RESEARCH ARTICLE Indian Journal of Agricultural Research, Volume 60 Issue 9: 1458-1465 (September 2026)

A Novel Hyperspectral Imaging based Framework for Tomato Leaf Disease Detection

Poonam Chaudhary[1], Sneha Kandacharam[1]

10.18805/IJARe.A-6519

ABSTRACT

Background: The importance of early identification of tomato bacterial leaf spot (BLS) is essential to minimize the losses in yields and enhance the use of precision agriculture.

Methods: The paper demonstrates a lightweight deep learning model, the Dual Multi-Layer Perceptron Feature Fusion Network (DMLPFFN) as a hyperspectral-based approach to the four tomato disease stages, which includes healthy, asymptomatic, early and late infection. The USDA hyperspectral data (168300 bands) has spectral heterogeneity which is solved by per-image Principal Component Analysis (PCA) to normalize the inputs into a single 128-dimensional space. Spectral-spatial patches of 9×9 are overlapping and normalized and then trained upon before training the model. The DMLPFFN combines the feature extraction process that consists of multi-scale dilated convolutions and global contextual modelling with lightweight element-wise fusion. Training strategies that are imbalance-aware are useful in increasing robustness.

Result: The model had a weighted F1-score of 0.9718 and a validation accuracy of 98.22% and test accuracy of 97.18%. These findings can be defined as good generalization and lower computational complexity, which makes the framework applicable to real-time agricultural application.

Key words: CNN-DMLP hybrid, Feature fusion, Hyperspectral imaging (HSI), PCA, Precision agriculture, Spectral-spatial classification, Tomato bacterial leaf spot.

INTRODUCTION

The agricultural industry in the world is changing at an alarming rate as it adopts new and modern technologies to help guarantee the crop health, high productivity and global food security. As the population is demanding sustainable methods of farming, the demand of diagnostic methods that are rapid, accurate and non-destructive is on the increase.

Tomato is one of the most popular and economical vegetables in the world among the major horticultural crops. But the Bacterial leaf spot (BLS) is a serious menace to tomato production that it can propagate rapidly in favourable environmental conditions and result into massive loss of yield (Abdulridha *et al.*, 2020, Sreedevi *et al.*, 2023). Timely and accurate diagnosis of BLS is thus important in reducing economic losses and ensuring massive loss in crops. Conventionally, experts, simple spectral vegetation indices and statistical analysis of reflectance data have been used as conventional methods to diagnose tomato diseases due to visual inspection. Although these methods are simple and inexpensive, they are usually not sensitive to the biochemical and structural variations that take place in the initial stages of infection. Over the last few years, hyperspectral imaging (HSI) has become a very promising agent in detecting plant diseases due to its ability to record spectral details in hundreds of narrow wavelength bands (Barreto *et al.*, 2023, Duhan *et al.*, 2025). Zhang *et al.* (2024); Krishnan and Kumar (2023). HSI allows automated identification of compound spectral-spatial patterns of plant stress and disease, when combined with deep learning (DL). Many studies have examined the different analytical models of identifying plant disease based on hyperspectral data, but each offers distinct advantages and limitations (Duhan *et al.*, 2025; Xie *et al.* (2022) proposed an early benchmark study with statistical band reduction based on Successive Projections Algorithm (SPA) for wavelength selection with an Extreme Learning Machine (ELM) classifier. Their approach proved useful in dimensionality reduction and enhanced classification. But it was very sensitive to manually chosen subsets of spectral and poor in scalability and flexibility. On the same note, Zhou *et al.* (2025) proved that it is possible to detect tomato leaf diseases early with the help of classical machine learning. They used Gaussian filtering as the method of reducing spectral noise and used optimization algorithms to select the discriminative wavelengths to use in SVM classification with an accuracy of 97%.

The paper has pointed out the usefulness and intuitiveness of spectral band optimization. Nevertheless, DMLP goes further than the dimensionality reduction

[1]The North Cap University, Gurugram-122 017, Haryana India.

Corresponding Author: Poonam Chaudhary, The North Cap University, Gurugram-122 017, Haryana India.
Email: poonamchaudhary@ncuindia.edu

ORCID: 0000-0001-5529-5561

How to cite this article: Chaudhary, P. and Kandacharam, S. (2026). A Novel Hyperspectral Imaging based Framework for Tomato Leaf Disease Detection. *Indian Journal of Agricultural Research.* **60(9):** 1458-1465. doi: 10.18805/IJARe.A-6519.

Submitted: 13-02-2026 **Accepted:** 19-06-2026 **Online:** 10-08-2026

methods like PCA by including spectral-spatial fusion and multiple-branch receptive field expansion. The current structure is more focused on computational efficiency and real-time spatial flexibility of hyperspectral diagnostics as compared with Zhou *et al*. (2025) who focused their attention on environmental stability and simplicity. Reis *et al*. (2024) investigated the effect of nitrogen fertilization on the spectral detectability of tomatoes with bacterial leaf blight. The data collected by them showed that nutrient stress alters spectral responses, which impact disease classification. They improved predictive performance with vegetation indices like NDVI with combination with other spectral characteristics. They, however, used methodology that involved multi-stage calculations and several data fusion steps. Conversely, the perceptron-based framework suggested makes processing simpler in the sense that implicit spectral-spatial fusion is carried out in a single unified network, which makes real-time implementation feasible without the need to run complex preprocessing pipelines. Shirly (2025) was interested in measuring severity of diseases by the changes in spectral signatures. Although their hybrid architecture was quite useful in tracking disease evolution with time, it was based on distinct severity mapping modules, which could not be easily combined into a simpler classification architecture. The DMLPFFN on the other hand is planned to give quick and precise projections within a unified model framework without the necessity to have additional computational units. Even though some researchers recorded high results with the help of DBO and BiLSTM models separately, the DMLP optimizes the incorporation of spatial features at the initial stages of detection that is critical to timely intervention (Zhou *et al*., 2025).

There are several other works, which have led to the changing environment of detection of hyperspectral diseases. Custom spectral features attracted (Guo *et al*. (2023) who integrated the play of random forest (RF) and Partial least squares regression (PLSR). Nevertheless, they could not generalize because they used handcrafted features. Mahlein *et al*. (2018) combined spectral reflectance and texture descriptors to enhance resiliency, however, at the expense of preprocessing. Zhang *et al*. (2024) investigated multi-temporal tracking, which allows tracking the disease development on repeating early foliar bacterial infections in tomato leaves for non-invasive hyperspectral insights. Such methods are more effective, but they need more than one data capture. Xu *et al*. (2026). also optimized the choice of bands to minimize the computational cost but could have lost subtle biochemical signals during the process. The DMLPFFN tackles those issues by relying on PCA-based compression to maintain the variance but also expand the spectral-spatial representations. Wei *et al*. (2025) enhanced the accuracy of classification using the combination of hyperspectral and RGB using optical imaging but added complexity to hardware. Xu *et al*. (2022) proposed spectral attention networks that did not only improve the emphasis on wavelengths but also added to the number of computations. Multi-branch perceptron the lightweight multi-branch perceptron architecture is an efficient feature weighting architecture which uses no costly attention mechanisms.

More recent models have tried new architectures including ResNet variants, transformers, GAN-based augmentation and multimodal fusion network. Although such systems can be accurate (more than 95%), they usually have deep backbones that require millions of parameters and complex fusion blocks, making the deployment of edges difficult. To overcome these drawbacks, the DMLPFFN is a multi-branch spectral-spatial fusion implementation of lightweight perceptron blocks, which is highly expressive at low computational cost. Hyperspectral tomato dataset provided by USDA Yadav *et al.* (2024) in this study was also well pre-processed to guarantee uniformity and strength. Water absorption bands were eliminated and band counts (*e.g.* 168 and 300 bands) were reduced to 128 components with principal component analysis (PCA) to 128 components. The last data set was made up of around 12,740 samples that were arranged in a format of 9×9 Spectro-spatial patches with a stride of 4. In order to ensure an equal representation, data were partitioned into stratified 70/15/15 train, validation and test split. To overcome the imbalance in classes, weighted loss and balanced sampling were used in training (Acheraiou *et al*., 2025; Reddy, 2026). The proposed dual multi-layer perceptron feature fusion network is a joint feature learning algorithm that is trained on full spectral-spatial patches. This removes the need to use manually made wavelength selection and increases feature richness. Its architecture includes 4D patch tensors of shape (N,128,9,9) to combine the features. Instead of using deep convolutional stacks or recurrent layers, it has a multi-branch architecture, which has low-, middle and high-level feature pathways. A Local Perceptron feature module which has dilated convolutions is an efficient expansion of the spatial receptive field. This architectural design allows the spectral and spatial features to be fused through element-wise addition, thus satisfying the need for greater efficiency in early-stage detection and better computational scalability when compared with previous methods.

After the introduction, Section 2 presents the suggested methodology of early disease detection of tomato bacterial leaf spot disease with the help of the USDA hyperspectral dataset and includes a block diagram of the research process along with explanations of evaluation measures. Section 3 addresses the findings of the experiments and their comparison with the current approaches. Lastly, the conclusion of Section 4 sums up the study by stating the opportunities of improving the work and extending it in the future.

MATERIALS AND METHODS

The proposed methodology transforms heterogeneous hyperspectral image (HSI) cubes into standardized spectral spatial patches to classify tomato disease stages correctly with the help of the Dual Multi-Layer Perceptron Feature Fusion Network (DMLPFFN). The entire process, such as the harmonization of datasets, preprocessing, architecture design, training settings and evaluation, is presented in section. It uses the hyperspectral tomato dataset funded by USDA that measures the reflectance changes in four disease stages that are measured using days after inoculation (DAI): Healthy (0 DAY), Asymptomatic (1 DAY), Early infection (23 DAY) and Late infection (47 DAY). The preliminary exploratory analysis showed the heterogeneity in spectral resolution as there were two configurations of sensors 168-band SWIR and 300-band VISNIR cubes. This dissimilarity makes it impossible to ingest this model directly due to a difference in spectral dimensionality amongst samples. In solving this, a harmonization pipeline is introduced. Individual hyperspectral scans in each of the raw ENVI-format are loaded and labels of diseases are detected in DAI metadata within the file name. To reduce sensor and illumination bias, images are spatially standardized by removing the image to a 64×64 pixel centered region and calibrated. Water absorptions are eliminated using spectral cleaning (13501460 nm and 18001950 nm) to reduce distortion.

Then Principal Component Analysis (PCA) is used to each cube in order to normalize dimensionality to 128 components.

$$Z = [X - \text{mean}(X)] \times G_k \quad \text{...(i)}$$

Where,

G_k= The best 128 eigenvectors which retain the greatest spectral variance.

The size of the spectral-spatial patches that are extracted is 9×9×128 and stride 4, giving about 12740 samples (Chen *et al*., 2014; Chang *et al*., 2014).

$$P(i, j) = I[i: i + h, j: j + w] \quad \text{...(ii)}$$

Where,

P(i, j)= The patch at (i, j).

h and w= The patch dimensions.

I= The input HSI.

Z-score normalization is applied to each band. The output from this step is the patches in a new shape (N, 128, 9, 9) suitable for PyTorch as it is now in channel-first format. Afterward, data is stratified and then the subsets for training (70%), validation (15%) and testing (15%) are created from it using train_val_test_split, which ensures that the four disease stages are equally represented and results in 1,912 test samples.

Fig 1 (Block diagram) depicts the entire methodological workflow that begins with raw heterogeneous HSI cubes then moves through spectral cleaning, PCA harmonization, patches extraction, normalisation, DMLPFFN-based features extraction and concludes with disease-stage classification and thus gives a pictorial overview of the end-to-end one.

Table 1 gives summary of dataset harmonization and pre-processing, standardization of heterogeneous spectral inputs to train the model. The preprocessing stages are visualized in Fig 2 (Preprocessing pipeline), which illustrates sequential processing of the stages such as ENVI file loading, label extraction by use of the DAI, 64×64 crop, calibration, elimination of water absorption bands, PCA dimensionality reduction, patch extraction and normalization prior to ingestion by the model.

Further, DMLPFFN is a sparse CNN based hybrid network, which is a lightweight multi-branch structure and is developed to effectively identify multi-scale spectral-spatial features. Each harmonic patch (128, 9, 9) is made from three parallel branches.

The low-level branch employs a 3 × 3 convolution to capture the fine lesion textures; the mid-level block employs a 5 × 5 convolution to detect contextual infection spreads and the high-level block employs a DMLP block to learn deeper spectral-spatial interactions:

$$f_{global} = FC_2\,[\text{ReLU}\,\{FC_1\,(\text{GAP}\,(f))\}] \quad \text{...(iii)}$$

and during the expansion of the receptive fields with the Local Perceptron three parallel dilated convolutions are used (dilation = 1, 2, 5):

$$f_{local} = f_1 + f_2 + f_5 \quad \text{...(iv)}$$

This type of multi-dilation strategy captures fine, mid-range and broad spatial dependencies and is not deeply stacked. Upon the independent processing of the branches, feature fusion is accomplished through the light weight element-wise addition:

$$f_{used} = f_{low} + f_{mid} + f_{high} \quad \text{...(v)}$$

Search fused representation is subjected to Global Average Pooling (GAP) and classification where $\hat{y}$ predicted probability distribution over four classes:

$$\hat{y} = \text{Softmax}\,(W_z + b) \quad \text{...(vi)}$$

The forward-processing pipeline is graphically described in Fig 3 (DMLP-FFN Model Architecture) in which the three parallel perceptual branches (low, mid, high), the DMLP blocks with Global and Local Perceptron modules, element-wise fusion and the ultimate GAP-based classifier which produces probabilities of the four stages of the disease are depicted.

The specialized module

Local Perceptron (inside the DMLPBlock see Table 2) is the main idea that is deeply integrated in each DMLPBlock. By multi-dilation convolutions, it broadens the spatial receptive field, thus the network can catch not only the local but also the remote spatial dependencies in spectral-spatial patches. In a Local Perceptron, instead of stacking deeper convolutional layers, three parallel dilated convolutions with different dilation factors are done and their outputs are summed up to get the final. The light weight multi-branch design which is a tradeoff between efficiency and representational richness is summarized in Table 2.

The optimizer function AdamW (lr=0.0002, batch size=16, weight decay= 0.0001) using Cosine Annealing Warm Restarts scheduling and early stopping (patience=25 epochs) is used in training. Weighted cross-entropy loss is used to deal with class imbalance and balanced sampling to make all classes represented in the mini-batch (size =16) proportionately. To measure convergence and resistance to overfitting, training accuracy and validation accuracy and loss curves are used to measure model performance. Checkpoint with the highest validation accuracy (0.9822 at epoch 166) is chosen to be finally tested, which proves to be very generalized.

RESULTS AND DISCUSSION

The outcome of the experiments shows that the presented DMLPFFN model can learn and extrapolate intricate spectral-spatial correlation derived utilizing the harmonized data of the hyperspectral dataset (as illustrated in Fig 4 and 5). The convergence of the training process was steady with the training and validation losses steadily declining with training epochs. The associated graphs of accuracy replicated the same trend upward, which reflected an effective optimization and controlled gradient flow. The maximum performance of the network was attained at the Epoch 166 when the model attained training accuracy of 99.88 (0.9988) and a validation accuracy of 98.22 (0.9822).

The relationship between training and validation curves is very close, which indicates that the model was not affected by serious overfitting, which is a problem in agricultural datasets that are usually small or unbalanced. The small distance between the two curves proves that the network has acquired transferable spectral spatial patterns and not just memorizing the training samples. The three key design considerations that have led to this stability are, PCA-based spectral harmonization, weighted loss

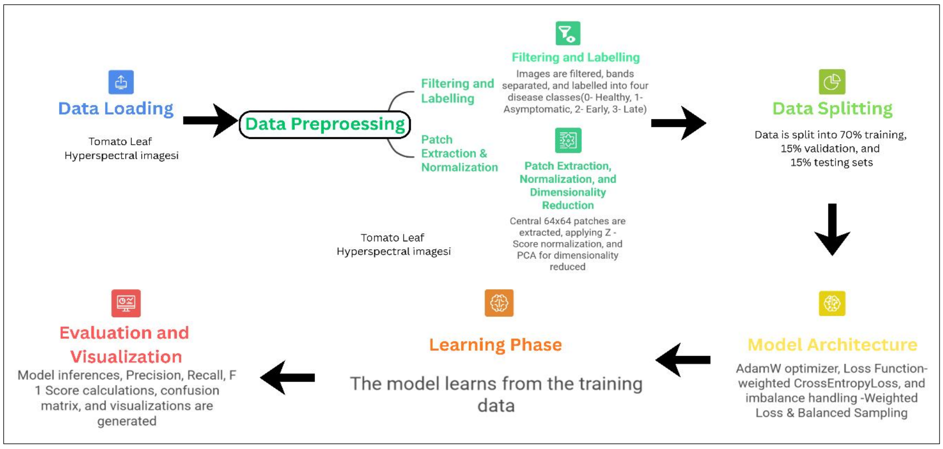

Fig 1: Block diagram.

Table 1: Dataset harmonization and pre-processing.

Block	Layer type	Input channels	Output filters	Kernel size	Stride	Padding/ dilation
Input	Patch tensor (B, H, W)	128	-	9 × 9	-	-
Low-level branch	Conv2D	128	16	3	1	1
	BatchNorm2d	16	-	-	-	-
	ReLU	-	-	-	-	-
	DMLPBlock (details below)	16	16	-	-	-
	Conv2D (Projection)	16	64	1	1	0
Middle-level branch	Conv2D	128	32	5	1	2
	BatchNorm2d	32	-	-	-	-
	ReLU	-	-	-	-	-
High-level branch	DMLPBlock (details below)	128	64	-	-	-
	Conv2D (Projection)	64	64	1	1	0
Fusion	Element-wise Addition	64, 32, 64	64	-	-	-
Output	AdaptiveAvgPool2d (GAP)	64	-	1	-	-
	Linear (Classifier)	64	4 (Classes)	-	-	-

balanced sampling and lightweight multi-branch fusion architecture. To make the final analysis, the checkpoint whose validation accuracy was the largest was chosen and ran on a separate test set of 1,912 samples.

The DMLPFFN also had a total accuracy of 97.18% (0.9718) and a weighted average F1-score of 0.9718 (Table 3). The Macro-average score of 0.9650 also shows that there was a steady performance level among the four disease stages.

These findings affirm the strength of the architecture to differentiate between health, Asymptomatic, early and late infection stages. The in-depth analysis of the data in terms of classes shows high levels of accuracy and recall of all groups. The model was accurate in determining 1,182 Late-stage (Class 3) and 312 Healthy (Class 0) samples (Table 4) indicating a high level of sensitivity to non-infected and severe conditions. Most of the classes had precision, recall and F1-scores greater than 0.95, indicating good discrimination even in intermediate stages like asymptomatic and early infection. These findings are also supported by the confusion matrix, the diagonal is strongly dominated, which means that it correctly classifies in most cases. Minor patterns of misclassification were only detected. Two notable trends include: Late-to-Healthy misclassifications (17 samples): There were few samples of late-stage leaves that were predicted to be healthy. This possibly is because of the similar reflection properties within the near-infrared range where the spectral decay of the necrotic tissue can resemble non-infected leaves. Early to Late misclassification (12 samples): There is some early-to-late misclassification in that the early-stage samples could be mixed up with late-stage samples, probably due to biochemical transitions in the course of infection that could lead to intermediate spectral patterns. These misclassifications combined make up the less than 3 per cent of the total predictions which is unsurprising considering the fine line spectral differences between progressive disease stages.

Overall, the confusion matrix (Fig 6) validates significant generalization as well as inter-class separability. As compared to the existing hyperspectral disease detection research that has achieved a hyperspectral disease detection with a range of accuracies of 90% to 98%, the DMLPFFN is competitive and at the same time, it is much less bulky in terms of its architecture. The proposed design does not use deep residual networks, LSTM-based models, or multimodal fusion systems, which have many computational resource requirements and can perform similarly when the size of the multi-branch design is small, with element-wise fusion. Local Perceptron: The Local Perceptron module, which uses dilated convolutions (dilations 1, 2, 5), is a simple expansion of the receptive field without stacking that does not require many parameters and will take less time to run inference in one. Overall, the findings confirm the fact that the DMLPFFN delivers a good compromise between precision,

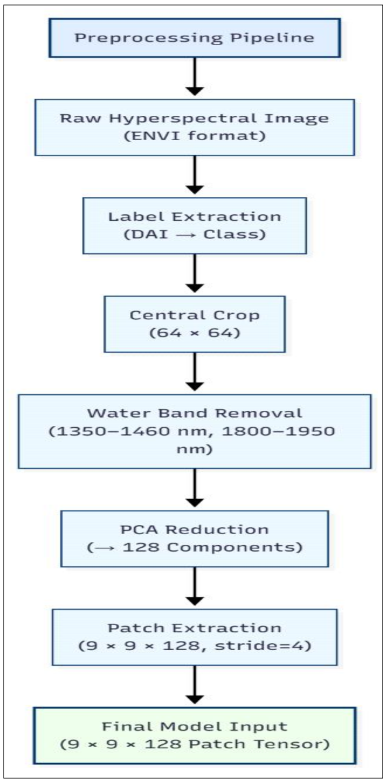

Fig 2: Preprocessing pipeline.

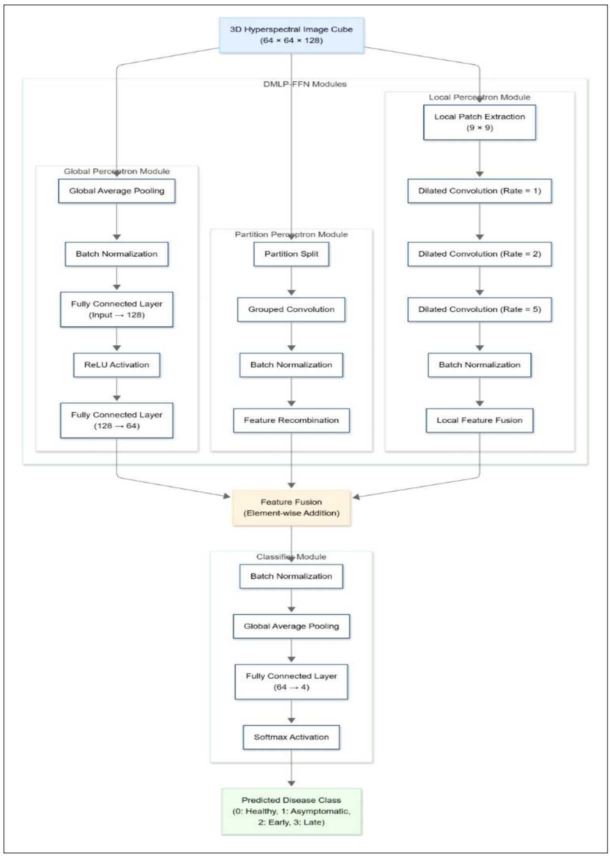

Fig 3: DMLP-FFN model architecture.

Table 2: Local perceptron structure.

Block	Layer type	Input channels	Output filters	Kernel size	Stride	Padding/dilation
Local perceptron	Conv2D (Dilated)	C (In)	C (Out)	3	1	1 (Dilation=1)
	BatchNorm2d	C (Out)	-	-	-	-
	Conv2D (Dilated)	C (In)	C (Out)	3	1	2 (Dilation=2)
	BatchNorm2d	C (Out)	-	-	-	-
	Conv2D (Dilated)	C (In)	C (Out)	3	1	5 (Dilation=5)
	BatchNorm2d	C (Out)	-	-	-	-

Table 3: Performance metrics on test data.

Metric	Value	Support
Accuracy	0.9718	1912
Macro avg	0.9650	1912
Weighted avg	0.9718	1912

Table 4: Classification report summarizing per-class metrics.

Class	Precision	Recall	F1-score	Support
0	0.9455	0.9630	0.9541	324
1	0.9912	0.9573	0.9739	117
2	0.9509	0.9509	0.9509	265
3	0.9817	0.9801	0.9809	1206

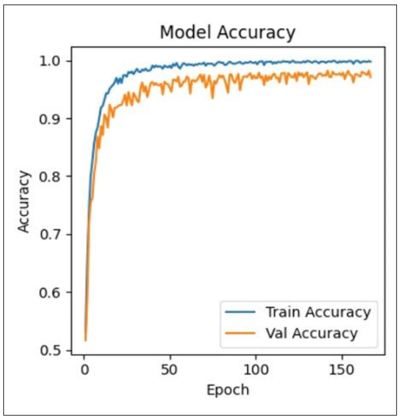

Fig 4: Model accuracy vs epoch.

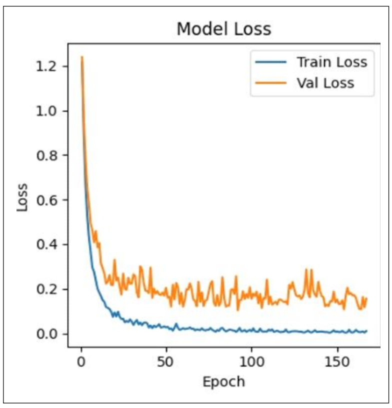

Fig 5: Model loss vs epoch.

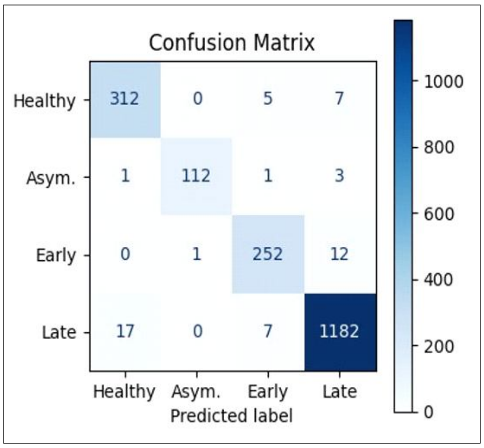

Fig 6: Confusion matrix.

computational performance and architectural complexity. Its high resolution of hyperspectral inputs harmonized shows that it is applicable in real-time agricultural diagnostics especially in situations where edge deployment and low resource settings are essential.

CONCLUSION

The paper presents a computationally efficient hyperspectral deep learning system of multi-stage tomato bacterial leaf spot detection. The DMLPFFN is a compromise in accuracy and efficiency of classification and model with the application of both PCA-based spectral harmonization and multi-branch spectral-spatial feature learning. Local Perceptron incorporates the expansion of spatial receptive fields with the help of the dilated convolutions and the element-wise fusion of the merging facilitates the integration of multiple scales in a compact manner without complex backbone networks. A balanced training policy which includes weighted loss, balanced sampling, cosine annealing scheduling and early termination, provided a stable convergence and high generalization. The model had 98.22% validation accuracy and 97.18% test accuracy, which ensured accurate discrimination in all stages of the disease. The proposed method has similar performance with much reduced computational cost as compared to deeper CNN and hybrid architectures. This is because of its lightweight design and high accuracy that can be used in scalable and real-time precision agriculture system. The further development will be concerned with field validation and cross-crop adaptability.

Disclaimers

The views and conclusions expressed in this article are solely those of the authors and do not necessarily represent the views of their affiliated institutions. The authors are responsible for the accuracy and completeness of the information provided, but do not accept any liability for any direct or indirect losses resulting from the use of this content.

Conflict of interest

The authors declare that there are no conflicts of interest regarding the publication of this article. No funding or sponsorship influenced the design of the study, data collection, analysis, decision to publish, or preparation of the manuscript.

REFERENCES

Acheraiou, L., Medjdoub-Bensaad, F., Guermah, D. and Ramdini, R. (2025). Arthropods diversity associated with the fig tree (*Ficus carica*, Linnaeus, 1753) in Tigzirt (Tizi-Ouzou, Algeria). *Indian Journal of Agricultural Research.* **59(Special Issue):** 92-99. doi: 10.18805/IJARe.AF-956.

Abdulridha, J., Ampatzidis, Y., Qureshi, J. and Roberts, P. (2020). Laboratory and UAV-based identification and classification of tomato yellow leaf curl, bacterial spot and target spot diseases in tomato utilizing hyperspectral imaging and machine learning. *Remote Sensing.* **12(17):** 2732. https://doi.org/10.3390/rs12172732.

Barreto, A., Ispizua Yamati, F.R., Varrelmann, M., Paulus, S. and Mahlein, A.K. (2023). Disease incidence and severity of Cercospora leaf spot in sugar beet assessed by multispectral unmanned aerial images and machine learning. *Plant Disease.* **107(1):** 188-200. https://doi.org/10.1094/PDIS-12-21-2734-RE.

Chen, Y., Lin, Z., Zhao, X., Wang, G. and Gu, Y. (2014). Deep learning-based classification of hyperspectral data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. **7(6):** 2094-2107.

Chang, Y.L., Liu, J.N., Han, C.C. and Chen, Y.N. (2014). Hyperspectral image classification using nearest feature line embedding approach. *IEEE Transactions on Geoscience and Remote Sensing.* **52(1):** 278-287. doi: 10.1109/TGRS.2013.2238635.

Duhan, S., Gulia, P., Gill, N.S., Aziz, A.L., Aljuaid, M., Banshal, S.K. and Bansal, R. (2025). LiteCShuffle: A lightweight and efficient deep learning approach for plant disease classification. *Cogent Food and Agriculture.* **11(1):** 2568197. https://doi.org/10.1080/23311932.2025.2568197.

Guo, Y., Mokany, K., Ong, C., Moghadam, P., Ferrier, S. and Levick, S.R. (2023). Plant species richness prediction from DESIS hyperspectral data: A comparison study on feature extraction procedures and regression models. *ISPRS Journal of Photogrammetry and Remote Sensing.* **196:** 120-133. https://doi.org/10.1016/j.isprsjprs.2022.12.028.

Krishnan, R.S. and Julie, E.G. (2023). Computer aided detection of leaf disease in agriculture using convolution neural network based squeeze and excitation network. *Automatika.* **64(4):** 1038-1053. https://doi.org/10.1080/00051144.2023.2241792.

Mahlein, A.K., Kuska, M.T., Behmann, J., Polder, G. and Walter, A. (2018). Hyperspectral sensors and imaging technologies in phytopathology: State of the art. *Annual Review of Phytopathology.* **56:** 535-558. https://doi.org/10.1146/annurev-phyto-080417-050100.

Mashamba, M.M., Telukdarie, A., Munien, I., Onkonkwo, U. and Vermeulen, A. (2024). Detection of bacterial spot disease on tomato leaves using a convolutional neural network (CNN). *Procedia Computer Science.* **237:** 602-609. https://doi.org/10.1016/j.procs.2024.05.145.

Reddy, K.S. and Kumari, K.P. (2026). Fruit disease detection using AI: A review of classical and deep learning approaches. *Indian Journal of Agricultural Research.* **60(2):** 159-165. doi: 10.18805/IJARe.A-6436.

Reis Pereira, M., Verrelst, J., Tosin, R., Rivera Caicedo, J.P., Tavares, F., Neves dos Santos, F. and Cunha, M. (2024). Plant disease diagnosis based on hyperspectral sensing: Comparative analysis of parametric spectral vegetation indices and nonparametric Gaussian process classification approaches. *Agronomy.* **14(3):** 493.

Shirly, E.A. (2025). MLVI-CNN: A hyperspectral stress detection framework using machine learning-optimized indices and deep learning for precision agriculture. *Frontiers in Plant Science.* **16:** 1631928. https://doi.org/10.3389/fpls.2025.1631928.

Shoaib, M., Shah, B., El-Sappagh, S., Ali, A., Ullah, A., Alenezi, F., Gechev, T., Hussain, T. and Ali, F.A. (2023). An advanced deep learning models-based plant disease detection: A review of recent research. *Frontiers in Plant Science.* **14:** 1158933. https://doi.org/10.3389/fpls.2023.1158933.

Sreedevi, A. and Manike, C. (2023). Development of weighted ensemble transfer learning for tomato leaf disease classification solving low resolution problems. *The Imaging Science Journal.* **71(2):** 161-187. https://doi.org/10.1080/13682199.2023.2178605.

Wei, Z., Jiang, Y. and Liu, H. (2025). Optical imaging combined with artificial intelligence in plant disease detection: A comprehensive review. *Spectroscopy Letters*. https://doi.org/10.1080/00387010.2025.2459246.

Xu, L., Chen, L., Luo, Q., Zhao, S., Huang, J., Wang, K., Yang, Z., Weng, X., Fang, K. and Feng, H. (2026). Leveraging UAV hyperspectral imaging for crop physiology and biochemistry: A comprehensive review of feature extraction and selection methods. *Plant Phenomics.* **8(1):** 100141. https://doi.org/10.1016/j.plaphe.2025.100141.

Xu, Y., Wu, W., Chen, Y., Zhang, T., Tu, K., Hao, Y., Cao, H., Dong, X. and Sun, Q. (2022). Hyperspectral imaging with machine learning for non-destructive classification of *Astragalus membranaceus* var. *mongholicus*, *Astragalus membranaceus* and similar seeds. *Frontiers in Plant Science.* **13:** 1031849. https://doi.org/10.3389/fpls.2022.1031849.

Xie, B., Zhang, Y., Mei, S., Zhang, G., Feng, Y. and Du, Q. (2022). Spectral variation augmented representation for hyperspectral imagery classification with few labeled samples. *IEEE Transactions on Geoscience and Remote Sensing*. **60:** 1-12. doi: 10.1109/TGRS.2022.3220579

Yadav, R., Seth, A. and Dembla, N. (2024). Optimizing crop yield prediction: Data-driven analysis and machine learning modeling using USDA datasets. *Current Agriculture Research Journal.* **12(1):** 272-285.

Zhang, X., Vinatzer, B.A. and Li, S. (2024). Hyperspectral imaging analysis for early detection of tomato bacterial leaf spot disease. *Scientific Reports.* **14:** 27666. https://doi.org/10.1038/s41598-024-78650-6.

Zhou, C., Zheng, L., Meng, K., Zheng, W., Zhang, K. and Shi, Q. (2025). Early detection of tomato leaf spot and wilt diseases based on hyperspectral imaging technology. *Vegetable Research.* **5:** e026. https://doi.org/10.48130/vegres-0025-0010.