatic detection model for in-field wheat diseases. *Precision Agriculture* **25**, 1–26 (2023).
18. Chekir, A., Djelouat, D. F. & Djemame, F.-J. Deep learning models for wheat diseases detection and recognition. In *Electrical Engineering and Computer Systems and Sciences* (2023).
19. Shafi, U. *et al.* Embedded ai for wheat yellow rust infection type classification. *IEEE Access* 23726–23738 (2023).
20. Kaur, G., Sharma, N. & Gupta, R. Wheat leaf disease classification using efficientnetb3 pre-trained architecture. In *International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering*, 1–5 (2023).
21. Khan, M. O. et al. Detection of wheat leaf disease using transfer learning: A comparative study. In *27th International Conference on Computer and Information Technology*, 2653–2658 (2024).
22. Ramkumar, G., Swarna, M., Vijayashanthi, R. S. & L, H. M. An intensive model for wheat crop disease prediction based on hybrid deep learning principle with image processing support. In *2nd International Conference on Emerging Trends in Engineering and Medical Sciences*, 597–601 (2024).
23. Kaur, R., Kumar, S., Joon, D. & Pundir, M. Advancing wheat leaf disease classification for sustainable agriculture using efficientnet. In *2024 International Conference on Computing, Sciences and Communications (ICCSC)*, 1–5 (2024).
24. Dong, T. et al. Wheat disease recognition method based on the sc-convnext network model. *Scientific Reports* (2024).
25. Islam, M. et al. Optimal features driven attention network with medium-scale benchmark for wheat diseases recognition. *IEEE Access* 150739–150753 (2024).

26. Ms, S. K., Santhiya, A., Sheebha Shanthini, R., Sheshapriya, N. & Varsha, V. Optimizing wheat rust disease detection with efficient net. *International Research Journal on Advanced Engineering Hub* **3** (2025).
27. Syeda, F., Jameel, A., Alani, N., Humayun, M. & Alwakid, G. N. Automated detection and severity prediction of wheat rust using cost-effective xception architecture. *Plant, Cell & Environment* **48**, 4126–4139 (2025).
28. Chad, F. Enhancing wheat disease surveillance: A scalable xception architecture for rust detection and severity prediction. *ResearchGate* (2025).
29. Chouhan, S. S., Kaul, A. & Singh, U. P. Image segmentation using fuzzy competitive learning based counter propagation network. *Multimedia Tools and Applications* **78**, 35263–35287. https://doi.org/10.1007/s11042-019-08094-y (2019).
30. Sellam, V., Kannan, N., Pandi, S. S. & Manju, I. Enhancing sustainable agriculture using attention convolutional bidirectional gatedrecurrent based modified leaf in wind algorithm: Integrating ai and iot for efficient farming. *Sustainable Computing: Informatics and Systems* 101160 (2025).
31. Kushagra3204. Wheat plant diseases. https://www.kaggle.com/ (2024).
32. Mao, R. et al. Dae-mask: a novel deep-learning-based automatic detection model for in-field wheat diseases. *Precision Agriculture* **25**, 1–26 (2023).
33. Kaur, R., Singh, K., Joon, D. & Pundir, M. Advancing wheat leaf disease classification for sustainable agriculture using efficientnet. In *2024 International Conference on Computing, Sciences and Communications (ICCSC)*, 1–5 (2024).

Acknowledgements

This work was supported by the Cooperative Research Program for Agriculture Science and Technology Development (Project No. RS-2025-02223510, Expansion of Big Data for Pest and Disease Image Diagnosis and Enhancement of Data Sharing Technology (Phase 2)) of the Rural Development Administration, Republic of Korea.

Author contributions

Conceptualization, A.R. (Ayesha Razaq), B.A., S.R., A.R. (Ali Raza), N.L.F., M.S. and Y.G.; Methodology, A.R. (Ayesha Razaq), B.A., S.R., A.R. (Ali Raza), N.L.F., M.S. and Y.G.; Software, A.R. (Ayesha Razaq), B.A., A.R. (Ali Raza) and N.L.F.; Validation, S.R., A.R. (Ali Raza), and M.S.; Formal analysis, A.R. (Ayesha Razaq), B.A., A.R. (Ali Raza) and N.L.F.; Investigation, S.R.; Data curation, A.R. (Ali Raza); Writing—original draft, A.R. (Ayesha Razaq), B.A., S.R., A.R. (Ali Raza) and N.L.F.; Writing—review & editing, M.S. and Y.G.; Visualization, A.R. (Ayesha Razaq) and N.L.F.; Supervision, S.R., M.S. and Y.G.; Funding acquisition, M.S. and Y.G. All authors have read and agreed to the published version of the manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Y.G. or M.S.