International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 13 Issue: 11 | Nov 2026 www.irjet.net p-ISSN: 2395-0072

Evaluating the Impact of Photometric Image Preprocessing on Crop Disease Classification in Real-World Images

Vikram Rawal[1], Utpal Kumar[2], Ujjwal Kumar[3], Sameer Raj[4], Suraj Kumar B P[5]

[1234]Dept. of Computer Science and Engineering, Sir M. Visvesvaraya Institute of Technology, Bangalore, India
[5]Assistant Professor, Dept. of Computer Science and Engineering, Sir M. Visvesvaraya Institute of Technology, Bangalore, India

---

Abstract - *Deep learning models for crop disease classification often achieve high accuracy on images collected under controlled, laboratory-like conditions, but their performance can degrade on real-world images that exhibit variable illumination, contrast, colour, and hue. This paper investigates whether simple, deterministic photometric preprocessing can improve the performance of a transfer-learning-based crop disease classifier evaluated on real-world images. Using the PlantDoc dataset, a single pretrained ResNet-18 model is fine-tuned under an identical architecture, data split, training procedure, and set of hyperparameters across every experimental condition, varying only the preprocessing applied to the input images. A raw-image baseline is compared against four photometric preprocessing conditions brightness, contrast, saturation, and hue adjustment each tested at three magnitudes and with two independent training seeds per configuration, for 26 training runs in total. On the held-out test set, the baseline achieves 58.5% mean accuracy; only a mild contrast adjustment (factor 1.2) modestly exceeds this, at 59.1%, while most other photometric adjustments particularly at higher magnitudes perform at or below the baseline. These findings suggest that the tested fixed photometric transformations did not consistently improve crop disease classification on PlantDoc under the evaluated experimental conditions.*

Key Words: crop disease classification, deep learning, transfer learning, image preprocessing, PlantDoc, ResNet-18

1. INTRODUCTION

Timely and accurate diagnosis of crop disease is central to effective agricultural management, yet manual diagnosis by agricultural extension workers or plant pathologists remains slow and inaccessible in many regions. Advances in computer vision and deep learning have made automated, image-based disease diagnosis an increasingly practical alternative, with convolutional neural networks (CNNs) reporting high classification accuracy when trained and evaluated on curated, laboratory-style benchmark datasets.

This high accuracy does not always transfer to real-world deployment conditions. Studies evaluating plant disease classifiers on PlantDoc have reported lower performance than is commonly achieved on controlled laboratory-style datasets [3]. PlantDoc contains field-oriented images with uncontrolled backgrounds, lighting, and leaf orientations [1]. This gap motivates closer attention to what, specifically, makes real-world images harder for these models to classify.

Photometric variability is a recurring part of that gap. Field-acquired plant images are routinely affected by uneven ambient lighting, differing camera sensors, and time-of-day or weather-driven shifts in brightness, contrast, colour saturation, and hue. Prior work on image preprocessing for plant disease detection has reported that content-adaptive contrast enhancement can measurably improve detection accuracy under uneven field lighting [6], and that preprocessing and augmentation strategies more broadly can improve classifier performance and robustness to class imbalance in disease detection pipelines [8], [9]. This raises a practical question for deployment: can simple, computationally inexpensive, fixed photometric adjustments to the input image brightness, contrast, saturation, or hue meaningfully improve classification performance on real-world images, without changing the model, the dataset, or the training procedure?

This paper investigates that question through a controlled, single-variable experiment. Using the PlantDoc dataset and a single pretrained ResNet-18 model fine-tuned via transfer learning. The model architecture, data split, random seed pool, training procedure, and hyperparameters are held fixed across every experimental condition, varying only the photometric preprocessing applied to the input images. A raw-image baseline is evaluated against brightness, contrast, saturation, and hue adjustments, each at three magnitudes and across two independent training seeds, for 26 training runs in total, with accuracy, macro-precision, macro-recall, and macro-F1-score reported for every condition. The remainder of this paper is organised as follows. Section II reviews related work on transfer learning for crop disease classification and on preprocessing for real-world plant images. Section III describes the PlantDoc dataset. Section IV details the experimental methodology. Section V presents the results, Section VI discusses them, Section VII discusses limitations, and Section VIII concludes.

2. RELATED WORK

2.1 Transfer Learning with ResNet-18 for Crop Disease Classification

Transfer learning from ImageNet-pretrained convolutional networks is widely used for crop disease classification, since labelled disease datasets are typically small relative to general-purpose vision benchmarks. ResNet-18 [2] is a residual convolutional neural network architecture with a relatively lightweight design. Its use in crop disease classification is demonstrated in studies such as [4] and [5]. Fine-tuning a pretrained ResNet-18 has been reported to reach 96% accuracy classifying four corn leaf disease categories [4] and 98.67% accuracy on a 16-class subset of the PlantVillage dataset [5], both using laboratory-style or curated images. These results illustrate the strong performance transfer learning with ResNet-18 can achieve under favourable, lab-style imaging conditions, which motivates testing the same architecture directly on real-world imagery, as this paper does.

2.2 Real-World and PlantDoc-Based Studies

Studies evaluating classifiers specifically on PlantDoc report substantially lower accuracy than equivalent lab-dataset results, reflecting the dataset's uncontrolled backgrounds and lighting. A comparative study of InceptionResNet, VGG-19, and EfficientNetLite on PlantDoc found EfficientNetLite with focal loss and augmentation achieved the best result at 70.71% test accuracy and a mean F1-score of 0.70 [3], while a more recent comparison of classification and object-detection pipelines on PlantDoc reported 78.6% accuracy for EfficientNet-B0 and 75.3% for ResNet50 on the classification task, alongside 63.7% mean Average Precision for a Faster R-CNN detector [7]. Both studies emphasise that performance on PlantDoc remains well below what the same model families achieve on laboratory-style datasets, and both identify dataset class imbalance and visual similarity between disease categories as recurring sources of error, consistent with the moderate absolute accuracy levels obtained in the present study.

2.3 Preprocessing and Augmentation for Disease Classification

Several recent studies report that preprocessing and augmentation can improve plant disease classification performance. A factorial study on calamansi leaf disease detection found that Contrast Limited Adaptive Histogram Equalization (CLAHE) a content-adaptive, locally-applied contrast enhancement improved model accuracy by 4.8% under uneven field lighting, alongside gains in precision and recall [6]. Separately, a plant disease detection pipeline study found that combining preprocessing with class-imbalance handling techniques improved classifier performance over an unprocessed baseline [8], and a review of data augmentation methods for crop disease recognition found that augmentation strategies generally improve generalisation and robustness when applied during training [9]. These studies differ from the present work in two respects: their enhancement techniques are either content-adaptive (CLAHE) or applied stochastically during training (augmentation), rather than the single, fixed, deterministic photometric shift applied uniformly to every image within a condition that this paper evaluates. These studies evaluate different interventions and experimental settings from the fixed photometric transformations investigated here. This paper's controlled, single-variable design is intended to test specifically whether fixed photometric shifts alone, independent of any other change, measurably affect real-world classification performance.

3. DATASET

This study uses a subset of PlantDoc [1], a dataset of 2,552 images spanning 13 plant species and up to 27 disease and healthy classes, constructed specifically to reflect the complex backgrounds, variable illumination, and uncontrolled camera conditions typical of real-world plant disease images, in contrast to laboratory-style datasets. Images were collected from internet sources and manually annotated, with class sizes that are known to be imbalanced [1], [7]. PlantDoc provides a predefined train/test split, used directly in this study: the provided training folder is further split into training and validation subsets (Section IV-C), and the provided test folder is held out and used only for final evaluation.

4. METHODOLOGY

4.1 Model Architecture and Transfer Learning

A single ResNet-18 model [2], pretrained on ImageNet, is used for every experimental condition. The final fully connected layer is replaced with a new linear layer matching the number of PlantDoc classes, and the entire network is fine-tuned end-to-end. The model architecture, its initialisation procedure, and the fine-tuning approach are identical across every condition reported in this paper; only the preprocessing applied to the input images (Section IV-B) differs between conditions.

4.2 Preprocessing Conditions

Every image, regardless of condition, is resized to 224×224 pixels and normalised using the ImageNet channel mean and standard deviation, as required by the ImageNet-pretrained weights. Beyond this required step, the baseline condition applies no further preprocessing. Four photometric preprocessing conditions are additionally evaluated, each applying a single deterministic transform to every image in the condition: brightness, contrast, and saturation, each tested at multiplicative factors of 1.2, 1.5, and 1.8 (1.0 =

unchanged), and hue, tested at additive shifts of 0.05, 0.10, and 0.15 (valid range [-0.5, 0.5]). Unlike the stochastic, training-time augmentation and content-adaptive enhancement discussed in Section II-C, these transforms are applied deterministically and identically to every image within a condition, modelling a persistent, uniform shift in imaging conditions rather than introducing training-time random variability or locally-adaptive enhancement.



Fig -1: Sample PlantDoc image under the baseline condition and under each photometric preprocessing condition at its largest tested magnitude (brightness, contrast, saturation, hue).

4.3 Experimental Protocol

The PlantDoc training folder is split into training (85%) and validation (15%) subsets using stratified sampling with a fixed random seed (random_state = 42), identical across every one of the 26 runs. The resulting split is identical across all experimental conditions and training runs. Each condition is trained using two training seeds, 1 and 2, which control model initialisation and training stochasticity. The PlantDoc test folder is held out entirely and used only once per run, after training is complete, for final evaluation. All runs share identical hyperparameters: batch size 32, the Adam optimiser with a learning rate of 1×10^{-4}, up to 10 training epochs, and 224×224 input resolution. Within each run, the model checkpoint with the highest validation accuracy is selected for final test-set evaluation. Each of the 13 conditions (baseline, plus four transforms × three magnitudes) is trained with two independent random seeds controlling model initialisation and training stochasticity, for 26 training runs in total, allowing preprocessing effects to be distinguished from ordinary run-to-run training variance. The implementation uses PyTorch, and evaluation metrics are computed using scikit-learn.

4.4 Evaluation Metrics

For every run, four metrics are computed on the held-out test set: overall accuracy, macro-averaged precision, macro-averaged recall, and macro-averaged F1-score. Macro averaging computes each metric independently per class and reports the unweighted mean across classes, so that performance on PlantDoc's less-represented classes contributes equally to the reported score, rather than being dominated by the dataset's more frequent classes.

5. RESULTS

Table I reports accuracy, macro-precision, macro-recall, and macro-F1-score for the baseline and each of the twelve transform × magnitude conditions, averaged across the two training seeds. Chart 1 shows the same accuracy comparison as a bar chart, with error bars spanning the range between the two seeds and a dashed line marking the baseline's mean accuracy.

Table -1: MEAN TEST-SET PERFORMANCE BY CONDITION (%, MEAN OF 2 SEEDS)

Condition	Magnitude	Accuracy	Precision	Recall	F1
Baseline	—	58.5	59.6	58.9	57.1
Brightness	1.2	57.6	58.1	57.6	56.5
Brightness	1.5	56.1	57.6	56.2	55.6
Brightness	1.8	55.1	55.5	55.1	53.8
Contrast	1.2	59.1	62.6	59.1	58.7
Contrast	1.5	55.1	56.7	54.7	53.9
Contrast	1.8	54.4	54.4	54.4	53.2
Saturation	1.2	57.0	59.0	57.4	56.6
Saturation	1.5	58.1	59.8	58.2	57.5
Saturation	1.8	56.8	57.8	56.8	55.8
Hue	0.05	55.9	57.0	55.8	54.4
Hue	0.10	55.5	55.3	55.6	53.8
Hue	0.15	55.9	58.0	56.1	54.8

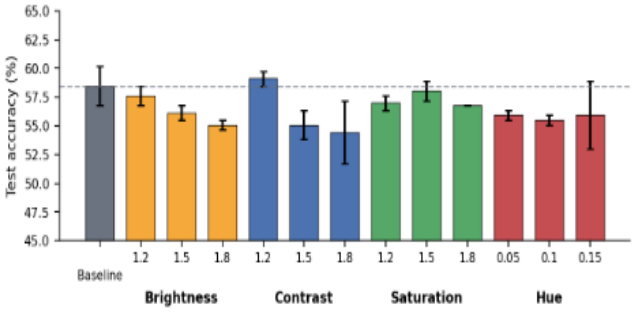

Chart -1: Test accuracy by preprocessing condition, averaged over two training seeds; error bars show the range between seeds; the dashed line marks baseline accuracy

The baseline condition achieves 58.5% mean accuracy. Of the twelve preprocessing conditions tested, only mild contrast adjustment (factor 1.2) exceeds this, at 59.1% a margin of 0.6 percentage points. Every other condition performs at or below the baseline. The decline is most pronounced at higher magnitudes: brightness and contrast both fall from their mildest to their strongest tested magnitude (57.6% to 55.1% for brightness; 59.1% to

54.4% for contrast), and the strongest contrast condition (1.8) records the lowest mean accuracy of all thirteen conditions. Saturation shows a smaller and non-monotonic pattern, peaking at magnitude 1.5 (58.1%) rather than at the mildest magnitude. Hue adjustment produces accuracy between 55.5% and 55.9% across all three magnitudes tested, consistently below baseline, with magnitude 0.15 showing the widest spread between seeds (52.9% to 58.7%) of any condition.

6. DISCUSSION

These results do not support the hypothesis that simple, fixed photometric preprocessing reliably improves real-world crop disease classification performance in this setting. Only mild contrast adjustment nominally exceeded the baseline. Every other condition performed at or below the unprocessed baseline, and the decline generally grew larger as the magnitude of the photometric shift increased, most clearly for brightness and contrast.

One possible explanation is that global photometric transformations alter image statistics relative to those encountered during pretraining. However, this experiment does not directly measure distribution shift or establish that it caused the observed performance changes. The results also differ from the positive effect reported for CLAHE in calamansi disease detection [6]. CLAHE enhances contrast locally and adaptively, whereas the present study applies fixed global photometric transformations. Differences in datasets, tasks, models, and experimental procedures may also contribute to the different outcomes. Similarly, the improvements reported when preprocessing is combined with class-imbalance handling [8] or when augmentation is applied stochastically during training [9] involve additional interventions beyond a fixed, deterministic photometric shift, which this paper's design deliberately isolated as the sole variable.

Taken together, these results suggest that the tested fixed global photometric transformations did not consistently improve classification performance in this experimental setting. Whether content-adaptive enhancement or combinations with other interventions would improve performance remains an open question for future evaluation.

7. LIMITATIONS

Several limitations should be considered when interpreting these results. First, the PlantDoc test set is relatively small, which limits the precision and generalisability of the reported performance estimates. Second, only a single lightweight architecture (ResNet-18) is evaluated; findings may not generalise to deeper or differently structured architectures, including the EfficientNet and detection-based architectures that performed comparatively well on PlantDoc in related work [3], [7]. Third, the photometric magnitudes tested were chosen as moderate, round-number adjustments rather than empirically calibrated to the illumination and colour variability actually present in PlantDoc or in a specific target deployment environment. Fourth, two training seeds per condition provide only a coarse estimate of run-to-run variance. Fifth, each photometric condition is evaluated in isolation; combinations of multiple simultaneous adjustments, or content-adaptive techniques such as CLAHE, were not tested and may behave differently, as suggested by the contrast in outcomes with related work in Section VI.

8. CONCLUSION AND FUTURE WORK

This paper presented a controlled, single-variable experimental design for evaluating whether simple, fixed photometric preprocessing brightness, contrast, saturation, and hue adjustment improves the performance of a transfer-learning-based crop disease classifier on real-world PlantDoc images, holding model architecture, data split, training procedure, and hyperparameters fixed across all conditions. Across 26 training runs, no tested preprocessing condition produced a consistent improvement over the raw-image baseline. Only mild contrast adjustment (factor 1.2) achieved a higher mean accuracy, increasing from 58.5% to 59.1%, while most other conditions performed at or below baseline. These findings suggest that, at least for a single lightweight architecture evaluated on PlantDoc, simple uniform photometric adjustments are not a reliable route to improved real-world classification accuracy, in contrast to the gains reported elsewhere for content-adaptive enhancement and training-time augmentation.

Future work could test content-adaptive techniques such as CLAHE directly within this same controlled design, evaluate combinations of multiple photometric adjustments rather than each in isolation, calibrate preprocessing magnitudes against measured illumination and colour statistics of real PlantDoc images, and extend the comparison to additional lightweight architectures and real-world datasets.

REFERENCES

[1] D. Singh, N. Jain, P. Jain, P. Kayal, S. Kumawat, and N. Batra, "PlantDoc: A dataset for visual plant disease detection," in Proc. 7th ACM IKDD CoDS and 25th COMAD, 2020, pp. 249–253.

[2] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778.

[3] A. Sethi, S. Gopal V, and J. Ranjani J, "Plant disease identification on real-world data using deep learning: A comparative study," in Proc. 2021 4th Int. Conf. on Computing and Communications Technologies (ICCCT), 2021, pp. 223–229.

[4] R. Singh, N. Sharma, and R. Gupta, "Classification and detection of corn leaf disease using ResNet 18 transfer learning model," in Proc. 2023 8th Int. Conf. on Communication and Electronics Systems (ICCES), 2023, pp. 1655–1661.

[5] S. Sailau, M. Z. Mohsin, R. H. Ali, and U. R. Syed, "High-accuracy plant disease detection: A ResNet18 transfer learning approach for multi-class classification," in Proc. 2025 27th Int. Multitopic Conf. (INMIC), 2025.

[6] H. Y. Bangalisan, J. A. Sneha, and V. Sivakumar, "Optimized image-based disease detection: Evaluating the effects of preprocessing, annotation methods, and dataset variability," in Proc. 2025 8th World Engineering Conf. on Contemporary Technologies (WECON), 2025.

[7] A. Saini, P. Madaan, P. Kumari, S. Panwar, and S. Kaushik, "Multi-crop plant disease detection using PlantDoc," in Proc. 2025 4th Int. Conf. on Automation, Computing and Renewable Systems (ICACRS), 2025, pp. 1657–1664.

[8] M. O. Ojo and A. Zahid, "Improving deep learning classifiers performance via preprocessing and class imbalance approaches in a plant disease detection pipeline," Agronomy, vol. 13, no. 3, p. 887, 2023.

[9] S. Lee and S. Lee, "Efficient data augmentation methods for crop disease recognition in sustainable environmental systems," Big Data and Cognitive Computing, vol. 9, no. 1, p. 8, 2025.