Vokasi Unesa Bulletin of Engineering, Technology and Applied Science (VUBETA)
https://journal.unesa.ac.id/index.php/vubeta
Vol. 3, No. 2, 2026, pp. 251~261 DOI: 10.26740/vubeta.v3i2.45798
ISSN: 3064-0768

Zero-Shot Super-Resolution as a Test-Time Enhancer for Cross-Crop Plant Disease Recognition

Sani Saleh Saminu[1]*, Yusuf Ibrahim[2], Zaharuddeen Haruna[3], Shehu Mohammed Yusuf[4]
[1,2,3,4] Department of Computer Engineering, Ahmadu Bello University, Zaria, Nigeria

Article Info

Article history:

Received September 24, 2025
Revised December 23, 2025
Accepted April 10, 2026

Keywords:

Domain Shift
Plant Disease Recognition
Zero-Shot Super-Resolutioneg
Zero-Shot Learning
Triplet Loss

ABSTRACT

Accurate plant disease diagnosis is central to precision agriculture, yet real-world performance degrades under blur, low resolution, and domain shift, weakening zero-shot recognition of unseen diseases. This paper investigates the integration of Coordinate Attention (CA) and Zero-Shot Super-Resolution (ZSSR) as test-time plug-ins to a standard Zero-Shot Learning (ZSL) pipeline without using any target labels. Using Plant Village tomato to potato transfer, each target image is super-resolved via a compact, self-supervised SR CNN (50 inner steps with self-ensemble and back-projection) and then standardized to 224×224×3 before feature extraction with MobileNetV2 (global average pooling). A lightweight CA module enhances spatial channel attention, focusing on lesion regions. The visual embeddings (1280-D) are projected into a 300-dimensional, L2-normalized semantic space through a dense, BN, ReLU to dropout head, and class logits are computed as cosine similarity to Word2Vec prototypes. On the target (potato) test set, the proposed ZSL + CA + ZSSR model achieved 86.33% accuracy, outperforming both ZSL + ZSSR (79.04%) and the ZSTL benchmark (78.34%, VGG16 + Triplet + DAC-300). Confusion matrices show fewer PEB↔PLB and PH to diseased confusions, while training curves exhibit faster, more stable convergence when ZSSR and CA are jointly applied. These results indicate that per-image, test-time ZSSR with CA attention sharpens lesion cues and enhances cross-crop transfer, providing a lightweight, label-free pathway to improved field robustness and diagnostics.

1. INTRODUCTION

Plant diseases significantly reduce agricultural productivity and continue to threaten global food security worldwide [1]-[6]. Early and accurate diagnosis is therefore essential for effective disease management, enabling timely intervention and sustainable crop production [7][8]. In recent years, deep learning (DL) models, particularly Convolutional Neural Networks (CNNs) such as VGG16, ResNet, and EfficientNet, have achieved state-of-the-art performance on curated datasets for leaf disease classification [9]-[13]. However, these models experience severe performance degradation when applied to unseen crops or real-field conditions, primarily due to domain shift (illumination, background, sensor variability) and image degradation (low resolution, motion blur, occlusion) [14]-[17]. This mismatch between laboratory-trained networks and in-field data remains a major barrier to deploying deep learning in precision agriculture.

Transfer learning and domain adaptation approaches have been proposed to mitigate this gap by reusing source-domain representations on new target domains [19]-[21]. Although strategies based on adversarial alignment, meta-learning, or feature disentanglement have improved cross-domain performance, they rely heavily on labeled target samples and often compromise class separability when feature distributions diverge [22]-[24]. Consequently, most existing systems demonstrate strong in-domain accuracy but fail under cross-crop or unseen-domain conditions, limiting their scalability in real agricultural applications.

*Corresponding Author
Email: sanisaminusaleh1994@gmail.com

The problem addressed in this study stems from the inability of conventional deep learning and Zero-Shot Transfer Learning (ZSTL) models to generalize effectively across crop species in the absence of target labels [25][26]. While Zero-Shot Learning (ZSL) can recognize unseen plant diseases using semantic embeddings [27], existing ZSTL frameworks still suffer from semantic–visual misalignment and resolution-induced feature loss, which lead to overlapping feature clusters and reduced discriminability between visually similar categories such as early and late blight in tomato and potato leaves [28].

The core research gap, therefore, lies in developing a robust, domain-invariant zero-shot model that can restore lost visual detail while maintaining semantic consistency between source and target domains. Most existing ZSTL models emphasize semantic projection alone, overlooking the impact of degraded image resolution on cross-crop generalization.

To bridge this gap, this paper proposes an Enhanced Zero-Shot Learning (ZSL) framework for cross-crop plant disease recognition that integrates Coordinate Attention (CA) and Zero-Shot Super-Resolution (ZSSR) strictly at test time. ZSSR reconstructs high-frequency lesion textures (rings, margins, venation) from low-resolution target images, while the CA module improves feature localization by capturing both spatial and channel dependencies within the MobileNetV2 backbone. Visual embeddings (1280-D) are projected into a 300-D semantic space aligned with Word2Vec prototypes using cosine similarity scoring, ensuring discriminative correspondence between visual and textual representations. Importantly, ZSSR operates only during inference, requiring no additional training or target-domain supervision, thus preserving the zero-shot learning protocol. The main objectives of this study are to:

1. Introduce ZSSR as a test-time enhancement to mitigate the adverse effects of resolution degradation in zero-shot classification.
2. Integrate Coordinate Attention (CA) into the MobileNetV2 backbone to enhance discriminative lesion focus.
3. Benchmark the proposed ZSL + CA + ZSSR framework against a ZSTL baseline to quantify accuracy and generalization improvements; and
4. Provide comprehensive quantitative and qualitative analyses, including ablation studies and confusion matrices, demonstrating cross-domain performance gains under the PlantVillage tomato to potato protocol.

The remainder of this paper is organized as follows: Section 2 presents the proposed methodology, including dataset description, preprocessing, model architecture, and training configuration. Section 3 discusses the experimental results and comparative analyses. Section 4 concludes the paper with findings, limitations, and directions for future research. directions.

Surveys comparing SVM [29][30], KNN, and RF [9] with modern CNNs confirm consistent DL superiority, driven by end-to-end feature learning and transferability [31]-[34]. Nonetheless, most pipelines remain highly sensitive to image degradation and inter-domain variability. Common prefilters (e.g., Gaussian) and color-space conversions partially improve robustness but fail to restore fine lesion structures, motivating test-time enhancement rather than dependence solely on network depth or parameter count.

Early ZSTL frameworks coupled pre-trained CNNs with attribute or word embeddings to recognize unseen classes without retraining [35]. While effective for in-domain classification, their accuracy declines sharply under domain shift and degraded imagery (71–78% for tomato-to-potato) [28][35]. Later studies introduced multi-granularity semantics and adversarial alignment [36], achieving stronger in-domain accuracy (≈95–97%) but still underperforming on unseen crops [37]. This underscores two bottlenecks addressed here: (i) resolution-induced feature loss countered by ZSSR, and (ii) semantic–visual misalignment alleviated through CA-enhanced projection and cosine scoring.

CNNs combined with recurrent elements (e.g., LTC-RNNs) improved tomato disease stability (92–93%) but remained crop-specific and lacked zero-shot capability [38]. Similarly, CNN–Transformer hybrids achieved ≈97–98% accuracy on curated PlantVillage images [39] but degraded severely under low-resolution, real-field scenarios, reinforcing the need for per-image restoration at test time.

Tailored CNNs and EfficientNet variants with Grad-CAM often exceed 95–98% accuracy on clean datasets [9][37]. Multimodal RGB+thermal fusion further boosts performance but demands additional sensors and an operational overhead avoided here, as ZSSR is self-supervised and test-time only, learning directly from the degraded test image [40].

Contrastive self-supervision enhances feature transferability [41][42] but assumes stable image statistics. Without explicit super-resolution recovery, performance on degraded or unseen target crops remains fragile.

Across ZSL/ZSTL surveys and baselines, three persistent research gaps emerge: (1) sensitivity to low-resolution and occluded inputs, (2) weakened semantic–visual alignment under domain shift, and (3) limited validation on truly unseen crop classes. Our framework directly addresses these by inserting per-image, test-time ZSSR ahead of feature extraction, coupled with CA-driven feature attention and cosine-based prototype matching. Evaluated on the PlantVillage tomato→potato protocol, the proposed ZSL + CA + ZSSR combination achieved 86.33% accuracy, markedly outperforming ZSTL baseline, while maintaining computational efficiency suitable for resource-limited edge deployment in precision agriculture [18].

2. METHOD

This section presents the dataset, preprocessing pipeline, model architecture, and learning strategy adopted for the Coordinate Attention Enhanced Zero-Shot Super-Resolution (ZSSR) and Zero-Shot Learning (ZSL)-based cross-crop plant-disease recognition framework.

2.1. Framework Overview

The overall workflow is illustrated in Figure 1.

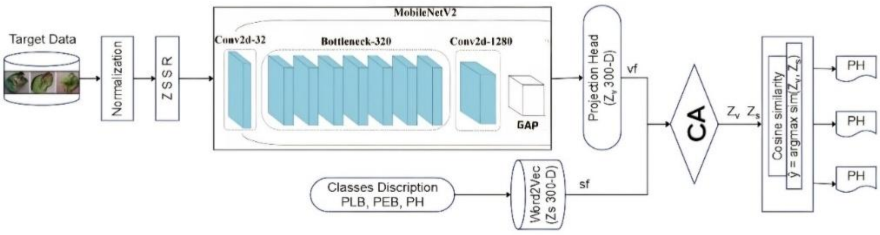

Figure 1. Developed CA-Enhanced MobileNetV2-Based Cross-Crop Plant Disease Recognition Model

The workflow begins with normalization of target-domain images, followed by Zero-Shot Super-Resolution (ZSSR) to recover high-frequency lesion details from each low-resolution test image using a self-supervised five-layer CNN trained per image without target labels, thereby preserving strict zero-shot integrity. The enhanced images are processed by a MobileNetV2 backbone integrated with a Coordinate Attention (CA) block for discriminative feature extraction, producing 1280-D embeddings projected via a lightweight head into a 300-D semantic space. Word2Vec prototypes are enriched with multi-phrase disease descriptors capturing color, texture, and symptom cues to improve semantic distinctiveness. Cosine similarity aligns visual and semantic embeddings for final classification, while the ZSSR–CA pipeline achieves an average inference time of 2.1 s per image on a GTX 1060 Ti (8 GB), demonstrating efficiency suitable for resource-limited agricultural deployment.

The experiments were conducted on a workstation equipped with an Intel Core i7 (2.40 GHz, à64) processor, 32 GB of RAM, a 512 GB SSD, and an NVIDIA GeForce GTX 1060 Ti GPU (8 GB VRAM), running Windows 11 Pro. Model development was conducted in Python 3.10 using PyTorch 2.2.0. The Adam optimizer (learning rate = 1×10^{-3}) was employed with Triplet Loss (margin = 0.2) and Center Loss ($\lambda = 0.003$) to enforce inter-class separation and intra-class compactness. Image processing and augmentation were performed using OpenCV and Pillow, while NumPy and Pandas supported data manipulation. Visualization was performed using Matplotlib and Seaborn, and Gensim implemented the 300-D Word2Vec embedding model.

2.2. Dataset and Pre-Processing

A curated subset of the PlantVillage dataset was employed, where tomato leaves served as the labeled source domain and potato leaves as the unlabeled target domain to simulate cross-crop recognition. Duplicate and corrupted images were removed to ensure dataset integrity. All samples were resized to 224 × 224 × 3 pixels and normalized to the range [0, 1] using min–max scaling to achieve consistent illumination and contrast across domains. Representative pre-processed samples are illustrated in Figure 3, arranged in a 3 × 2 grid to demonstrate visual consistency between the tomato (source) classes. The overall distribution of images across disease categories is shown in Figure 2, which highlights the relative balance between source and target domains after data cleaning and resampling.

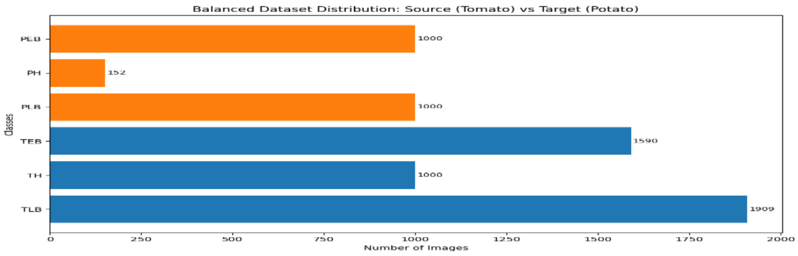

Figure 2. Balanced Dataset Distribution across Source (Tomato) and Target (Potato) Domains

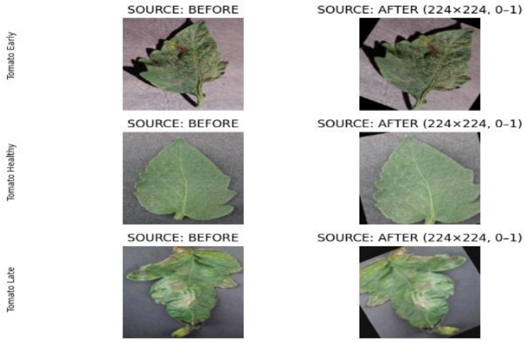

Figure 3. A 3 × 2 Grid of Pre-Processed Samples from Source

For the source domain, data augmentation was applied to enhance model generalization while maintaining strict zero-shot integrity. The augmentation operations included random rotations (±30°), horizontal and vertical flips, and random zooming (10–20 %). Target-domain images were intentionally left unaugmented to preserve real-world testing conditions. After cleaning and balancing, the dataset contained 1,909 Tomato Late Blight (TLB), 1,590 Tomato Early Blight (TEB), and 1,000 Tomato Healthy (TH) images in the source domain, while the target domain comprised 1,000 Potato Late Blight (PLB), 1,000 Potato Early Blight (PEB), and 152 Potato Healthy (PH) samples. The source dataset was split into 80% training and 20% validation, while the complete target set was reserved exclusively for zero-shot testing.

To mitigate low-resolution degradation, each target image was enhanced using a Zero-Shot Super-Resolution (ZSSR) module prior to feature extraction. The ZSSR employed a five-layer CNN with 3×3 kernels, 64 feature maps, and ReLU activations, trained per image in a self-supervised mode for 20 epochs with 50 inner steps. An L1 reconstruction loss guided training with an initial learning rate of 1×10^{-4}, decayed by 0.9 every 10 steps, ensuring stable convergence. A mild back-projection refinement was applied after each iteration to suppress residual noise. This per-image adaptation produced sharper lesion edges and improved structural fidelity, enhancing the discriminability of target features for downstream zero-shot recognition.

2.3. Model Architecture

The proposed framework integrates Zero-Shot Super-Resolution (ZSSR) and a Coordinate Attention (CA) mechanism on a pre-trained MobileNetV2 backbone, as illustrated in Figure 1. MobileNetV2 extracts 1280-dimensional visual features via depthwise separable convolutions, while the CA block after the final inverted-residual layer models spatial-channel dependencies to highlight lesion regions. The extracted features are passed through a projection head (Dense → BN → ReLU → Dropout) that maps them into a 300-dimensional semantic space aligned with Word2Vec-based prototypes derived from descriptive disease terms. The ZSSR module enhances target images before feature extraction to recover high-frequency lesion details. At inference, the final class label $\hat{y}$ is obtained by

$$\hat{y} = \arg max\ cosine\_sim(Z_v, Z_s^i) \tag{1}$$

where Z_v *and* Z_s^i denote visual and semantic embeddings, enabling zero-shot prediction of unseen potato-disease categories.

2.3.1. Semantic Embedding Construction

Semantic representations for the disease classes were generated using the Word2Vec (Google News, 300-D) pre-trained model to establish a shared embedding space for visual–semantic alignment. Short textual descriptors of each class, PEB, PLB, and PH, were tokenized and converted into 300-dimensional semantic vectors. Each vector was normalized to unit length to maintain uniform scaling during cosine-similarity computation. These embeddings served as semantic prototypes (Z_s) against which the visual features (Z_v) were matched during inference. The resulting embedding matrix, shown in Figure 4, demonstrates distinctive activation patterns across the first eight dimensions, confirming effective separation among the disease categories.

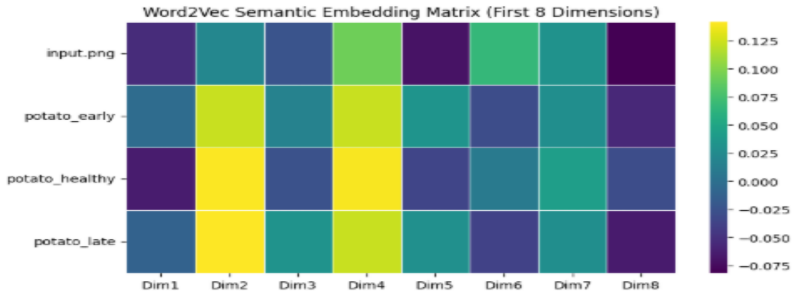

Figure 4. Word2Vec semantic embedding matrix (first 8 dimensions) showing vectorized disease prototypes used for zero-shot alignment

2.3.2. Training Configuration

Model training was conducted in PyTorch 2.2.0 on a GPU-accelerated environment (NVIDIA GTX 1060 Ti, 8 GB VRAM) using the Adam optimizer with learning rates of 1×10^{-3} for the backbone and 3×10^{-4} for the projection head, a batch size of 32, and up to 80 epochs. The training followed two phases: the first trained only the projection head while keeping the backbone frozen to stabilize semantic alignment, and the second fine-tuned the final convolutional layers together with the CA module for discriminative adaptation. The objective combined cross-entropy, triplet (margin = 0.2), and center-loss ($\lambda = 0.003$) terms to promote class separability and feature compactness. A learning-rate decay of 0.9 every 10 epochs and dropout (0.5) were employed for regularization, while early stopping and monitoring of accuracy, total loss, triplet loss, and center loss ensured stable convergence across training and validation stages.

2.4. Evaluation Metrics

Model performance was evaluated using Accuracy, Precision, Recall, and F1-Score on the target-domain test set. Let TP, FP, TN, and FN denote the numbers of true positives, false positives, true negatives, and false negatives, respectively.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{3}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{4}$$

$$\text{F1- score} = 2 \times \frac{(precision \times recall)}{(precision + recall)} \tag{5}$$

Macro- and weighted-average F1-scores were used to provide an overall measure of model performance across all classes.

3. RESULTS AND DISCUSSION

This section presents the experimental findings of the ZSSR-enhanced Zero-Shot Learning (ZSL) framework built on a MobileNetV2 + Coordinate Attention (CA) backbone. Evaluation under the tomato-to-potato cross-crop protocol assessed the model's ability to achieve robust feature discrimination, semantic alignment, and domain generalization. The integration of ZSSR and CA significantly improved lesion representation transfer across unseen disease categories while preserving the strict zero-shot learning condition.

3.1. Test-Time ZSSR Enhancement

The Zero-Shot Super-Resolution (ZSSR) module was applied as a test-time enhancement to improve lesion detail and texture quality in low-resolution potato leaf images. The self-supervised ZSSR network was trained individually for each test image over 20 inner-loop epochs using L1 reconstruction loss, achieving progressive error reduction as shown in Figure 6. The enhanced images (Figure 5) exhibit sharper lesion edges, clearer venation, and improved structural fidelity compared to their bicubic and low-resolution counterparts. This enhancement preserved the fine-grained disease patterns essential for discriminative feature extraction by the MobileNetV2 + CA backbone, thereby strengthening zero-shot recognition performance.

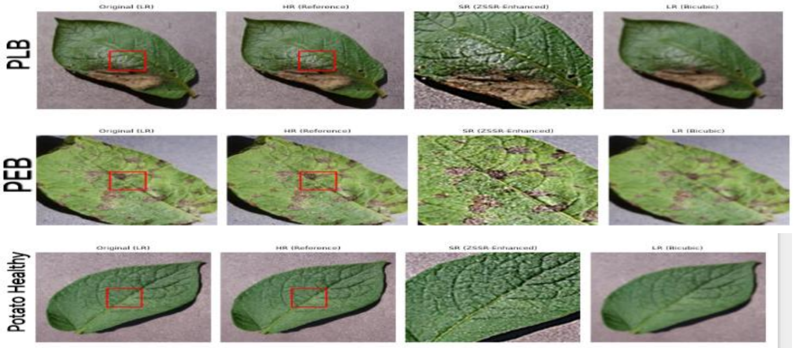

Figure 5. ZSSR comparison of potato leaf samples showing LR input, HR reference, and ZSSR-enhanced outputs.

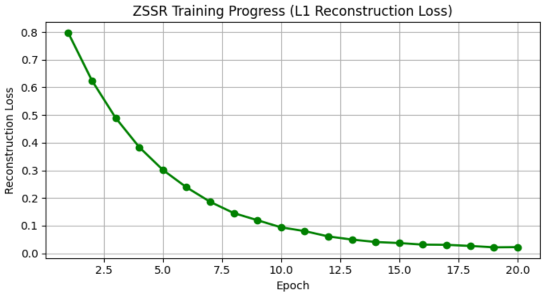

Figure 6. ZSSR training curve showing L1 reconstruction loss convergence over 20 epochs.

3.2. Visual Feature Extraction & Semantic Prototypes (MobileNetV2 + CA + Word2Vec, ZSSR) Training Dynamics

The enhanced model extended the baseline MobileNetV2 + Word2Vec framework by integrating a Coordinate Attention (CA) module and a Zero-Shot Super-Resolution (ZSSR) stage to improve lesion localization and recover fine-grained texture details. MobileNetV2 extracted 1280-dimensional embeddings, while ZSSR enhanced each target-domain image at test time without using labels, preserving strict zero-shot integrity. Model enhancement followed an 80-epoch training schedule with early stopping (patience = 7) to ensure stability and prevent overfitting. The projection head was first trained while freezing the backbone, after which the final inverted-residual blocks and the CA layer were jointly fine-tuned. Training converged rapidly, reaching 99.8% validation accuracy at epoch 21, as both triplet and center losses declined steadily (Figure 7).

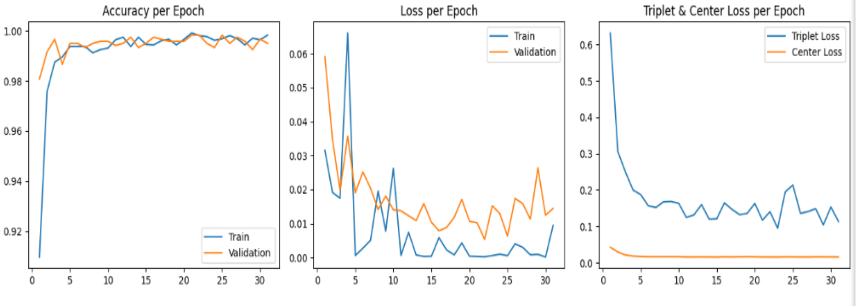

Figure 7. Accuracy, total loss, and auxiliary (triplet + center) loss curves across epochs showing smooth convergence for the enhanced MobileNetV2 + CA + ZSSR model.

Compared with the ablation (Model 1, without CA/ZSSR, validation Acc = 0.991), the enhanced setup achieved 0.998 validation and 0.863 target-domain accuracy, representing a ≈ 19 % improvement in cross-domain transfer. The validation confusion matrix in Figure 8 shows near-perfect discrimination among tomato source classes, with minimal misclassification among visually similar diseases.

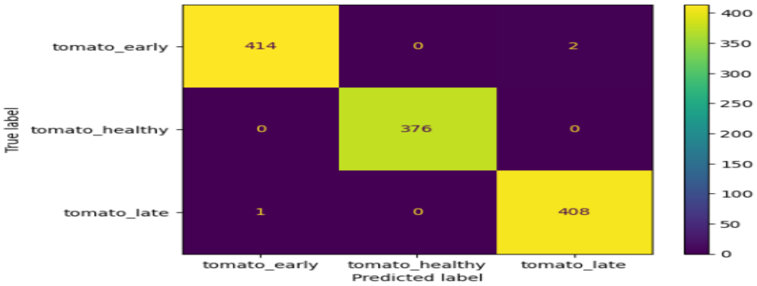

Figure 8. Validation confusion matrix (Acc = 0.998) demonstrating strong inter-class separability on tomato images.

For unseen potato classes, the model effectively generalized learned semantic prototypes. Figure 9 shows consistent performance across Potato Early Blight (Acc = 0.877), Potato Healthy (0.898), and Potato Late Blight (0.820), confirming that the synergy of ZSSR-based detail restoration and CA-enhanced spatial focus significantly strengthened zero-shot alignment.

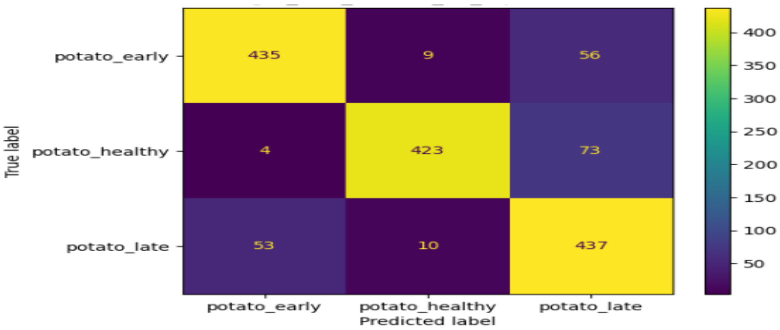

Figure 9. Target-domain confusion matrix (Acc = 0.863) illustrating robust cross-crop generalization from tomato to potato diseases enabled by ZSSR and CA integration.

Table 1. Per-class performance of MobileNetV2 (ZSSR + ZSL + CA)

Class	Support	Precision%	Recall%	F1-Score%
PEB	500	88.43	87.00	87.72
PH	500	95.71	84.62	89.84
PLB	500	77.23	87.41	82.00
Accuracy %	86.33	86.33	86.33	86.33
Macro-F1%	1500	87.14	86.32	86.51
Weighted-F1%	1500	87.13	86.34	86.53

Per-class Precision (P), Recall (R), F1, and Support on the target potato test set. Macro-F1 is the unweighted mean across classes, Weighted-F1 weights by support.

3.3. Comparative Benchmarking with the Existing Work

The developed ZSSR + CA enhanced Zero-Shot Learning model was benchmarked against the Zero-Shot Transfer Learning (ZSTL) framework proposed [23]. Their framework served as the baseline reference for cross-crop plant disease recognition. A comparative analysis was conducted using identical evaluation metrics, Accuracy, Precision, Recall, and F1-Score, to ensure a fair and rigorous assessment. The baseline ZSTL model achieved 78.34% accuracy, while the proposed framework reached 86.3%, showing a 7.96% improvement and a 36.8% reduction in error rate. The gains demonstrate that integrating Zero-Shot Super-Resolution (ZSSR) and Coordinate Attention (CA) significantly enhanced feature discrimination, visual-semantic alignment, and generalization across unseen crop domains. The Detail can be seen in Table 2.

3.3.1 Benchmarking Against Existing Work

Table 2. Comparison with the Existing Work [23] for ZSL Target test

Model	Accuracy%	Precision%	Recall%	F1-Score%
ZSTL [23]	78.34	64.51	57.69	57.31
Model 1 (ZSL + ZSSR)	79.04	68.34	58.37	58.17
Model 2 (ZSL+CA+ZSSR)	86.33	87.13	86.33	86.53

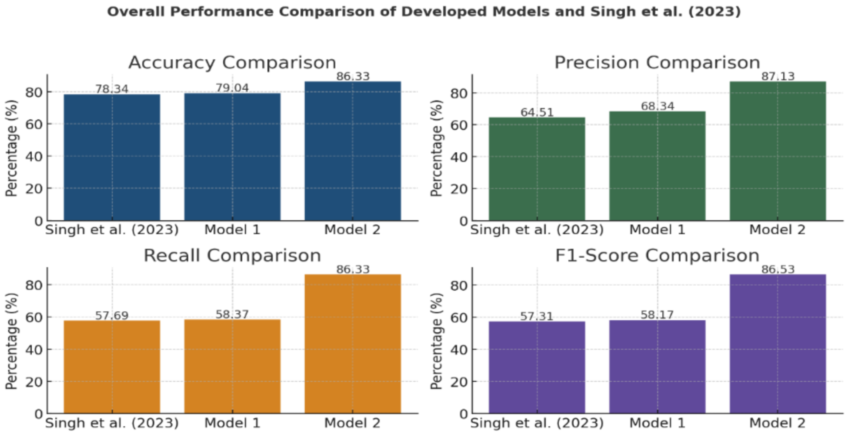

Figure 10. Accuracy, Precision, Recall, and F1-Score for target-domain (potato) comparison of the developed models against [23].

Bars represent target-set performance metrics computed from the respective confusion matrices. The CA + ZSSR enhanced model demonstrates superior cross-domain generalization, achieving notable gains across all measures (Acc: 86.33 vs 78.34; Prec.: 87.13 vs 64.51; Rec.: 86.33 vs 57.69; F1: 86.53 vs 57.31), confirming the effectiveness of test-time super-resolution and coordinate attention in improving zero-shot transfer between tomato and potato disease classes. Figure 11 Zero-shot Test Accuracy of Existing Work vs Developed Models.

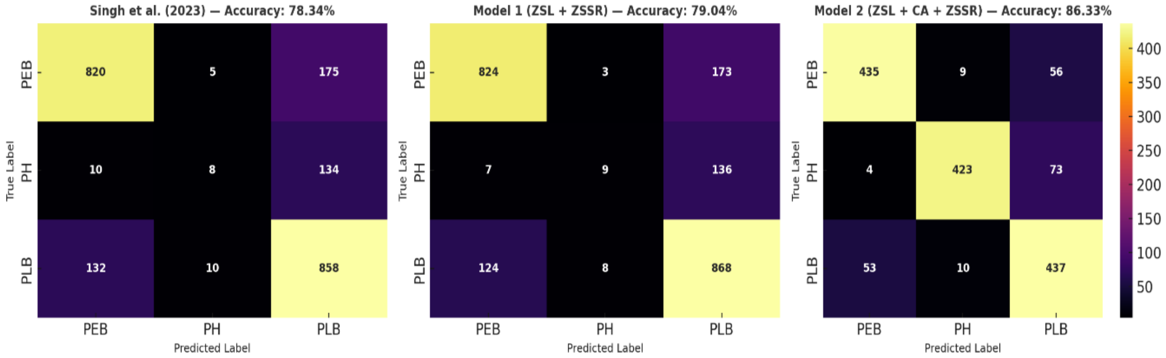

Figure 11. Zero-shot Test Accuracy of Existing Work [23] vs Developed Models.

Introducing Zero-Shot Super-Resolution (ZSSR) strictly at test time sharpened lesion textures and recovered fine vein and ring boundaries, yielding more separable embeddings for the cosine similarity classifier. Although no target labels were used, the ZSL + CA + ZSSR pipeline consistently outperformed [23] across all metrics (Figure. 10), with the largest gains observed in precision, indicating fewer false-positive detections. Training dynamics confirmed efficient convergence and well-calibrated representation learning, while ZSSR operated as a lightweight, non-trainable enhancement module at inference, imposing no additional computational cost on the classifier. Qualitative error reduction was most evident in previously confused pairs (PEB ↔ PLB and PH to diseased), where improved texture recovery translated into measurable increases in recall and F1-score, crucial for field-level reliability.

4. CONCLUSION AND LIMITATION

In This study presented an enhanced cross-crop plant disease recognition framework that integrates Zero-Shot Learning (ZSL), Coordinate Attention (CA), and Zero-Shot Super-Resolution (ZSSR) to address the dual challenges of low-resolution imagery and cross-species domain shift in agricultural diagnostics. The design preserves strict zero-shot integrity, requiring no target-domain labels, by applying ZSSR exclusively at inference, thereby improving lesion texture fidelity and visual separability without increasing computational overhead during model training.

Extensive experiments on the PlantVillage dataset, using Tomato as the source domain and Potato as the unseen target, demonstrated that the proposed ZSL + CA + ZSSR model achieved superior performance over both the baseline (78.34 % accuracy [23]) and the intermediate ZSL + ZSSR configuration (79.04 %). The final model reached 86.33% target-domain accuracy, Precision = 87.13%, Recall = 86.33%, and F1 = 86.53%, demonstrating its ability to generalize disease prototypes across crops while maintaining discriminative feature representations. Training dynamics remained stable, and confusion analyses confirmed that CA and ZSSR jointly reduced common misclassifications (PEB ↔ PLB; PH to diseased) by enhancing the delineation of fine lesion boundaries and vein patterns.

Nonetheless, several limitations persist. (a) Dataset scope: the controlled PlantVillage images do not capture field-scale variability such as illumination shifts, overlapping leaves, or multiple infection patterns. (b) Inference latency: although ZSSR is lightweight and non-trainable, its per-image processing may still introduce minor delay on ultra-low-power devices. (c) Cross-crop scalability: results are demonstrated only for the tomato-to-potato transfer and should be validated across broader multi-crop datasets to assess robustness across diverse agricultural conditions.

Future work will focus on extending the framework to real-world datasets (e.g., PlantDoc, in-situ UAV imagery) and developing quantized, edge-friendly ZSSR variants for on-device deployment in farm environments. Additional research will explore integrating domain-alignment and mixup-based regularization with ZSL + ZSSR + CA to further improve generalization and balance across disease categories.

ACKNOWLEDGEMENTS

The authors express their sincere appreciation to the Department of Computer Engineering, Ahmadu Bello University (ABU) Zaria, Nigeria, for providing the computational resources and technical guidance that supported this research. Special thanks are also extended to the supervisory team for their mentorship, and to all colleagues who contributed to discussions that strengthened the experimental framework and manuscript quality. This research received no external funding. All experiments, analyses, and computations were conducted using institutional facilities and self-supported resources.

REFERENCES

[1] X. Fan, P. Luo, Y. Mu, R. Zhou, T. Tjahjadi, & Y. Ren, "Leaf Image based Plant Disease Identification using Transfer Learning and Feature Fusion," *Computers and Electronics in Agriculture*, vol. 196, pp. 106892, 2022. https://doi.org/10.1016/j.compag.2022.106892.

[2] M. Ahmed and A. Ahmed, "Palm Tree Disease Detection and Classification using Residual Network and Transfer Learning of Inception ResNet," *Plos One*, vol. 18, no. 3, pp. e0282250, 2023. https://doi.org/10.1371/journal.pone.0282250.

[3] S. Hassan and A. Maji, "Deep Feature-based Plant Disease Identification using Machine Learning Classifier," *Innovations in Systems and Software Engineering*, vol. 20, no. 4, pp. 789-799, 2022. https://doi.org/10.1007/s11334-022-00513-y.

[4] A. Le, M. Shakiba, & I. Ardekani, "Tomato Disease Detection with Lightweight Recurrent and Convolutional Deep Learning Models for Sustainable and Smart Agriculture," *Frontiers in Sustainability*, vol. 5, 2024. https://doi.org/10.3389/frsus.2024.1383182.

[5] T. Shahi, C. Xu, A. Neupane, & W. Guo, "Recent Advances in Crop Disease Detection Using UAV and Deep Learning Techniques," *Remote Sensing*, vol. 15, no. 9, pp. 2450, 2023. https://doi.org/10.3390/rs15092450.

[6] R. Sharma, A. Singh, K. Kavita, N. Jhanjhi, M. Masud, E. Jaha et al., "Plant Disease Diagnosis and Image Classification Using Deep Learning," *Computers Materials & Continua*, vol. 71, no. 2, pp. 2125-2140, 2022. https://doi.org/10.32604/cmc.2022.020017.

[7] M. Saleem, J. Potgieter, & K. Arif, "Plant Disease Detection and Classification by Deep Learning," *Plants*, vol. 8, no. 11, pp. 468, 2019. https://doi.org/10.3390/plants8110468.

[8] M. Hossen, M. Awrangjeb, S. Pan, & A. Mamun, "Transfer Learning in Agriculture: A Review," *Artificial Intelligence Review*, vol. 58, no. 4, 2025. https://doi.org/10.1007/s10462-024-11081-x.

[9] W. Jhoo and J. Heo, "Collaborative Learning with Disentangled Features for Zero-shot Domain Adaptation," *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 8876-8885, 2021. https://doi.org/10.1109/iccv48922.2021.00877.

[10] Y. Gulzar, "Fruit Image Classification Model Based on MobileNetV2 with Deep Transfer Learning Technique," *Sustainability*, vol. 15, no. 3, pp. 1906, 2023. https://doi.org/10.3390/su15031906.

[11] C. Jackulin and S. Murugavalli, "A Comprehensive Review on Detection of Plant Disease using Machine Learning and Deep Learning Approaches," *Measurement Sensors*, vol. 24, pp. 100441, 2022. https://doi.org/10.1016/j.measen.2022.100441.

[12] N. Ullah, J. Khan, S. Almakdi, M. Alshehri, M. Qathrady, N. El-Rashidy et al., "An Effective Approach for Plant Leaf Diseases Classification based on a Novel DeepPlantNet Deep Learning Model," *Frontiers in Plant Science*, vol. 14, 2023. https://doi.org/10.3389/fpls.2023.1212747.

[13] B. Yang, M. Li, F. Li, Y. Wang, Q. Liang, R. Zhao et al., "A Novel Plant Type, Leaf Disease and Severity Identification Framework using CNN and Transformer with Multi-Label Method," *Scientific Reports*, vol. 14, no. 1, 2024. https://doi.org/10.1038/s41598-024-62452-x.

[14] H. Lu, M. Du, K. Qian, X. He, & K. Wang, "GAN-Based Data Augmentation Strategy for Sensor Anomaly Detection in Industrial Robots," *IEEE Sensors Journal*, vol. 22, no. 18, pp. 17464-17474, 2022. https://doi.org/10.1109/jsen.2021.3069452.

[15] Y. Ma, Z. Zhang, H. Yang, & Z. Yang, "An Adaptive Adversarial Domain Adaptation Approach for Corn Yield Prediction," *Computers and Electronics in Agriculture*, vol. 187, pp. 106314, 2021. https://doi.org/10.1016/j.compag.2021.106314.

[16] A. Prasad, N. Mehta, M. Horak, & W. Bae, "A Two-Step Machine Learning Approach for Crop Disease Detection Using GAN and UAV Technology," *Remote Sensing*, vol. 14, no. 19, pp. 4765, 2022. https://doi.org/10.3390/rs14194765.

[17] Y. Lu, D. Chen, E. Olaniyi, & Y. Huang, "Generative Adversarial Networks (GANs) for Image Augmentation in Agriculture: A Systematic Review," *Computers and Electronics in Agriculture*, vol. 200, pp. 107208, 2022. https://doi.org/10.1016/j.compag.2022.107208.

[18] R. Zhao, Y. Zhu, & Y. Li, "CLA: A Self-Supervised Contrastive Learning Method for Leaf Disease Identification with Domain Adaptation," *Computers and Electronics in Agriculture*, vol. 211, pp. 107967, 2023. https://doi.org/10.1016/j.compag.2023.107967.

[19] A. Goyal and K. Lakhwani, "Integrating Advanced Deep Learning Techniques for Enhanced Detection and Classification of Citrus Leaf and Fruit Diseases," *Scientific Reports*, vol. 15, no. 1, 2025. https://doi.org/10.1038/s41598-025-97159-0.

[20] R. Mohana, C. Reddy, P. Anisha, & B. Murthy, "WITHDRAWN: Random Forest Algorithms for the Classification of Tree-Based Ensemble," *Materials Today Proceedings*, 2021. https://doi.org/10.1016/j.matpr.2021.01.788.

[21] A. Farahani, S. Voghoei, K. Rasheed, & H. Arabnia, "A Brief Review of Domain Adaptation," *Transactions on Computational Science and Computational Intelligence*, pp. 877-894, 2021. https://doi.org/10.1007/978-3-030-71704-9_65.

[22] Y. Li and X. Chao, "Semi-Supervised Few-Shot Learning Approach for Plant Diseases Recognition," *Plant Methods*, vol. 17, no. 1, 2021. https://doi.org/10.1186/s13007-021-00770-1.

[23] R. Singh and R. Sanodiya, "Zero-Shot Transfer Learning Framework for Plant Leaf Disease Classification," *IEEE Access*, vol. 11, pp. 143861-143880, 2023. https://doi.org/10.1109/access.2023.3343759.

[24] P. Wang, K. Han, X. Wei, L. Zhang, & L. Wang, "Contrastive Learning based Hybrid Networks for Long-Tailed Image Classification," *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 943-952, 2021. https://doi.org/10.1109/cvpr46437.2021.00100.

[25] A. Lengyel, S. Garg, M. Milford, & J. Gemert, "Zero-Shot Day-Night Domain Adaptation with a Physics Prior," *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 4399-4409, 2021. https://doi.org/10.48550/arxiv.2108.05137.

[26] X. Li, M. Fang, & B. Chen, "Generalized Zero-Shot Domain Adaptation with Target Unseen Class Prototype Learning," *Neural Computing and Applications*, vol. 34, no. 20, pp. 17793-17807, 2022. https://doi.org/10.1007/s00521-022-07413-z.

[27] B. Gülmez, "A Comprehensive Review of Convolutional Neural Networks based Disease Detection Strategies in Potato Agriculture," *Potato Research*, vol. 68, no. 2, pp. 1295-1329, 2024. https://doi.org/10.1007/s11540-024-09786-1.

[28] X. Zhang, X. Zhang, X. Li, Z. Li, & S. Wang, "Classify Social Image by Integrating Multi-Modal Content," *Multimedia Tools and Applications*, vol. 77, no. 6, pp. 7469-7485, 2017. https://doi.org/10.1007/s11042-017-4657-2.

[29] M. Chandra and S. Bedi, "Survey on SVM and Their Application in Image Classification," *International Journal of Information Technology*, vol. 13, no. 5, pp. 1-11, 2018. https://doi.org/10.1007/s41870-017-0080-1.

[30] M. Tanveer, T. Rajani, R. Rastogi, Y. Shao, & M. Ganaie, "Comprehensive Review on Twin Support Vector Machines," *Annals of Operations Research*, vol. 339, no. 3, pp. 1223-1268, 2022. https://doi.org/10.1007/s10479-022-04575-w.

[31] M. Ouhami, A. Hafiane, Y. Es-Saady, M. Hajji, & R. Canals, "Computer Vision, IoT and Data Fusion for Crop Disease Detection Using Machine Learning: A Survey and Ongoing Research," *Remote Sensing*, vol. 13, no. 13, pp. 2486, 2021. https://doi.org/10.3390/rs13132486.

[32] W. Shafik, A. Tufail, L. Silva, & R. Apong, "Using Transfer Learning-Based Plant Disease Classification and Detection for Sustainable Agriculture," *BMC Plant Biology*, vol. 24, no. 1, 2024. https://doi.org/10.1186/s12870-024-04825-y.

[33] S. Vallabhajosyula, V. Sistla, & V. Kolli, "Transfer Learning-Based Deep Ensemble Neural Network for Plant Leaf Disease Detection," *Journal of Plant Diseases and Protection*, vol. 129, no. 3, pp. 545-558, 2021. https://doi.org/10.1007/s41348-021-00465-8.

[34] X. Wu, X. Fan, P. Luo, S. Choudhury, T. Tjahjadi, & C. Hu, "From Laboratory to Field: Unsupervised Domain Adaptation for Plant Disease Recognition in the Wild," *Plant Phenomics*, vol. 5, pp. 0038, 2023. https://doi.org/10.34133/plantphenomics.0038.

[35] M. Shoaib, B. Shah, S. El–Sappagh, A. Ali, A. Ullah, F. Alenezi et al., "An Advanced Deep Learning Models-Based Plant Disease Detection: A Review of Recent Research," *Frontiers in Plant Science*, vol. 14, 2023. https://doi.org/10.3389/fpls.2023.1158933.

[36] Y. Zhang, J. Wang, Z. Zhang, B. Yu, L. Wang, D. Tao et al., "Domain-Specific Risk Minimization for Out-of-Distribution Generalization," *ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2022. https://doi.org/10.48550/arxiv.2208.08661.

.

[37] P. Kumar, J. Mathew, R. Sanodiya, T. Setty, & B. Bhaskarla, "Zero Shot Plant Disease Classification with Semantic Attributes," *Artificial Intelligence Review*, vol. 57, no. 11, 2024. https://doi.org/10.1007/s10462-024-10950-9.

[38] A. Le, M. Shakiba, & I. Ardekani, "Tomato Disease Detection with Lightweight Recurrent and Convolutional Deep Learning Models for Sustainable and Smart Agriculture," *Frontiers in Sustainability*, vol. 5, 2024. https://doi.org/10.3389/frsus.2024.1383182.

[39] W. Hagar, N. Eladawi, D. Sabry, & H. Salaheldine, "Enhancing Glaucoma Detection Using Convolutional Neural Networks: A Comparative Study of Multi-Class and Binary Classification Approaches," *Alfarama Journal of Basic & Applied Sciences*, 2024. https://doi.org/10.21608/ajbas.2024.324901.1232.

[40] S. Rao, K. Sreekala, P. Rao, N. Shirisha, G. Srinivas, & E. Sreedevi, "Plant Disease Classification using Novel Integration of Deep Learning CNN and Graph Convolutional Networks," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 36, no. 3, pp. 1721, 2024. https://doi.org/10.11591/ijeecs.v36.i3.pp1721-1730.

[41] W. Yang, C. Yang, S. Huang, L. Wang, & M. Yang, "Few-Shot Unsupervised Domain Adaptation via Meta Learning," *2022 IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1-6, 2022. https://doi.org/10.1109/icme52920.2022.9859804.

[42] Y. F. Hendrawan and O. Chekuri, "Exploring Supervised Learning Methods for Predicting Cuisines from Their Ingredients", *Vokasi UNESA Bull. Eng. Technol. Appl. Sci.*, vol. 2, no. 1, pp. 67–76, Mar. 2025. https://doi.org/10.26740/vubeta.v2i1.34153.

BIOGRAPHIES OF AUTHORS

Sani Saleh Saminu is currently pursuing an MSc degree in Artificial Intelligence in the Department of Computer Engineering, Ahmadu Bello University (ABU) Zaria, Nigeria. He received his BSc(Ed) in Computer Science from the same institution. His research interests include plant disease recognition, computer vision, deep learning, zero-shot learning, and intelligent systems for agricultural and medical applications. He can be contacted via:
email: sanisaminusaleh1994@gmail.com

Yusuf Ibrahim is a Senior Lecturer in the Department of Computer Engineering at Ahmadu Bello University, Zaria, Nigeria, with over a decade of teaching and research experience. He holds a B.Eng. in Electrical Engineering (First Class Honours), an MSc and a Ph.D. in Computer Engineering. He is also a Huawei Certified ICT Associate (AI), Huawei Certified Academy Instructor, and a COREN registered engineer. His research interests span Artificial Intelligence, Natural Language Processing, Computer Vision, and Computing.
Email**:** yibrahim@abu.edu.ng

Zaharuddeen Haruna is a lecturer in the Department of Computer Engineering at Ahmadu Bello University, Zaria, Nigeria, with over seven years of teaching and research experience. He holds a B.Eng. in Electrical Engineering, an MSc and a Ph.D. in Control Engineering. He is also a COREN registered engineer. His research interests include Control Systems Design, System Modeling, Intelligent Robotics, Autonomous Systems, and Embedded Systems.
Email: hzaharuddeen@abu.edu.ng

Shehu Mohammed Yusuf is a Senior Lecturer in the Department of Computer Engineering at Ahmadu Bello University, Zaria, Nigeria, with over a decade of teaching and research experience. He holds a B.Eng. in Electrical Engineering, an MSc in Electrical Engineering and a Ph.D. in Computer Engineering. He is also a Huawei Certified ICT Associate (AI), Huawei Certified Academy Instructor, and a COREN registered engineer. His research interests span Artificial Intelligence, Natural Language Processing, Computer Vision, and Bioinformatics.
Email: smyusuf@abu.edu.ng