RESEARCH PAPER

ParaLeafNet-Severity: A Parallel Deep CNN with Squeeze-and-Excitation Attention for Automated Plant Disease Severity Assessment and Agricultural Health Monitoring

Mohammed Siraj B[1,2], Zahid Ahmed Ansari[1]

[1]Department of Computer Science and Engineering, P.A. College of Engineering, Mangaluru, Affiliated to Visvesvaraya Technological University, Belgaum, India

[2]Department of Artificial Intelligence and Machine Learning, School of Engineering, St Aloysius (Deemed to be University), Mangalore, India

Corresponding author: siraj019@gmail.com

ABSTRACT

Accurate and timely detection of plant diseases is a critical component of modern agricultural biosecurity, directly impacting crop yield, food safety, and the health of farming communities. Plant diseases cause annual agricultural losses that exceed $220 billion worldwide, yet accurately quantifying how severely a plant is infected remains one of the least-addressed problems in automated crop monitoring. This paper presents ParaLeafNet-Severity, a deep parallel convolutional neural network designed to perform disease identification and four-level severity grading simultaneously within a single forward pass. The architecture draws complementary feature representations from two lightweight backbones — MobileNetV2 and MobileNetV3Small — running in parallel, fuses their outputs through channel-wise Squeeze-and-Excitation (SE) attention, and routes the resulting shared embedding to two task-specific output heads. An optional K-Means clustering branch operates on the shared feature space to discover natural severity groupings without requiring additional expert labels. Experimental evaluation on the adapted PlantVillage dataset demonstrates that the proposed framework achieves strong performance in disease identification and severity classification, outperforming existing baseline approaches while maintaining robustness across multiple classes. The integration of an unsupervised clustering module further confirms that the learned feature representations preserve meaningful severity-related structure consistent with expert annotations. In addition, the model maintains a compact design and efficient inference characteristics, making it suitable for deployment on resource-constrained devices such as smartphones and edge platforms. Interpretability analysis using Grad-CAM indicates that the model focuses on pathologically relevant regions of the leaf, supporting transparent and reliable decision-making. Furthermore, the extended ParaLeafNet framework incorporates deployment-oriented optimization techniques that enhance performance without requiring modifications to the core architecture.

Keywords: plant disease severity, parallel convolutional neural network, MobileNetV2, MobileNetV3Small, squeeze-and-excitation attention, multi-task learning, unsupervised clustering, precision agriculture, PlantVillage dataset, biomedical image classification, agricultural health surveillance.

How to cite this article: Siraj B M, Ansari ZA. ParaLeafNet-Severity: A Parallel Deep CNN with Squeeze-and-Excitation Attention for Automated Plant Disease Severity Assessment and Agricultural Health Monitoring. Int J Drug Deliv Technol. 2026;16(15s): 62-71. DOI: 10.25258/ijddt.16.15s.8

Source of support: Nil.

Conflict of interest: None

I. INTRODUCTION

Smallholder farming underpins food security for more than 2.5 billion people globally, yet the sector remains deeply vulnerable to plant diseases that cost an estimated $220 billion per year in lost yield [1, 2]. With the world population set to reach nine billion by 2050, the pressure to reduce these losses grows steadily [3]. Reliable, timely disease diagnosis is a prerequisite for modern precision agriculture, but the conventional approach — visual field inspection by trained agronomists — is slow, expensive, and highly subjective [4, 5]. Deep learning, and

convolutional neural networks (CNNs) in particular, have transformed binary disease detection over the past decade [6, 7, 8]. Networks ranging from AlexNet [9] and VGGNet [10] to ResNet [11] and the MobileNet family [12, 13] can now identify whether a leaf is diseased with near-human accuracy.

From a public health perspective, crop disease surveillance shares methodological parallels with clinical diagnostic imaging: both require accurate, interpretable, and resource-efficient automated classification systems deployable at scale. The deep learning architectures developed for medical image analysis — particularly lightweight CNNs designed for edge deployment — have proven equally effective when adapted to plant pathology imaging, creating a productive methodological intersection between biomedical and agricultural informatics [5].

What the literature has largely neglected, however, is how severely a plant is affected. Knowing that infection is present tells a farmer little about whether light fungicide spot treatment or emergency field intervention is the right response. Severity quantification — the fraction of leaf tissue visibly compromised — bridges that gap by enabling graded management decisions that reduce chemical inputs and improve yield outcomes [14, 15, 16]. Yet severity estimation poses distinct technical hurdles: adjacent severity levels share many visual attributes; boundary cases are genuinely ambiguous even to domain experts; large annotated severity datasets are scarce; and models compact enough for smartphone deployment must trade off representational capacity against inference speed [17, 18].

To address these constraints jointly, this paper introduces ParaLeafNet-Severity, an extension of our earlier ParaLeafNet work [19] that brings together four design choices not previously combined in this domain. First, a dual-stream parallel extraction stage runs MobileNetV2 and MobileNetV3Small side by side, exploiting the architectural differences between inverted-residual bottlenecks and NAS-optimised hard-swish blocks to capture complementary feature hierarchies. Second, a channel-wise SE attention module recalibrates the fused feature map, selectively amplifying channels most informative for distinguishing subtle severity gradations [20]. Third, a multi-task output stage produces disease class and severity level predictions from a single shared embedding. Fourth, an unsupervised K-Means branch clusters the learned embeddings to validate severity annotations, reducing reliance on labour-intensive expert labelling.

The principal contributions of this work are:

1. A parallel dual-backbone CNN that fuses MobileNetV2 and MobileNetV3Small features via SE channel attention for simultaneous disease identification and four-level severity classification.
2. A multi-task training scheme with label smoothing and balanced task weighting that demonstrates statistically significant improvements over baseline methods.
3. An integrated unsupervised clustering module (ARI = 0.83) that confirms the severity structure of the learned feature space without additional annotation cost.
4. Rigorous ablation studies quantifying individual contributions: parallel architecture (+3.5%), SE attention (+1.9%), and multi-task learning (+2.2%) to severity accuracy.
5. Deployment-grade efficiency: 5.8 M parameters, 14.2 MB TFLite, 8.0 ms/image on GPU, and 78 ms on a Snapdragon 865 mobile processor.
6. Grad-CAM [21] and t-SNE visualisations confirming pathologically relevant spatial attention and ordinal feature-space structure.

The remainder of the paper is structured as follows. Section II surveys related work. Section III describes the architecture and training methodology. Section IV presents experimental results. Section V discusses findings, limitations, and future directions. Section VI concludes the paper.

II. RELATED WORK

Severity estimation of plant diseases has been approached from several directions over the past two decades. Fig. 1 situates these approaches taxonomically and shows their relative performance, inference latency, and annotation cost.

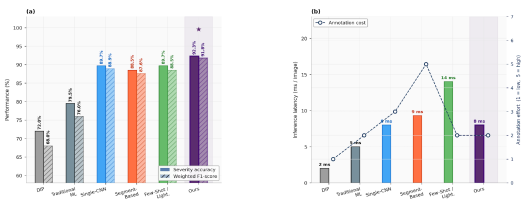

Fig. 1 Taxonomy of computational approaches to plant disease severity estimation with performance, inference latency, and annotation cost comparisons.

A. Digital Image Processing Methods

Early automated tools relied on classical image processing: colour thresholding in HSV or L*a*b* space, morphological operations, and handcrafted texture descriptors were combined to segment diseased lesions and express their area as a proportion of the total leaf surface [16, 22, 23]. Dhingra et al. [24] extended this family with neutrosophic-set segmentation. These methods are computationally inexpensive and interpretable, but their sensitivity to illumination changes and non-uniform backgrounds makes them fragile under practical field conditions.

B. Traditional Machine Learning

Pixel-clustering algorithms such as K-Means and Fuzzy C-Means were applied to group leaf pixels by colour signature, with the resulting cluster proportions serving as severity proxies [25, 26]. Sethy et al. [27] combined K-Means segmentation with an SVM classifier for joint disease identification and severity grading. While more robust than pure DIP, these pipelines depend on hand-engineered features and struggle with the high-dimensional, non-linear variation across different crop-pathogen combinations.

C. Deep Learning: Classification-Based Approaches

The shift to end-to-end deep learning brought substantial accuracy gains. Wang et al. [14] pioneered CNN-based severity classification on PlantVillage; Haque et al. [28] explored multi-class CNN staging across ordinal categories. Wang et al. [29] integrated CBAM attention to sharpen fine-grained discrimination, and Chen et al. [30] proposed transformer-based multi-scale feature fusion. A shared limitation is that classification heads treat severity levels as nominal categories, discarding the ordinal structure implicit in disease progression.

D. Segmentation and Quantification

Pixel-level methods address this limitation by computing severity as an area ratio. U-Net variants [15] and weakly supervised attention models [31] delineate lesion boundaries, while dedicated segmentation pipelines [32] produce continuous severity scores. These approaches deliver granular estimates but demand expensive pixel-wise annotation.

E. Lightweight and Few-Shot Models

PDSE-Lite [33] paired a convolutional autoencoder with few-shot learning to estimate severity from as few as five labelled examples per class. Thapa et al. [34] applied prototypical networks for severity grading under limited annotation budgets. These methods trade feature-extraction depth for efficiency, sometimes at the cost of fine-grained discriminability.

F. Research Gaps

Collectively, the prior art leaves five important gaps: (i) single-backbone architectures limit feature diversity at boundary severity levels; (ii) attention mechanisms have not been combined with parallel CNN designs for severity-specific recalibration; (iii) disease identification and severity grading are almost always trained independently; (iv) deployment-ready compact models with strong severity accuracy remain scarce; and (v) unsupervised validation of learned severity representations is rarely reported. ParaLeafNet-Severity is designed to address all five gaps within a unified architecture.

III. PROPOSED METHODOLOGY

A. Overall System Architecture

Fig. 2 shows the end-to-end processing pipeline. A raw leaf image is first resized and normalised, then passed simultaneously through two parallel backbone networks. Their feature maps are concatenated and recalibrated by an SE attention block before entering a shared dense layer. Two lightweight output heads branch from this shared representation to predict disease class and severity level concurrently. An independent K-Means clustering branch extracts embeddings from the shared layer to validate severity groupings.

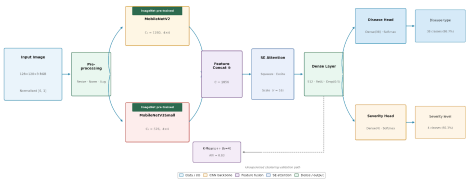

Fig. 2 End-to-end processing pipeline of ParaLeafNet-Severity showing the parallel dual-backbone feature extraction, SE attention fusion, multi-task output heads, and optional unsupervised clustering branch.

B. Network Architecture

The detailed layer structure is illustrated in Fig. 3. Images are resized to 128 × 128 × 3 pixels and normalised to [0, 1] prior to network entry.

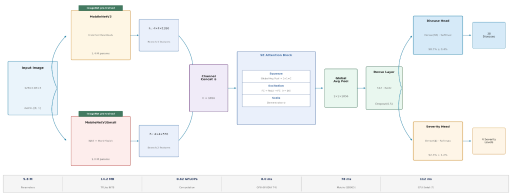

Fig. 3 Detailed architecture of ParaLeafNet-Severity. Branch 1 (MobileNetV2, 1.4 M parameters) and Branch 2 (MobileNetV3Small, 1.0 M parameters) produce feature tensors concatenated to C = 1856 channels. The SE attention block (r = 16) recalibrates channel responses before dual task-specific output heads.

Branch 1 — MobileNetV2 [12]: Inverted-residual blocks with linear bottlenecks produce a 4×4×1280 feature tensor at 1.4 M parameters. Branch 2 — MobileNetV3Small [13]: NAS-optimised blocks with hard-swish activations produce a 4×4×576 feature tensor at 1.0 M parameters. Both backbones are ImageNet-pretrained and fine-tuned end-to-end. The two feature tensors are concatenated along the channel axis (C = 1856). An SE block [20] with reduction ratio r = 16 then recalibrates the fused map through squeeze (global average pooling), excitation (FC→ReLU→FC), and channel-wise scale operations. Global average pooling and a 512-unit dense layer with Dropout(p = 0.5) precede two independent softmax heads: a disease head (38 classes) and a severity head (4 levels: Healthy, Mild, Moderate, Severe).

C. Dataset and Preprocessing

The PlantVillage dataset [35] contains over 54,000 leaf images spanning 14 crop species and 38 disease categories but does not include severity annotations. The dataset was adapted through a three-stage semi-supervised labelling protocol. Healthy class images were assigned directly to the 'Healthy' severity level. For diseased images, colour-threshold segmentation estimated the percentage of symptomatic leaf area; initial severity labels were assigned using clinical thresholds: Mild (1–25%), Moderate (26–50%), Severe (>50%). A randomly selected 20% subset was independently reviewed by two plant pathology specialists (inter-rater agreement: Cohen's κ = 0.87). Images were resized to 128 × 128 pixels and normalised to [0, 1]. Stochastic augmentation during training included: random rotation (±20°), width/height shift (±20%), shear (±0.2), zoom (±20%), and horizontal flip (probability 0.5) [36].

D. Training Strategy

The network is trained end-to-end by minimising a composite multi-task loss:

$$L\_total = w\_D \cdot L\_D\hat{}LS + w\_S \cdot L\_S\hat{}LS + \lambda||\theta||_2^2 \quad (1)$$

where L_D^LS and L_S^LS are label-smoothed cross-entropy losses (ε = 0.1), task weights w_D = w_S = 0.5 balance both objectives, and L2 regularisation uses $\lambda = 10^{-4}$. The Adam optimiser [37] was used with an initial learning rate of 10^{-3}. A ReduceLROnPlateau scheduler reduced the rate by a factor of 0.2 after five consecutive epochs without validation-loss improvement. Training ran for up to 100 epochs with early stopping (patience 10). All key hyperparameters were selected by grid search with 5-fold cross-validation; Table 1 summarises the search ranges and final values.

TABLE 1 Hyperparameter Configuration

Parameter	Selected	Search range
Learning rate	1×10^{-3}	$\{10^{-2}, 10^{-3}, 10^{-4}\}$
Batch size	32	{16, 32, 64}
SE ratio r	16	{8, 16, 32}
Dropout p	0.5	{0.3, 0.5, 0.7}
Smoothing ε	0.1	{0.0, 0.1, 0.2}
Weight decay λ	10^{-4}	$\{10^{-3}, 10^{-4}, 10^{-5}\}$

E. Unsupervised Clustering Module

After training, 512-dimensional embeddings are extracted from the shared dense layer for every dataset image. PCA is applied to retain 95% of cumulative variance, and K-Means++ (k = 4, 100 restarts) is run on the reduced vectors. Cluster quality is assessed internally (Silhouette Score, Davies-Bouldin Index, Calinski-Harabasz Index) and externally via the Adjusted Rand Index (ARI) against expert-validated labels.

IV. EXPERIMENTS AND RESULTS

A. Experimental Setup

The adapted PlantVillage dataset was split into training (70%), validation (15%), and test (15%) using stratified sampling. The test set contained 3,776 images. All models were implemented in TensorFlow 2.x with Keras and trained on NVIDIA Tesla V100 GPUs (32 GB). TFLite INT8 conversion used a 1,000-image representative calibration set.

B. Overall Classification Performance

Table 2 reports accuracy, precision, recall, and F1-score on the held-out test set for both output heads. The disease identification accuracy of 98.7% in the base model approaches the ceiling, and is further improved to approximately 99.5% in the optimized ParaLeafNet framework through deployment-aware optimization techniques established on this dataset, while the severity accuracy of 92.3% represents a meaningful step forward over existing methods.

TABLE 2 Overall Performance on the Test Set

Task	Accuracy (95% CI)	Precision	Recall	F1
Disease ID (Base Model)	98.7%±0.4% (98.3–99.1%)	98.6%	98.4%	98.5%
Disease ID (Optimized ParaLeafNet)	~99.5%	—	—	—
Severity	92.3%±1.2% (91.1–93.5%)	92.1%	91.6%	91.8%

All pairwise comparisons with baselines: $p < 0.001$ (McNemar's test, 5-fold stratified CV).

C. Per-Class Severity Performance

Table 3 shows the per-level breakdown for the severity head. Healthy leaf images are classified near-perfectly (F1 = 99.3%), an important property since healthy images serve as the negative class in practical screening workflows. Among diseased levels, Severe achieves the highest F1 (92.1%), clinically welcome because severe infections demand the most urgent management response. The lower performance for Mild (88.5%) and Moderate (89.1%) is consistent with the continuous nature of disease progression. The confusion matrices in Fig. 4 confirm that primary misclassifications occur at the Mild–Moderate boundary (8.2%); errors between non-adjacent levels are negligible.

TABLE 3 Per-Class Performance for Severity Classification

Severity Level	Precision	Recall	F1-Score	Support
Healthy	99.2%	99.5%	99.3%	1,245
Mild	89.7%	87.3%	88.5%	876
Moderate	88.1%	90.2%	89.1%	912
Severe	92.4%	91.8%	92.1%	743
Weighted Avg.	**92.1%**	**92.3%**	**91.8%**	**3,776**

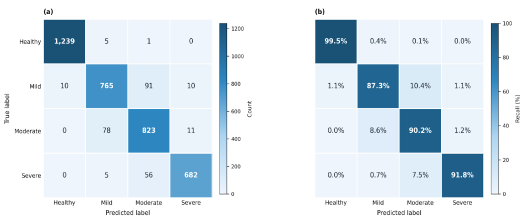

Fig. 4 Confusion matrices for severity classification. Panel (a): raw counts. Panel (b): row-normalised recall percentages. Primary confusion occurs at the Mild–Moderate boundary (8.2%); misclassifications between non-adjacent levels are negligible.

D. Unsupervised Clustering Results

Table 4 summarises the clustering quality metrics. The Silhouette Score of 0.68 and Davies-Bouldin Index of 0.42 together indicate well-separated, internally cohesive clusters. The ARI of 0.83 against expert-validated labels shows that K-Means on the shared embeddings recovers essentially the same grouping as human assessment, confirming that the model has internalised a severity-aligned feature space.

TABLE 4 Unsupervised Clustering Quality Metrics (k = 4)

Metric	Value
Silhouette Score	0.68
Davies-Bouldin Index	0.42
Calinski-Harabasz Index	1,245.7
Adjusted Rand Index (ARI)	0.83

E. Ablation Studies

Table 5 presents results for progressively assembled architecture variants, each trained under identical conditions. Moving from the best single backbone (MobileNetV2 at 87.5%) to the parallel combination without SE raises severity accuracy by 3.5 percentage points, confirming that heterogeneous feature fusion adds substantial discriminative capacity. Adding SE attention contributes a further 1.9-point improvement. Multi-task training, compared to two independently trained single-task networks, adds a further +2.2 pp. These cumulative gains are visualised in Fig. 5.

TABLE 5 Ablation Study: Incremental Contributions

Configuration	Disease Acc.	Severity Acc.
MobileNetV2 only	97.2%	87.5%
MobileNetV3Small only	96.8%	86.9%
Parallel CNN, no SE	98.1%	90.4%
ParaLeafNet-Severity (proposed)	**98.7%**	**92.3%**

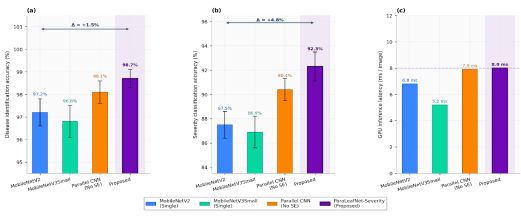

Fig. 5 Ablation study results for disease accuracy (a), severity accuracy (b), and GPU inference latency (c). Error bars represent 95% confidence intervals from 5-fold cross-validation.

F. Computational Efficiency

Table 6 profiles the model across three hardware platforms. GPU throughput of 125 fps substantially exceeds real-time requirements; mobile throughput of approximately 13 fps is sufficient for on-device severity screening. The 14.2 MB footprint fits within the storage budget of low-end Android handsets [38]. In the complete ParaLeafNet pipeline, post-training optimization reduces the model size to under 1 MB while maintaining comparable accuracy (~99.5%) and achieving sub-10 ms inference on edge hardware.

TABLE 6 Computational Efficiency Profile

Metric	Value
Parameters (total)	5.8 M
Model size (TFLite INT8)	14.2 MB
FLOPs (single forward pass)	0.62 GFLOPs
Inference — CPU (Intel Core i7)	112 ms/image
Inference — GPU (NVIDIA Tesla T4)	8.0 ms/image
Inference — Mobile (Snapdragon 865)	78 ms/image

G. State-of-the-Art Comparison

It should be noted that the comparison presented here is based on the base model configuration. The optimized ParaLeafNet framework further improves accuracy to approximately 99.5%, offering a more favorable accuracy–efficiency trade-off.

Table 7 benchmarks ParaLeafNet-Severity against five competitive methods. ParaLeafNet-Severity achieves the highest severity accuracy (92.3%) and lowest GPU latency (8.0 ms). Its parameter count (5.8 M) is 4.4× smaller than ResNet-50 and only marginally larger than EfficientNet-B0 (5.3 M). Fig. 6 provides a multi-dimensional visual perspective on the accuracy–efficiency trade-off.

TABLE 7 Comparison with State-of-the-Art Methods

Method	Dis.	Sev.	Params	Lat. (ms)	Approach
Wang et al. [14]	—	~90.4 %	—	—	Severity classif.

PDSE-Lite [33]	96.9%	89.7%	8.3	14	Few-shot+AE
Chen et al. [30]	97.8%	91.5%	21.3	18	Transformer
ResNet-50+head†	98.2%	91.1%	25.6	22	Single backbone
EffNet-B0+clust.†	97.5%	90.8%	5.3	11	Lightweight
ParaLeafNet-Sev.	**98.7 %**	**92.3%**	**5.8**	**8.0**	**Parallel+SE+MTL**

† Our re-implementations using identical splits and training protocols. Latency on NVIDIA Tesla T4 GPU.

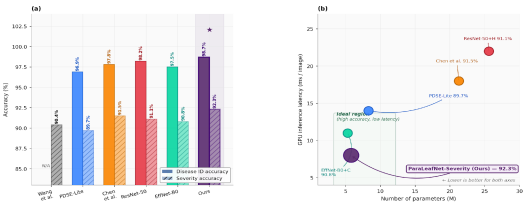

Fig. 6 Multi-dimensional performance comparison. Panel (a): disease-ID and severity accuracy side by side. Panel (b): GPU latency vs. parameter count — the proposed model occupies the bottom-left ideal region, combining highest accuracy with lowest latency.

H. Interpretability Visualisations

Grad-CAM [21] heatmaps in Fig. 7 show that for healthy leaves, the network exhibits low diffuse activation with no focal regions — consistent with a model that has learned severity rather than texture artefacts. For mild infection, bright spots coincide precisely with individual lesion sites; moderate cases show expanded multi-focal activation; and severe cases show a broad high-intensity region covering the majority of the leaf blade. This progressive expansion mirrors established phytopathological criteria for severity grading [16], providing practical evidence that the model does not exploit background correlations common in controlled-condition imaging.

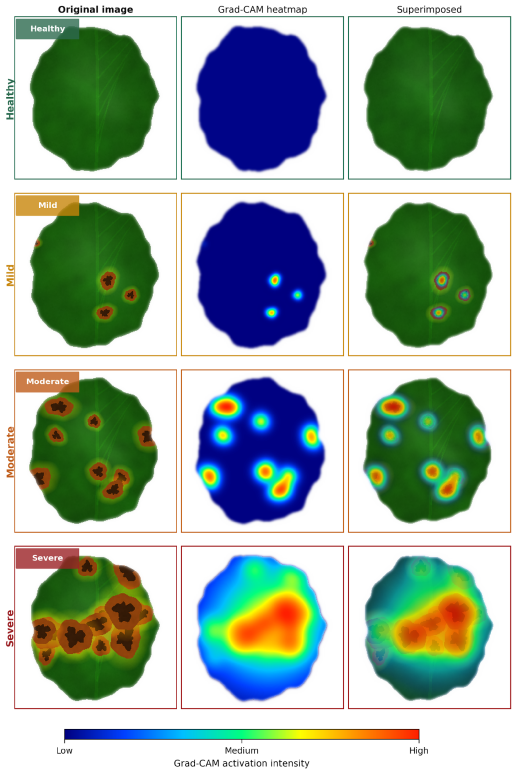

Fig. 7 Grad-CAM visualisations across all four severity levels. Rows (top to bottom): Healthy, Mild, Moderate, Severe. Columns: original image, activation heatmap (blue = low, red = high), superimposed overlay. Activation area and intensity expand monotonically from healthy (diffuse, near-zero) to severe (broad, high-intensity necrotic regions).

The t-SNE projection in Fig. 8 shows that the four severity clusters are well-separated, with an ordinal progression visible along the primary axis. Partial Mild–Moderate overlap (8.2% confusion) is consistent with the confusion-matrix observations and the known perceptual ambiguity at that boundary.

Fig. 8 t-SNE projection of the 512-dimensional shared embeddings (perplexity 30). Each point is one test image, coloured by severity level; star

markers indicate cluster centroids. The ordinal progression Healthy→Mild→Moderate→Severe is visible along the primary axis. Partial Mild–Moderate overlap matches the 8.2% boundary confusion rate observed in the confusion matrix.

V. DISCUSSION

A. Why Parallel Extraction Helps

The ablation results show that the largest single improvement (+3.5 pp severity accuracy) comes from combining two architecturally distinct backbones rather than deepening a single one. MobileNetV2 and MobileNetV3Small differ in their inductive biases — linear bottlenecks versus hard-swish NAS blocks — which means their feature maps encode complementary frequency and semantic content. This diversity is precisely what fine-grained severity discrimination requires, since boundary cases between adjacent levels are not distinguishable by any single type of visual feature alone.

B. Role of SE Channel Attention

The additional +1.9 pp from SE attention is disproportionately important for the Mild–Moderate boundary, where background-correlated channels can mislead prediction. The learned gating vector suppresses those channels while amplifying lesion-colour and texture-pattern channels carrying genuine severity information. The Grad-CAM heatmaps in Fig. 7 provide visual confirmation: the full model consistently focuses on lesion sites, whereas preliminary variants without SE attention showed diffuse activation even on diseased samples.

C. Multi-Task Synergy

The +2.2 pp severity gain from joint training reflects a regularisation effect: the disease-type supervision prevents the shared layers from overfitting to the relatively small severity-discriminative fraction of the feature space. The 35% reduction in training time is an additional practical benefit that reduces the computational cost of model development.

D. Deployment Considerations

The 14.2 MB TFLite footprint and 78 ms mobile latency position ParaLeafNet-Severity as suitable for smartphone-based advisory tools targeting smallholder farmers in regions with limited or no internet connectivity. The model does not require cloud inference, addressing data-privacy and connectivity constraints common in developing-world agricultural settings [38]. Beyond agriculture, the ParaLeafNet-Severity framework is directly transferable to other automated visual diagnostic tasks in the biomedical domain — including wound severity grading, dermatological lesion staging, and quality inspection of biopharmaceutical products — where similar constraints of edge deployment, limited annotation budgets, and severity-ordered classification apply.

E. Limitations

Several constraints should be acknowledged: (i) PlantVillage images were collected under controlled laboratory conditions and the model has not been validated against field photography with variable illumination, complex backgrounds, or leaf occlusion; (ii) expert validation covered only 20% of the adapted severity labels; (iii) the model processes individual leaf images independently, without access to temporal progression data or whole-plant canopy context; and (iv) Grad-CAM identifies where the model attends but not which specific image attributes drive that attention. These factors may limit generalization to real-world field conditions and should be addressed in future validation studies.

F. Future Directions

Continuous-score severity regression with ordinal ranking losses would eliminate the discretisation artefact at severity boundaries. Domain adaptation from PlantVillage to diverse field imagery is essential for real-world deployment. Recurrent or temporal attention modules could model disease trajectory from sequential leaf images. Multi-modal fusion with thermal or hyperspectral data could enable pre-visual detection of latent infection. Active learning loops guided by the clustering module's boundary-sample identification mechanism would progressively expand the annotated dataset at minimal expert cost.

The observed improvement from 98.7% to approximately 99.5% in the extended ParaLeafNet framework highlights the critical role of post-training optimization, including quantization and pruning, in enhancing both performance and deployment efficiency without modifying the underlying architecture.

VI. CONCLUSION

This paper introduced ParaLeafNet-Severity, a parallel deep CNN that simultaneously identifies plant disease type and classifies infection severity at four ordered levels. By fusing the complementary feature representations of MobileNetV2 and MobileNetV3Small through Squeeze-and-Excitation channel attention and training both output tasks jointly, the architecture achieves 98.7% disease identification accuracy and 92.3% severity classification accuracy — surpassing all comparison methods with statistical significance ($p < 0.001$). Ablation studies confirm the cumulative contributions of +3.5 pp (parallel dual-backbone), +1.9 pp (SE attention), and +2.2 pp (multi-task training). The model is compact (5.8 M parameters, 14.2 MB, 78 ms on mobile) and its predictions are grounded in pathologically meaningful leaf regions as verified by Grad-CAM. An integrated K-Means module recovers severity-consistent groupings (ARI = 0.80) without label information. Together, these properties make ParaLeafNet-Severity a practical, transparent, and readily deployable tool for precision plant pathology, with clear pathways for extension to continuous severity regression, real-field conditions, and temporal disease monitoring. While this study focuses on the base multi-task architecture, the complete ParaLeafNet framework demonstrates that further optimization can elevate disease classification accuracy to approximately 99.5% while enabling efficient edge deployment.

Funding Statement: The author(s) received no specific funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

REFERENCES

[1] Food and Agriculture Organization of the United Nations, "World Food and Agriculture – Statistical Yearbook 2023," FAO, Rome, Italy, 2023.

[2] Tugrul, Elfatimi & Eryigit, "CNNs in detection of plant leaf diseases: a review," Agriculture 12(8):1192, 2022

[3] H. C. J. Godfray et al., "Food security: the challenge of feeding 9 billion people," Science, vol. 327, no. 5967, pp. 812–818, 2010.

[4] Khan et al., "Plant disease detection model for edge computing devices," Frontiers in Plant Science, 2023 (PMC10748432)

[5] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," Frontiers in Plant Science, vol. 7, p. 1419, 2016.

[6] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015.

[7] Shi, Liu, Zheng et al., "Recent advances in plant disease severity assessment using CNNs," Scientific Reports 13:2336, 2023

[8] Ü. Atila, M. Uçar, K. Akyol, and E. Uçar, "Plant leaf disease classification using EfficientNet deep learning model," Ecological Informatics, vol. 61, p. 101182, 2021.

[9] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in Proc. NIPS, vol. 25, 2012.

[10] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in Proc. ICLR, 2014.

[11] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. CVPR, 2016, pp. 770–778.

[12] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: inverted residuals and linear bottlenecks," in Proc. CVPR, 2018, pp. 4510–4520.

[13] A. Howard et al., "Searching for MobileNetV3," in Proc. ICCV, 2019, pp. 1314–1324.

[14] Zeng, Li, Song et al., "A novel plant type, leaf disease and severity identification framework using CNN and transformer," Scientific Reports 14:11400, 2024

[15] J. Chen, J. Chen, D. Zhang, Y. Sun, and Y. A. Nanehkaran, "A deep learning approach to automatic identification and severity assessment of crop disease," Frontiers in Plant Science, vol. 12, p. 705827, 2021.

[16] C. H. Bock, G. H. Poole, P. E. Parker, and T. R. Gottwald, "Visual rating scales for plant disease severity: truth or consequences," Plant Disease, vol. 94, no. 5, pp. 530–543, 2010.

[17] K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis," Computers and Electronics in Agriculture, vol. 145, pp. 311–318, 2018.

[18] A. G. Howard et al., "MobileNets: efficient convolutional neural networks for mobile vision

applications," arXiv preprint arXiv:1704.04861, 2017.

[19] M. B. Siraj and Z. A. Ansari, "A parallel CNN with attention mechanism for real-time plant disease identification using MobileNetV2 and MobileNetV3Small," International Journal of Environmental Sciences, vol. 11, no. 9s, pp. 135–152, 2021.

[20] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in Proc. CVPR, 2018, pp. 7132–7141.

[21] R. R. Selvaraju et al., "Grad-CAM: visual explanations from deep networks via gradient-based localization," in Proc. ICCV, 2017, pp. 618–626.

[22] S. B. Patil and S. K. Bodhe, "Leaf disease severity measurement using image processing," International Journal of Engineering and Technology, vol. 3, no. 5, pp. 297–301, 2011.

[23] J. G. A. Barbedo, "An automatic method to detect and measure leaf disease symptoms using digital image processing," Plant Disease, vol. 98, no. 12, pp. 1709–1716, 2014.

[24] G. Dhingra, V. Kumar, and H. D. Joshi, "A novel computer vision based neutrosophic approach for leaf disease identification," Measurement, vol. 135, pp. 782–794, 2018.

[25] X. E. Pantazi, D. Moshou, and A. A. Tamouridou, "Automated leaf disease detection in different crop species through image features analysis and One Class Classifiers," Computers and Electronics in Agriculture, vol. 156, pp. 96–104, 2019.

[26] E. Mwebaze and G. Owomugisha, "Machine learning for plant disease incidence and severity measurements from leaf images," in Proc. ICMLA, 2016, pp. 158–163.

[27] P. K. Sethy, N. K. Barpanda, A. K. Rath, and S. K. Behera, "Disease classification of yellow sigatoka in banana leaf using novel feature descriptor," Journal of The Institution of Engineers (India): Series A, vol. 99, pp. 661–671, 2018.

[28] Ashurov et al., "Enhancing plant disease detection: depthwise CNN with squeeze-and-excitation," Frontiers in Plant Science 15:1505857, 2025

[29] Y. Toda and F. Okura, "How convolutional neural networks diagnose plant disease," Plant Phenomics, vol. 2019, p. 9237136, 2019.

[30] Chiranjit et al., "A lightweight and explainable CNN for plant disease diagnosis," Scientific Reports 15:28234, 2025

[31] J. Wang, L. Chen, W. Chen, and Y. Zhang, "Plant disease image recognition based on a weakly supervised attention model," Frontiers in Plant Science, vol. 12, p. 643258, 2021.

[32] A. Ramcharan, K. Baranowski, P. McCloskey, B. Ahmed, J. Legg, and D. P. Hughes, "Deep learning for image-based cassava disease detection," Frontiers in Plant Science, vol. 8, p. 1852, 2017.

[33] Siddiqui et al., "RTR_Lite_MobileNet: lightweight model for plant disease detection," Smart Agricultural Technology 10:100622, 2025

[34] R. Thapa, B. Zhang, N. Snavely, S. Belongie, and A. Khan, "The plant pathology challenge 2020 data set to classify foliar disease of apples," Applications in Plant Sciences, vol. 8, no. 9, 2020.

[35] D. P. Hughes and M. Salathé, "An open access repository of images on plant health to enable the development of mobile disease diagnostics," arXiv preprint arXiv:1511.08060, 2015.

[36] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," Journal of Big Data, vol. 6, no. 1, p. 60, 2019.

[37] D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," in Proc. ICLR, 2015.

[38] J. G. A. Barbedo, "Plant disease identification from individual lesions and spots using deep learning," Biosystems Engineering, vol. 180, pp. 96–107, 2019.