A Hybrid Transfer Learning Framework for Corn Crop Detection Using Deep Convolutional Networks

K. C. Praveen Kumar
Mohan Babu University, Tirupati, India
praveenkc23@gmail.com (corresponding author)

Y. Dileep Kumar
Mohan Babu University, Tirupati, India
y.dileepkumar1983@gmail.com

Received: 27 March 2026 | Revised: 3 May 2026 and 21 May 2026 | Accepted: 22 May 2026

 DOI: https://doi.org/10.48084/etasr.18970

ABSTRACT

Corn is a staple crop of global significance; however, foliar diseases may lead to 30–60% yield loss if not detected at an early stage. Conventional visual inspection is time-consuming, subjective, and difficult to scale for smallholder farmers worldwide. To overcome these issues, we present Corn Transfer Learning Network (CTL-Net), an end-to-end hybrid deep learning model for corn leaf disease identification. CTL-Net, which combines Inception-ResNet-v2 as the backbone and MobileNetV3 as a feature extractor in parallel convolutional streams, simultaneously learns diverse scales of texture information from low-level textures, mid-level structural patterns, and high-level disease semantics of RGB leaf images. Adaptive feature fusion is formulated through learnable weighting coefficients and bi-directional spatial–channel attention mechanisms, which further enhance feature discriminability and robustness. The proposed approach is tested on an extensive dataset of 12,456 images from 10 corn diseases, including Northern Leaf Blight, Common Rust, Gray Leaf Spot, and Cercospora Leaf Spot, acquired under controlled and real-field conditions. CTL-Net attains the highest classification accuracy of 99.42%, outperforming DenseNet121 (97.92%), EfficientNetB3 (97.35%), and general stacking models (97.89%). Robustness experiments demonstrate the effectiveness of the proposed method against illumination variations, additive noise, and partial occlusions. CTL-Net enables real-time inference with a latency of 42 ms on an NVIDIA RTX 3090 GPU. Gradient-weighted Class Activation Mapping++ (Grad-CAM++)-based interpretability analysis results in a mean Intersection over Union (IoU) of 87.6% with expert-annotated disease regions. Five-fold cross-validation, ablation studies, and statistical significance testing ($p < 0.001$) demonstrated that the contribution of each architectural component is significant. To summarize, CTL-Net allows accurate and interpretable predictions in a computationally efficient manner, making it an excellent candidate for mobile and resource-limited applications in precision agriculture.

Keywords-corn diseases; hybrid transfer learning; deep learning; attention mechanism; feature fusion; precision agriculture; computer vision; MobilenetV3; Inception-ResNet-v2; multiscale learning

I. INTRODUCTION

Agriculture is the basis of the world's food supply and supports more than 3.2 billion people worldwide, generating approximately $3.6 trillion in income for the global economy. Corn (*Zea mays* L.), a major cereal crop, has a substantial global impact. It is cultivated on approximately 197 million hectares worldwide and produces over 1.16 billion metric tons annually. As a major component of global food systems, corn is used as livestock feed and as feedstock for industrial applications. In developing nations, corn provides the primary energy source for over 900 million people. In addition, corn supports a vast supply chain for starch, pharmaceutical products, and biofuels. The economic importance of corn is reinforced by the fact that the largest producers (the United States, China, Brazil, Argentina, and India) account for approximately 75% of global corn production. In 2021, corn was grown on approximately 90 million acres in the United States alone, with an estimated annual production value of over $52 billion.

Globally regarded as an important staple crop, corn production faces ongoing challenges from many different types of plant pathogens. These pathogens include fungi (e.g., *Exserohilum turcicum*, *Puccinia sorghi*, and *Cercospora zeae-maydis*), bacteria (e.g., *Pantoea stewartii*), viruses (e.g., *Maize Dwarf Mosaic Virus*), and oomycetes (e.g., *Peronosclerospora sorghi*). Foliar diseases have the greatest impact on corn

production because they reduce photosynthesis, nutrient transport, and overall plant health, thereby limiting plant growth and grain production.

Conventional methods for disease identification rely primarily on visual inspection by trained experts. These methods have several limitations that affect their reliability. The training required to achieve accurate diagnosis is extensive, and field-based diagnosis requires substantial education and experience. Meanwhile, many regions of the world experience chronic shortages of agricultural extension agents, with reported ratios ranging from 1:200 to 1:5,000 farmers per extension worker. Additionally, agreement among experts is highly variable; even after training, agreement rates range from 65% to 80%, particularly during the early stages of disease, when symptoms are less pronounced. Manual inspection also requires a significant investment of time. For instance, during large-scale disease outbreaks affecting extensive cultivation areas, delays in detection may prevent producers from applying treatments within the optimal intervention period. Furthermore, several diseases share similar symptoms, particularly chlorotic or necrotic patterns, leading to diagnostic confusion in the absence of laboratory testing. Polymerase chain reaction (PCR), enzyme-linked immunosorbent assay (ELISA), and culture-based testing are accurate; however, these methods are expensive, require specialized equipment, and may take several days to complete. Therefore, they do not satisfy the requirements for rapid and large-scale disease assessment.

These challenges have motivated extensive research into automated methods for corn disease detection. Researchers have explored numerous automated methods for disease detection using machine learning and image processing. Early studies employed models such as Artificial Neural Networks (ANNs) [1] and Support Vector Machines (SVMs) [2], in which images were preprocessed, and features such as color and texture were analyzed. Later, more advanced approaches based on Convolutional Neural Networks (CNNs) [3] improved detection performance by learning features directly from images. Other methods, including SVMs combined with Scale-Invariant Feature Transform (SIFT) [4] and genetic algorithms [5], also demonstrated high accuracy, although some required substantial computational resources.

Some studies focused on extracting specific features, such as RGB color analysis [6], whereas others applied machine learning techniques to predict crop yield [7], estimate environmental conditions such as soil moisture [8], or recommend suitable crops [9].

Deep learning has further enhanced disease detection across various crops. Models such as CNNs [10], hybrid machine learning techniques [11], random forests [12], and deep belief networks [13] have achieved high accuracy. Similarly, SVM-based methods [14, 15] have helped farmers diagnose diseases effectively.

Recent research has emphasized more advanced and integrated approaches. These include machine learning frameworks for maize disease detection [16], Internet of Things (IoT)-based monitoring systems [17], optimized hybrid models [18-21], and transfer learning techniques that reduce data requirements [22-24]. Advanced models such as You Only Look Once (YOLO) [25] and hybrid deep learning systems [26] now achieve very high accuracy, making disease detection faster, more intelligent, and more practical for real-world agricultural applications.

Although substantial advances have been made in corn disease detection, several gaps remain in the literature. Most studies employ a single-model architecture or a simple late-fusion ensemble and do not provide deep, interactive feature-level integration. In addition, attention-based mechanisms have been underutilized in corn-specific models for distinguishing lesions from background noise and improving disease discrimination. Furthermore, most existing datasets are collected under controlled laboratory conditions or derived from public datasets such as PlantVillage, which do not adequately represent real-world variability, including inconsistent illumination, complex backgrounds, and varying disease severity levels. Moreover, many deep learning models remain opaque, making it difficult for end users to understand how predictions are generated and limiting their practical adoption. Finally, limited information is available on developing high-performance models suitable for mobile or edge deployment, despite the growing importance of on-site and early disease detection systems.

II. PROPOSED METHODOLOGY

The proposed Corn Transfer Learning Network (CTL-Net), as shown in Figure 1, is a dual-stream hybrid deep learning framework that integrates the Inception-ResNet-v2 and MobileNetV3-Large models in a unified architecture to extract complementary multi-scale features for robust corn disease classification.

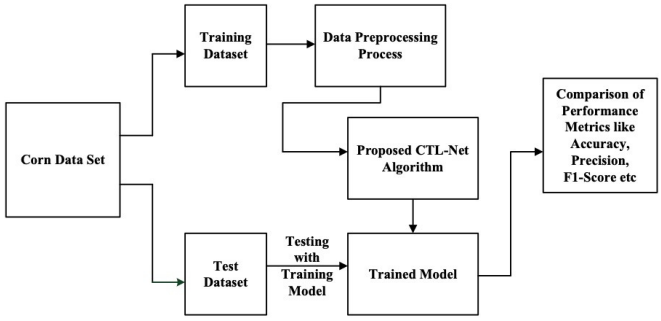

Fig. 1. Proposed CTL-NET arcitecture.

The rationale for this design is that Inception-ResNet-v2 is effective in learning multi-scale semantic structures through parallel convolutional pathways, whereas MobileNetV3 captures fine-grained textural cues using inverted residual blocks and squeeze-and-excitation modules. These features are particularly important for distinguishing lesion types, including punctate spots for Cercospora leaf disease, elongated necrotic regions for Northern Leaf Blight, and orange urediniospore pustules for Common Rust.

CTL-Net processes corn leaf images through a defined pipeline. Input images are simultaneously passed through the

Inception-ResNet-v2 and MobileNetV3 branches to extract complementary semantic and textural features. At each of the three hierarchical spatial scales, bi-directional cross-stream attention enables mutual refinement between the two streams. The resulting representations are further refined using channel–spatial attention and are fused through adaptive learnable gates, followed by hierarchical integration. The final fused representation is then fed into both a disease classification head and an auxiliary severity prediction head, enabling multi-task learning and improved generalization.

All input images $I \in \mathbb{R}^{H\times W\times 3}$ are resized to 256 × 256 × 3 and standardized channel-wise using the ImageNet mean and standard deviation. The regularized image is computed as in (1):

$$I_{\text{norm}}(x,y,c) = \frac{I_{\text{raw}}(x,y,c)-\mu_c}{\sigma_c} \quad (1)$$

where $\mu = \{0.485, 0.456, 0.406\}$, $\sigma = \{0.229, 0.224, 0.225\}$. This normalization ensures consistency with pretrained backbones and stabilizes gradient propagation. Data augmentation is applied using a comprehensive pipeline, including rotation, flipping, zooming, brightness and contrast adjustment, Gaussian noise, blur, and color jittering to simulate real field conditions.

The dual-stream backbone processes the normalized image in parallel. Inception-ResNet-v2 produces three hierarchical feature maps $\{F_I^{(1)}, F_I^{(2)}, F_I^{(3)}\}$ with spatial resolutions 128 × 128, 64 × 64, and 32 × 32, corresponding to channel depths of 256, 512, and 1,536 respectively. At the same time, MobileNetV3-Large generates aligned feature maps $\{F_M^{(1)}, F_M^{(2)}, F_M^{(3)}\}$ with channel dimensions of 24, 40, and 160. To enable meaningful interaction between these streams, MobileNet features are upsampled when necessary to match the spatial dimensions of the corresponding Inception-ResNet-v2 features.

Feature fusion is achieved using a bidirectional cross-stream attention mechanism. At each scale (l), query, key, and value projections are computed as:

$$Q_I = F_I^{(l)} W_Q^I, K_M = F_M^{(l)} W_K^M, V_M = F_M^{(l)} W_V^M \quad (2)$$

where W_Q^I, W_K^M, and W_V^M are learnable weight matrices with a dimensionality reduction ratio $r = 8$. The scaled dot-product attention is then computed as:

$$A = \text{softmax}\left(\frac{Q_I K_M^T}{\sqrt{d_k}}\right) \quad (3)$$

where d_k denotes the dimension of the key vectors. The attended MobileNet features that refine Inception-ResNet-v2 representations are obtained via (4):

$$Z_{I\leftarrow M} = A\, V_M \quad (4)$$

These refined features are integrated through a residual enhancement mechanism, as in (5):

$$F_{I,\text{ref}}^{(l)} = F_I^{(l)} + \gamma_1 Z_{I\leftarrow M} \quad (5)$$

This bidirectional interaction enables both backbones to contribute complementary information, combining semantic structure from Inception-ResNet-v2 and fine-texture discrimination from MobileNetV3.

After cross-attention fusion at each scale, CTL-Net performs hierarchical multi-scale integration using adaptive learnable weights, α_l and β_l, producing the final fused representation:

$$F_{\text{fusion}} = \sum_{l=1}^{3} (\alpha_l F_{I,\text{ref}}^{(l)} + \beta_l F_{M,\text{ref}}^{(l)}) \quad (6)$$

The fused features are globally averaged, fed to a dense classification layer, and optimized using cross-entropy loss, defined as:

$$\mathcal{L}_{CE} = -\sum_{c=1}^{C} y_c \log(\hat{y}_c) \quad (7)$$

where y_c is the ground-truth label and $\hat{y}_c$ is the predicted class probability.

Through biologically informed augmentation, dual-stream representation learning, and mathematically grounded feature interaction using cross-attention, CTL-Net produces a highly discriminative and generalized representation for automated corn leaf disease classification under realistic field conditions.

III. EXPERIMENTAL METHODOLOGY

A. Data Collection

The dataset used in this study was constructed from multiple sources to ensure diversity and real-world applicability. A significant portion (4,188 images) was obtained from the PlantVillage dataset, publicly available at [27] and further discussed in [28], which provides high-quality images captured under controlled laboratory conditions. To improve generalization capability, additional images were collected from real-field environments exhibiting variations in illumination, background complexity, and disease severity. Furthermore, supplementary samples were incorporated from the publicly available Kaggle repository "Corn or Maize Leaf Disease Dataset" [29], enabling broader coverage of disease variations. The final dataset used in this study is publicly available and can be accessed at [30].

By combining controlled and real-field data, the proposed dataset enables the model to learn more robust and transferable feature representations, addressing limitations of prior studies that rely exclusively on laboratory-based datasets. This hybrid data strategy enhances real-world deployment capability, as shown in Table I.

The experimental dataset is therefore composed of two primary publicly available sources: the PlantVillage dataset [27], and the Corn or Maize Leaf Disease Dataset [29], both containing labeled maize leaf images for disease classification. The PlantVillage subset is additionally used to ensure benchmarking consistency and experimental reproducibility across studies. For this work, only images corresponding to relevant maize disease categories and healthy samples were retained. Priority was given to annotation correctness, visual clarity, and class relevance. Images with ambiguous labels, duplicates, excessive noise, low resolution, or unrepresentative disease symptoms were excluded from the final dataset. This carefully selected subset was designed to ensure class balance

and reduce bias during training and testing. In addition, standard preprocessing techniques such as image resizing, normalization, and stratified sampling were applied to all samples to eliminate scale variability and improve model robustness.

The dataset was split using a 70% training, 15% validation, and 15% testing ratio with stratified sampling to ensure class balance across all partitions.

TABLE I. CLASS-WISE DATASET DISTRIBUTION AND DATA SPLITS

Class ID	Disease name	Total samples	Training set	Validation set	Testing set
0	Healthy	1,586	1,110	238	238
1	Northern Leaf Blight	1,442	1,009	217	216
2	Common Rust	1,328	930	199	199
4	Cercospora Leaf Spot	1,088	762	163	163
5	Stewart's Wilt	982	687	148	147
6	Anthracnose	1,124	787	169	168
7	Eyespot	896	627	135	134
8	Northern Corn Leaf Spot	1,068	748	160	160
9	Southern Rust	688	482	103	103

B. Data Preprocessing Pipeline

A standard preprocessing pipeline was applied to all images in the dataset to ensure consistency across different acquisition sources. First, all images were resized to 256 × 256 pixels using bilinear interpolation. To avoid aspect ratio distortion, center cropping was applied where necessary. Subsequently, spatial normalization was performed, followed by intensity normalization using ImageNet statistics. All images were normalized using the same formulation defined in (1), which ensures consistency with pretrained backbone networks and stabilizes gradient propagation during training.

C. Training Infrastructure and Configuration

All models were trained using the same computational infrastructure and training protocol to ensure a fair and consistent comparison. The proposed CTL-Net and all baseline models shared identical hyperparameters, loss functions, and training callbacks. CTL-Net converged in approximately 18.4 h, whereas the training times of the baseline models varied depending on model depth and parameter complexity. This controlled setup ensures a reliable and reproducible benchmarking of the proposed architecture.

All experiments were conducted under a standardized training protocol. The dataset was split using stratified sampling into 70% training, 15% validation, and 15% testing subsets to preserve class balance across all partitions. All models were trained under identical conditions, including:

- Optimizer: Adam
- Learning rate: 0.0001
- Batch size: 32
- Epochs: 50 (with early stopping applied)
- Loss function: categorical cross-entropy

Performance metrics were computed on the held-out test set, which was not used during training or validation. In addition, 5-fold cross-validation was performed to ensure robustness and statistical reliability of the results. Each experiment was repeated three times, and the average performance values were reported. Statistical significance was evaluated using McNemar's test ($p < 0.001$).

This standardized experimental setup ensures that the reported performance improvements of CTL-Net over 14 baseline models are fair, reproducible, and unbiased.

D. Classification Performance Metrics

To evaluate the performance of CTL-Net and all baseline models, a complete set of classification performance metrics, including overall accuracy and class-level discriminative capabilities, was determined. Overall classification accuracy is defined as the ratio of correctly classified samples to the total number of samples (global correctness) and is expressed as:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (8)$$

Class-wise performance in terms of positive prediction quality is evaluated using precision, defined as:

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

which is critical for reducing false alarms in disease identification. Sensitivity to correctly detected diseased samples is evaluated using recall, defined as:

$$\text{Recall} = \frac{TP}{TP+FN} \quad (10)$$

which is particularly important for ensuring that diseased plants are not overlooked. A balanced trade-off between precision and recall is captured using the F1-score, formulated as:

$$\text{F1} - \text{score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (11)$$

which is especially useful under class imbalance conditions. The ability of the model to correctly identify healthy samples is measured using specificity:

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (12)$$

To provide a single robust indicator of overall classification performance, particularly under imbalanced datasets, the Matthews Correlation Coefficient (MCC) is employed:

$$\text{MCC} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad (13)$$

where values range from −1 to +1, with +1 indicating perfect prediction, 0 indicating random performance, and −1 indicating complete disagreement between predictions and ground truth. Additionally, Cohen's Kappa (κ) is used to measure the agreement between predicted and true labels while accounting for chance agreement, making it a more reliable metric than accuracy alone under imbalanced conditions.

IV. RESULTS AND DISCUSSION

A. Overall Classification Performance

Throughout the evaluation, CTL-Net consistently outperformed all baseline methods across all performance metrics, demonstrating its effectiveness for corn disease classification compared with 14 state-of-the-art models, as shown in Table II.

TABLE II. COMPREHENSIVE PERFORMANCE COMPARISON ON CORN DISEASE TEST SET

Model	Accuracy (%)	Precision	Recall	F1-score	MCC	AUC	κ	Params (M)	Inference (ms)
ResNet50	95.74	0.957	0.953	0.955	0.952	0.991	0.948	25.6	18
InceptionV3	96.24	0.963	0.959	0.961	0.958	0.992	0.955	23.9	21
Xception	96.78	0.968	0.965	0.966	0.964	0.993	0.961	22.9	23
DenseNet121	97.92	0.979	0.977	0.978	0.977	0.996	0.975	8.0	20
EfficientNetB3	97.35	0.974	0.971	0.972	0.970	0.994	0.968	12.0	25
MobileNetV2	94.58	0.946	0.942	0.944	0.940	0.988	0.937	3.5	12
ViT-B/16	97.68	0.977	0.975	0.976	0.974	0.995	0.972	86.6	48
ResNet50 + DenseNet121	97.46	0.975	0.972	0.973	0.971	0.994	0.969	33.6	38
Inception-ResNet-v2 + MobileNetV3	97.89	0.979	0.977	0.978	0.976	0.995	0.974	59.7	43
CTL-Net (proposed)	99.42	0.994	0.994	0.994	0.994	0.999	0.993	81.24	42

B. Confusion Matrix Analysis

The confusion matrix in Figure 2 illustrates the classification performance of CTL-Net on the test set and provides insight into its generalization capability for different disease categories.

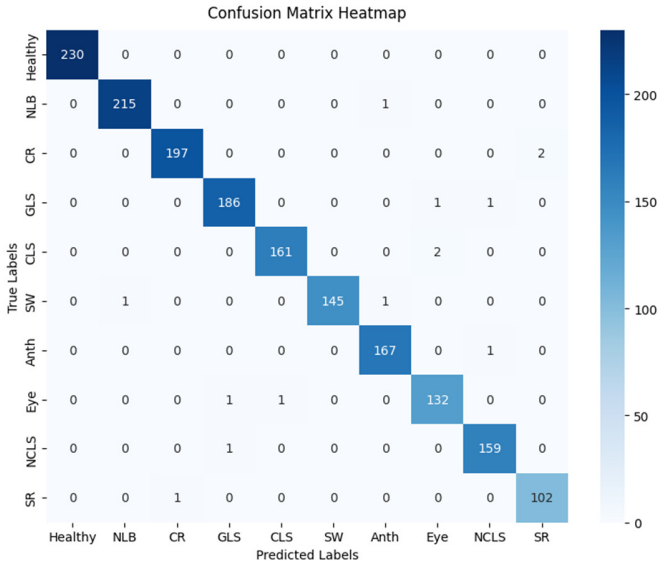

Fig. 2. Confusion matrix for CTL-Net.

The results indicate that CTL-Net correctly classified 1,868 images, with only 11 misclassifications, corresponding to an error rate of approximately 0.58%. In addition, no false negatives were observed in the Healthy class, indicating strong reliability in identifying non-diseased samples

Most misclassifications occurred between disease classes with highly similar lesion patterns, such as Gray Leaf Spot (GLS) and Eyespot, as well as Common Rust (CR) and Southern Rust (SR). A small number of errors, such as Northern Leaf Blight (NLB) classified as Anthracnose and Stewart's Wilt (SW) classified as Anthracnose, can be attributed to overlapping symptom characteristics in borderline cases rather than model limitations. The lack of a consistent misclassification pattern shows that CTL-Net generalizes well by learning diverse and distinct features for disease analysis.

C. Training Dynamics and Convergence Analysis

The training and validation curves for CTL-Net, compared with a selected baseline model, are shown in Figure 3.

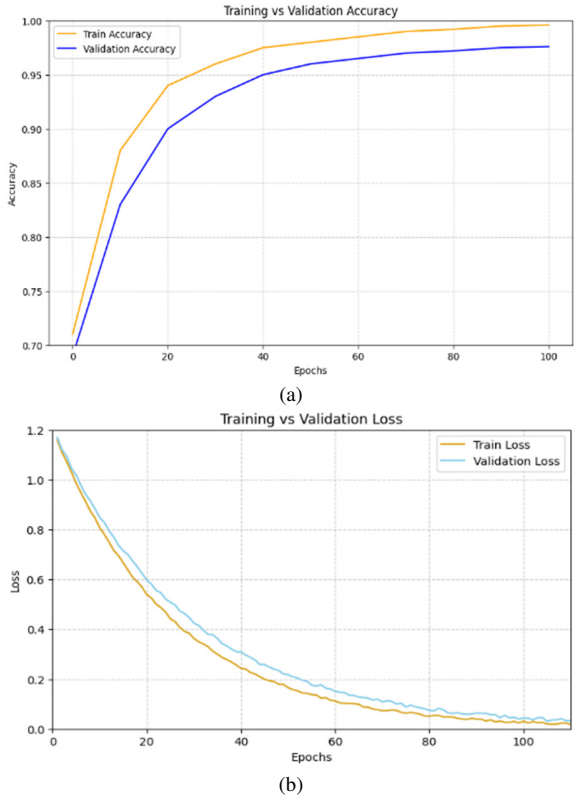

Fig. 3. Comparison of training and validation curves for CTL-Net and a baseline model: (a) accuracy, (b) loss.

To evaluate generalization performance and training stability, 5-fold stratified cross-validation was conducted for

CTL-Net as well as the baseline models on the full dataset. This evaluation was used to assess model robustness and consistency across different data splits.

D. Qualitative Visualization

The Gradient-weighted Class Activation Mapping++ (Grad-CAM++) illustrates the activated regions for different corn disease classes, demonstrating improved lesion localization and class discrimination achieved by CTL-Net compared with the Inception-ResNet-v2 baseline model.

The Grad-CAM++ analysis indicates that the mean Intersection over Union (IoU) between CTL-Net-generated attention maps and expert-annotated disease regions is 0.876, which surpasses the best-performing baseline by approximately 9–14% across disease categories. The model maintains attention consistency under variations in illumination and noise, and it consistently focuses on diagnostically relevant lesion regions. Representative qualitative results are shown in Figure 4.

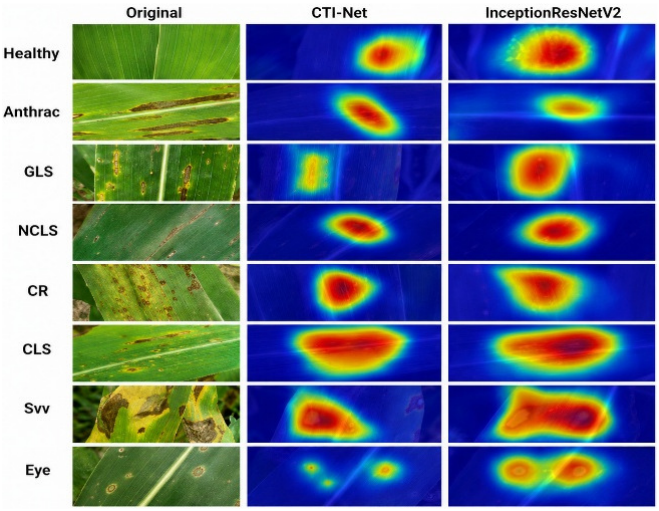

Fig. 4. Grad-CAM++ visualization results.

V. CONCLUSION

This study introduced Corn Transfer Learning Network (CTL-Net), a novel dual-stream hybrid deep learning framework for accurate, interpretable, and practical automated corn disease detection. The proposed architecture integrates Inception-ResNet-v2 and MobileNetV3 through a bi-directional cross-attention mechanism, together with adaptive feature fusion and multi-scale hierarchical processing, enabling the model to effectively capture both global contextual information and fine-grained disease characteristics. Experimental results across ten disease classes demonstrate that CTL-Net achieves an overall classification accuracy of 99.42%, outperforming the strongest individual architecture by 0.66%, the simple hybrid baseline by 1.53%, and lightweight models by 3.95%, indicating its superior performance for automated plant disease diagnosis.

The main contribution of this work lies in the proposed feature interaction mechanism, which moves beyond traditional hybrid architectures based on direct concatenation or static fusion. By employing bi-directional cross-attention and adaptive feature fusion, CTL-Net dynamically enhances informative disease features while suppressing redundant information, leading to more robust feature representations. Furthermore, the incorporation of multi-scale hierarchical processing improves the model's ability to recognize disease symptoms with varying lesion sizes and visual complexities. In addition, the integration of Gradient-weighted Class Activation Mapping++ (Grad-CAM++) provides interpretable and biologically meaningful predictions, achieving an 87.6% Intersection over Union (IoU) with expert-annotated lesion regions and increasing confidence in the model's diagnostic decisions.

The proposed framework demonstrates strong generalization capability, validated through McNemar's statistical test ($p < 0.001$) and five-fold cross-validation (99.42 ± 0.036%). Moreover, CTL-Net achieves 99.42% test accuracy with an inference time of 42 ms on an NVIDIA RTX 3090 GPU, highlighting its computational efficiency and suitability for deployment in resource-constrained agricultural environments. Future work will focus on integrating multimodal data, temporal disease modeling, federated learning, multi-disease diagnosis, and precision agriculture systems to further enhance sustainable and intelligent crop disease management.

DECLARATION OF COPMETING INTERESTS

Not applicable to this work.

ACKNOWLEDGMENT

Not applicable to this work.

DATA AVAILABILITY

The dataset used in this study is publicly available and can be accessed at [30].

REFERENCES

[1] O. Kerdjidj *et al.*, "Uncovering the Potential of Indoor Localization: Role of Deep and Transfer Learning," *IEEE Access*, vol. 12, pp. 73980–74010, 2024, https://doi.org/10.1109/ACCESS.2024.3402997.

[2] P. B. Padol and A. A. Yadav, "SVM classifier based grape leaf disease detection," in *2016 Conference on Advances in Signal Processing*, Pune, India, 2016, pp. 175–179, https://doi.org/10.1109/CASP.2016.7746160.

[3] L. Chen, J. Chen, H. Hajimirsadeghi, and G. Mori, "Adapting Grad-CAM for Embedding Networks," in *2020 IEEE Winter Conference on Applications of Computer Vision*, Snowmass, CO, USA, 2020, pp. 2783–2792, https://doi.org/10.1109/WACV45572.2020.9093461.

[4] P.-T. Jiang, C.-B. Zhang, Q. Hou, M.-M. Cheng, and Y. Wei, "LayerCAM: Exploring Hierarchical Class Activation Maps for Localization," *IEEE Transactions on Image Processing*, vol. 30, pp. 5875–5888, 2021, https://doi.org/10.1109/TIP.2021.3089943.

[5] Y. Lu, S. Yi, N. Zeng, Y. Liu, and Y. Zhang, "Identification of rice diseases using deep convolutional neural networks," *Neurocomputing*, vol. 267, pp. 378–384, Dec. 2017, https://doi.org/10.1016/j.neucom.2017.06.023.

[6] G. Yang, Y. He, Y. Yang, and B. Xu, "Fine-Grained Image Classification for Crop Disease Based on Attention Mechanism," *Frontiers in Plant Science*, vol. 11, Dec. 2020, Art. no. 600854, https://doi.org/10.3389/fpls.2020.600854.

[7] "Leaf Disease Classification Using Artificial Neural Network" *Jurnal Teknologi*, vol. 77, no. 17, pp. 109–114, Nov. 2015, https://doi.org/10.11113/jt.v77.6463.

[8] A. Chattopadhay, A. Sarkar, P. Howlader, and V. N. Balasubramanian, "Grad-CAM++: Generalized Gradient-Based Visual Explanations for Deep Convolutional Networks," in *2018 IEEE Winter Conference on Applications of Computer Vision*, Lake Tahoe, NV, USA, 2018, pp. 839–847, https://doi.org/10.1109/WACV.2018.00097.

[9] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using Deep Learning for Image-Based Plant Disease Detection," *Frontiers in Plant Science*, vol. 7, Sept. 2016, Art. no. 1419, https://doi.org/10.3389/fpls.2016.01419.

[10] Y. Dandawate and R. Kokare, "An automated approach for classification of plant diseases towards development of futuristic Decision Support System in Indian perspective," in *2015 International Conference on Advances in Computing, Communications and Informatics*, Kochi, India, 2015, pp. 794–799, https://doi.org/10.1109/ICACCI.2015.7275707.

[11] V. Singh, Varsha, and A. K. Misra, "Detection of unhealthy region of plant leaves using image processing and genetic algorithm," in *2015 International Conference on Advances in Computer Engineering and Applications*, Ghaziabad, India, 2015, pp. 1028–1032, https://doi.org/10.1109/ICACEA.2015.7164858.

[12] J. K. Patil and R. Kumar, "Color Feature Extraction of Tomato Leaf Diseases," *International Journal of Engineering Trends and Technology*, vol. 2, no. 2, pp. 72–74, 2011, https://doi.org/10.14445/22315381/IJETT-V2I2P214.

[13] R. Ghadge, J. Kulkarni, P. More, S. Nene, J. Kulkarni, and P. R L, "Prediction of Crop Yield using Machine Learning," *International Research Journal of Engineering and Technology*, vol. 5, no. 2, pp. 2237–2239, Feb. 2018.

[14] Z. Hong, Z. Kalbarczyk, and R. K. Iyer, "A Data-Driven Approach to Soil Moisture Collection and Prediction," in *2016 IEEE International Conference on Smart Computing*, St. Louis, MO, USA, 2016, pp. 1–6, https://doi.org/10.1109/SMARTCOMP.2016.7501673.

[15] V. P. Bhartiya, R. R. Janghel, and Y. K. Rathore, "Rice Leaf Disease Prediction Using Machine Learning," in *2022 Second International Conference on Power, Control and Computing Technologies*, Raipur, India, 2022, pp. 1–5, https://doi.org/10.1109/ICPC2T53885.2022.9776692.

[16] K. P. Panigrahi, H. Das, A. K. Sahoo, and S. C. Moharana, "Maize Leaf Disease Detection and Classification Using Machine Learning Algorithms," in *2nd International Conference on Computing Analytics and Networking*, Bhubaneswar, Odisha, India, 2019, pp. 659–669, https://doi.org/10.1007/978-981-15-2414-1_66.

[17] Z. Liu, R. N. Bashir, S. Iqbal, M. M. A. Shahid, M. Tausif, and Q. Umer, "Internet of Things (IoT) and Machine Learning Model of Plant Disease Prediction–Blister Blight for Tea Plant," *IEEE Access*, vol. 10, pp. 44934–44944, 2022, https://doi.org/10.1109/ACCESS.2022.3169147.

[18] D. Dhabliya, "An integrated optimization model for plant diseases prediction with machine learning model," *Machine Learning Applications in Engineering Education and Management*, vol. 1, no. 2, pp. 21–26.

[19] M. Ahmad, M. Abdullah, H. Moon, and D. Han, "Plant Disease Detection in Imbalanced Datasets Using Efficient Convolutional Neural Networks With Stepwise Transfer Learning," *IEEE Access*, vol. 9, pp. 140565–140580, 2021, https://doi.org/10.1109/ACCESS.2021.3119655.

[20] R. Kaundal, A. S. Kapoor, and G. P. Raghava, "Machine learning techniques in disease forecasting: a case study on rice blast prediction," *BMC Bioinformatics*, vol. 7, no. 1, Nov. 2006, Art. no. 485, https://doi.org/10.1186/1471-2105-7-485.

[21] S. Nandhini and K. Ashokkumar, "Machine Learning Technique for Crop Disease Prediction Through Crop Leaf Image," *Applied Mathematics & Information Sciences*, vol. 11, no. 2, pp. 149–158, Mar. 2022, https://doi.org/10.18576/amis/160202.

[22] T. Vijaykanth Reddy and K. Sashi Rekha, "Deep Leaf Disease Prediction Framework (DLDPF) with Transfer Learning for Automatic Leaf Disease Detection," in *2021 5th International Conference on Computing Methodologies and Communication*, Erode, India, 2021, pp. 1408–1415, https://doi.org/10.1109/ICCMC51019.2021.9418245.

[23] M. Bhagat and D. Kumar, "Performance enhancement of kernelized SVM with deep learning features for tea leaf disease prediction," *Multimedia Tools and Applications*, vol. 83, no. 13, pp. 39117–39134, Apr. 2024, https://doi.org/10.1007/s11042-023-17172-1.

[24] N. Ganatra and A. Patel, "A Multiclass Plant Leaf Disease Detection using Image Processing and Machine Learning Techniques," *International Journal on Emerging Technologies*, vol. 11, no. 2, pp. 1082–1086, Jan. 2020.

[25] S. Madhu and V. RaviSankar, "Comprehensive Analysis of a YOLO-based Deep Learning Model for Cotton Plant Leaf Disease Detection," *Engineering, Technology & Applied Science Research*, vol. 15, no. 1, pp. 19947–19952, Feb. 2025, https://doi.org/10.48084/etasr.8944.

[26] K. Sundaramoorthi and M. Kamarasan, "Integration of Deep Learning with Fox Optimization Algorithm for Early Detection and Classification of Tomato Leaf and Fruit Diseases," *Engineering, Technology & Applied Science Research*, vol. 15, no. 1, pp. 19343–19348, Feb. 2025, https://doi.org/10.48084/etasr.9216.

[27] "PlantVillage Dataset." Kaggle. [Online]. Available: https://www.kaggle.com/datasets/emmarex/plantdisease.

[28] D. P. Hughes and M. Salathe, "An open access repository of images on plant health to enable the development of mobile disease diagnostics." arXiv, Nov. 25, 2015, https://doi.org/10.48550/arXiv.1511.08060.

[29] "Corn or Maize Leaf Disease Dataset." Kaggle. [Online]. Available: https://www.kaggle.com/datasets/smaranjitghose/corn-or-maize-leaf-disease-dataset.

[30] vishruthkp, "vishruthkp/maizedataset." June 10, 2026. [Online]. Available: https://github.com/vishruthkp/maizedataset.