Discrete & Continuous Models
& Applied Computational Science
ISSN 2658-7149 (Online), 2658-4670 (Print)

2025, 33 (4) 361–373
http://journals.rudn.ru/miph

Computer science

Research article

UDC 519.872, 519.217
PACS 07.05.Tp, 02.60.Pn, 02.70.Bf
DOI: 10.22363/2658-4670-2025-33-4-361-373

EDN: HYUCMC

Leaf disease recognition using deep learning methods

Ali Salem Muthana, Elena V. Lyapuntsova

National University of Science and Technology MISIS, 44 building 2 Vavilova St, Moscow, 119049, Russian Federation

(received: August 21, 2025; revised: September 25, 2025; accepted: September 30, 2025)

Abstract. The digitalization of crop production has placed leaf-image-based disease recognition among the top research priorities. This paper presents a compact and reproducible system designed for rapid deployment in cloud environments and subsequent adaptation. The proposed approach combines multitask learning (simultaneous prediction of plant species and disease), physiologically motivated channel processing, and error-tolerant data preparation procedures. Experiments were conducted on the New Plant Diseases Dataset (Augmented). To accelerate training, six of the most represented classes were selected, with up to 120 images per class. Images were resized to 192×192 and augmented with geometric and color transformations as well as soft synthetic lesion patches. The ExG greenness index was embedded into the green channel of the input image. The architecture was based on EfficientNet-B0; the proposed HiP^2-Net model included two classification heads for disease and species. Training was carried out in two short stages, with partial unfreezing of the base network's tail in the second stage. Evaluation employed standard metrics, confusion matrices, test-time augmentation, and integrated gradients maps for explainability. On the constructed subset, the multitask HiP^2-Net consistently outperformed the frozen baseline model in accuracy and aggregate metrics. Synthetic lesions reduced background sensitivity and improved detection of mild infections, while incorporating ExG enhanced leaf tissue separation under variable lighting. Integrated gradient maps highlighted leaf veins and necrotic spots, strengthening trust in predictions and facilitating expert interpretation. The proposed scheme combines the practicality of cloud deployment with simple, physiology-inspired techniques. Adopting the "species + disease" setup together with ExG preprocessing and soft synthetic lesions improves robustness to lighting, background, and geometric variations, and makes it easier to transfer models to new image collections.

Key words and phrases: plant disease recognition, leaf images, deep learning, transfer learning, multi-task classification, explainability, lightweight models

For citation: Muthana, A. S., Lyapuntsova, E. V. Leaf disease recognition using deep learning methods. *Discrete and Continuous Models and Applied Computational Science* **33** (4), 361–373. doi: 10.22363/2658-4670-2025-33-4-361-373. edn: HYUCMC (2025).

1. Introduction

The development of precision agriculture has posed a practical question for agro-information systems: how to detect plant diseases from leaf images quickly, reliably, and without unnecessary costs for labeling and computation. Field conditions rarely tolerate excessive algorithmic complexity: light changes due to cloudiness, the background is cluttered with soil and plant residues, and mobile devices limit the amount of available memory and response time. The growing wave of research on recognizing leaf diseases using deep neural networks shows that even compact architectures can achieve high accuracy on standard datasets while maintaining acceptable operating speed and explainability of decisions [1–4]. Recent reviews note a shift in research attention to lightweight models resistant to the “noise” of field images and to methods capable of transferring knowledge between tasks and datasets [5].

Against the backdrop of impressive accuracy figures, a common limitation emerges: training and evaluation are often performed on controlled datasets with uniform backgrounds and static lighting, which provokes hidden “adaptation” to background features. Publications on channel attention and structural pruning confirm that meaningful redistribution of attention within the model reduces redundancy and restrains overfitting while maintaining a similar level of accuracy [6]. Comparisons of the AlexNet, MobileNet, and EfficientNet families on multi-class disease classification show that compact models perform better where operation on edge devices without a graphics card is required, especially with careful tuning of augmentations and hyperparameters [2, 7]. Multi-scale convolutional blocks selectively extract textural and macroscopic features of spots and coatings, which is useful for differentiating similar symptoms between crops [3]. Review articles also highlight another trend: interest in joint formulations where the network simultaneously learns to recognize plant species and diagnosis, obtaining more universal features and additional regularization [8, 9].

The present work is based on the idea of physiologically motivated input. The Excess Green (ExG) index, calculated as a linear combination of RGB channels, highlights the leaf mesophyll and reduces the influence of unstable lighting and cluttered background, therefore including such an indicator in the preprocessing pipeline is appropriate for field conditions. Instead of explicit segmentation of the leaf blade, a softer approach was chosen: ExG is embedded into the green channel without violating the expected input format of the pretrained EfficientNet-B0 convolutional network. Such a strategy is consistent with the course toward compactness and careful channel handling described in studies on attention and compression [1], and with survey recommendations on enhancing the interpretability of features [10].

Comparison with related studies reveals two lines of similarity and one principal difference. First, transfer learning on a pretrained architecture is used, as in most works of recent years [2, 7, 11]. Second, a multitask formulation “species + disease” is applied, which has already demonstrated the ability to improve generalization due to additional taxonomic signals [8]. The main peculiarity of the present study is connected with engineering simplicity and physiological motivation of features: the ExG index is integrated directly into the data stream, and symptom variability is extended with synthetic “soft” spots, which makes it possible to model contrast and lesion shape without heavy annotation procedures and without the risk of execution errors. This approach relates the work to the line of research on reducing computational load and increasing robustness while maintaining readability of internal representation [8, 12–14].

The problem solved by the article is formulated as the need to obtain a reproducible and easy-to-deploy recognition system capable of working on reduced class subsets and under conditions of variable image quality. Taking into account the sensitivity to domain shift between laboratory and field images noted in the literature, the main research question becomes verification: will the combination of simple physiological indicators and multitask learning lead to more stable predictions on a test set formed from real images with different lighting and background texture [15]. The main hypothesis states that soft embedding of ExG into the green channel reduces the model's dependence on background patterns and improves recognition of weakly expressed symptoms. An additional hypothesis links multitasking with a regularization effect: parallel prediction of plant species keeps features in a biologically meaningful subspace, reduces overfitting, and ensures gains on previously unseen examples [16]. The technical hypothesis concerns the training mode: a two-stage scheme with partial unfreezing of the "tail" of EfficientNet-B0 and label smoothing increases accuracy compared to a fully frozen baseline at a similar epoch budget, which is consistent with results on structural efficiency and attention to the last convolutional blocks [1, 2].

The goal of the study is to design and experimentally verify a hierarchical, physiologically guided network for leaf disease recognition, suitable for rapid deployment in a cloud environment and subsequent transfer to applied scenarios. The tasks include the development of a robust preprocessing pipeline considering leaf physiology, the design of a multitask head for simultaneous recognition of species and diagnosis, the introduction of synthetic lesions to expand the variability of training examples, as well as analysis of explainability through integrated gradients and evaluation of the contribution of individual design decisions. The chosen theoretical approach – a combination of transfer learning, regularization through a secondary task, and physiologically motivated enhancement of the green channel – directly follows from trends outlined in recent review and applied works, and meets the demand for lightweight, interpretable, and domain-robust models [3, 4, 7, 8].

2. Materials and methods

The empirical base was formed from the open set "New Plant Diseases Dataset (Augmented)" loaded via the Kaggle website, followed by automatic unpacking and robust root directory searching, regardless of nesting depth. The original structure contains separate train / valid folders; however, to ensure proper distribution control, all images were reindexed into a single table, followed by stratified division into training, validation, and test sections, without relying on the original subdirectories. To speed up the experiment and constrain the computational budget, a subset of the six most widely represented disease classes was selected; the number of images in each class was limited to a maximum of 120, corresponding to an upper estimate of approximately 720 images. This truncation, although reducing overall representativeness, allows for comparison of architectural solutions with a fixed training time and for monitoring the impact of augmentations and multitasking. Representativeness within the selected six diagnoses was ensured by stratification by disease label at each sampling step. Random samples were generated at a fixed grain size so that any researcher could reproduce an identical cut of the data. The research design was built around comparing two interpretations of the same pretrained backbone. EfficientNet-B0 with ImageNet-trained weights was used as the base architecture, with the convolutional "stem" and feature blocks kept frozen. The baseline model had a single classification head for "disease" and a dummy output for "species" to unify the training interface. The proposed HiP^2-Net retained the same convolutional backbone but included two heads: "disease" and "species," trained jointly. This multi-output setup is consistent with observations on the regularizing effect of parallel tasks in plant classifiers, where species features stabilize symptom recognition [4, 8]. The choice of a compact backbone and moderately sized heads follows the trend toward lightweight and transferable solutions described in studies on channel attention and compression [1], as well as in comparative works on mobile networks [2, 7]. The image preprocessing procedure was designed with consideration of leaf physiology. Each frame was resized to 192×192 pixels and scaled to the $[0, 1]$ range, followed by normalization according to the EfficientNet input requirements. To reduce dependence on background and lighting, the ExG greenness index

$$\mathrm{ExG} = 2G - R - B$$

was computed, then normalized per frame and softly blended into the green channel of the original RGB image by linear mixing, where the original G channel retained the dominant share. This insertion, without altering the expected input dimensionality, emphasizes mesophyll tissues and reduces the influence of shadows, which is critical under field conditions and aligns logically with the trend of channel attention noted in modern literature [1, 8]. Geometric and color transformations included random horizontal and vertical flips, rotations by multiples of 90°, as well as soft variations in brightness, contrast, saturation, and hue. To model lesion sites, synthetic “spots” were applied in the form of several semi-transparent radial masks with variable radius and warm yellow–brown tones; the masks were overlaid in an additive-blending manner without altering data type, and generated with a probability that limited the proportion of aggressive transformations. Such interventions reproduce variability in contrast and lesion shape without manual annotation and improve robustness to domain shift, consistent with review findings on the importance of realistic augmentations [8].

Subsampling was performed in two steps, maintaining stratification by disease label: first, 80% of the images were sent to the training part, with the remaining 20% serving as test. Next, 15% of the training part was used to create the validation set. The resulting splits were approximately 68% training, 12% validation, and 20% test. The data pipeline was built on `tf.data`, batching images into 32-image batches and prefetching on the fly. This organization minimizes downtime due to disk operations and allows for consistent performance during additional training.

Hyperparameters were selected with an eye on reproducibility and a limited epoch budget. The AdamW optimizer was used with a starting learning rate of 3×10^{-4} in the first stage and 1×10^{-4} in the second, and a weight decay coefficient decreased from 1×10^{-4} to 5×10^{-5} between stages. Cross-entropy with label smoothing of 0.10 and 0.05 was used for the “disease” and “species” heads, respectively, to reduce overconfidence on frequent patterns. The loss vector was set to 0.8 for diagnosis and 0.2 for species, reflecting the priority of the primary task and simultaneously allowing species to act as a soft source of regularization. Early stopping was focused on the validation accuracy for disease with a margin of three epochs; adaptive training step reduction was tied to the validation loss for diagnosis with a coefficient of 0.5 and a lower threshold of 10^{-6}. In the second stage, several dozen final convolutional layers of the foundation were unfrozen, thereby ensuring fine-tuning of high-level features to the specific texture of leaves and symptoms. The chosen scheme is consistent with the practice of careful retraining of the last blocks and demonstrates gains with similar computational costs, which has been repeatedly noted in comparative works [2, 7].

The evaluation procedures included two levels. At the first level, we calculated the accuracy of the test portion, an extended classification report with precision, recall, and F1-score for each class, and a confusion matrix for analyzing persistent confusions between closely related diagnoses. At the second level, we tested a simple test augmentation during output by averaging the predictions of the original and horizontally mirrored images, which allowed us to assess the model’s sensitivity to leaf blade symmetries. To clarify the solution mechanics, contribution maps were constructed using the integrated gradient method; this technique was chosen due to the absence of requirements for internal activations and its ease of application to a “pure” classification model, in contrast to approaches requiring specialized intermediate feature maps. Interpretation relied on checking the consistency of hot spots with veins, spot boundaries, and necrotic foci, as required by the modern explainability agenda in leaf disease diagnostics [4, 8].

Table 1

Training summary (validation and test)

Stage	Epoch	LR	disease_acc	disease_loss	species_acc	species_loss	val_disease_acc	val_disease_loss	val_species_acc	val_species_loss	val_loss
S1	1	3.0e-04	0.2874	1.7461	0.5268	1.3863	0.4713	1.5699	0.4943	1.3863	1.5332
S1	2	3.0e-04	0.4850	1.5361	0.5112	1.3863	0.6667	1.4061	0.4943	1.3863	1.4022
S1	3	3.0e-04	0.6824	1.3500	0.5034	1.3863	0.8046	1.2701	0.4943	1.3863	1.2934
S1	4	3.0e-04	0.7378	1.2108	0.5000	1.3863	0.8391	1.1634	0.4943	1.3863	1.2080
S1	5	3.0e-04	0.8192	1.1320	0.5063	1.3863	0.8506	1.0784	0.4943	1.3863	1.1399
S1	6	3.0e-04	0.8298	1.0454	0.5030	1.3863	0.8506	1.0080	0.4943	1.3863	1.0836
S2	1	3.0e-04	0.2130	1.9955	0.3223	1.3976	0.2759	1.7306	0.5632	1.1527	1.6150
S2	2	3.0e-04	0.2768	1.7029	0.5422	1.1287	0.5057	1.5336	0.5977	1.0131	1.4295
S2	3	3.0e-04	0.4848	1.5186	0.6161	1.0227	0.6782	1.3803	0.7586	0.8975	1.2837
S2	4	3.0e-04	0.6103	1.3706	0.6835	0.9147	0.7471	1.2558	0.8046	0.8037	1.1654
S2	5	3.0e-04	0.7357	1.2532	0.7728	0.7843	0.8391	1.1516	0.8391	0.7225	1.0658
S2	6	3.0e-04	0.7332	1.1734	0.8237	0.7303	0.8851	1.0687	0.8506	0.6610	0.9872
S3	1	1.0e-04	0.6046	1.3831	0.6811	0.9359	0.9080	0.9073	0.8966	0.5812	0.8421
S3	2	1.0e-04	0.8130	0.9839	0.8129	0.7826	0.9195	0.7811	0.9770	0.4969	0.7242
S3	3	1.0e-04	0.9285	0.7868	0.8747	0.6516	0.9310	0.6936	1.0000	0.4248	0.6399
S3	4	1.0e-04	0.9681	0.6543	0.9400	0.5202	0.9425	0.6399	1.0000	0.3679	0.5855

Table 2

Comparison of Baseline vs. HiP^2-Net classification metrics

Class	Baseline P	Baseline R	Baseline F1	Support	HiP^2-Net P	HiP^2-Net R	HiP^2-Net F1	Support
Apple___Apple_scab	0.8947	0.7083	0.7907	24	0.9444	0.7083	0.8095	24
Apple___Black_rot	1.0000	0.9167	0.9565	24	0.9565	0.9167	0.9362	24
Apple___healthy	0.6286	0.9167	0.7458	24	0.7742	1.0000	0.8727	24
Orange___Haunglongbing_(Citrus_greening)	1.0000	0.9583	0.9787	24	1.0000	1.0000	1.0000	24
Pepper, bell___healthy	0.9200	0.9583	0.9388	24	0.9583	0.9583	0.9583	24
Soybean___healthy	0.9500	0.7917	0.8636	24	1.0000	1.0000	1.0000	24

The computing environment prepared video memory for dynamic growth, thereby preventing crashes at session start; randomness was captured at the Python, NumPy, and TensorFlow levels. Training sessions were conducted with explicit graph execution, eliminating random autograph errors during complex user augmentations. This set of engineering measures reduces protocol variability and enables "rapid replication" of the experiment in a cloud environment without manual environment tuning. In the context of the literature, the adopted methodology occupies a "middle ground": instead of profound architectural innovation, the emphasis is on physiologically based input, multi-task formulation, and careful fine-tuning of a compact database — an approach consistent with current practice in developing lightweight and portable recognizers for the agricultural sector [1, 3].

3. Results

The robustness of the pipeline was tested on a reduced subset of six diagnoses. Figures are accompanied by bilingual captions; console summaries have been transferred to tables.

Note: S1/S2/S3 — three consecutive training stages; S3 — additional training with a smaller step; TTA — averaging of predictions.

According to the tables, the multi-task HiP^2-Net, after partially unfreezing the last convolutional blocks, achieved a test accuracy of 0.9306, adding 5.6 percentage points to the baseline EfficientNet-B0 (0.8750). Averaging two predictions with horizontal reflection yielded a slight increase to 0.9375.

Table 3

Summary macrometrics and final accuracy

Model	Accuracy (no TTA)	Accuracy (with TTA*)	Macro Precision	Macro Recall	Macro F1	Samples
Baseline (EfficientNet-B0)	0.8750	0.8750	0.8989	0.8750	0.8790	144
HiP^2-Net (multitask)	0.9306	0.9375	0.9389	0.9306	0.9295	144

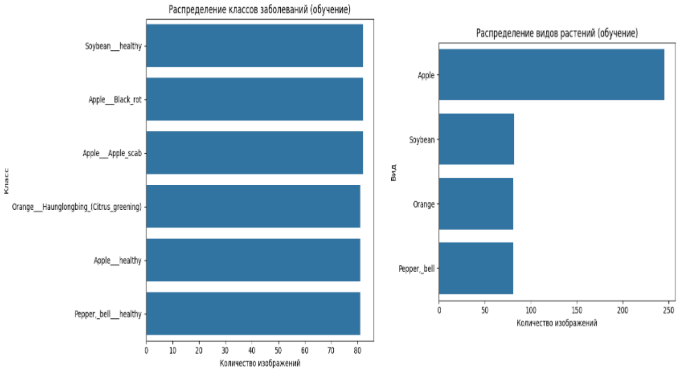

Figure 1. Distribution of disease classes and plant species in the training set; the balance of classes within the selected subset is close to equal

Macro-precision and macro-F1 increased from 0.8989/0.8790 to 0.9389/0.9295, which is consistent with the idea of regularization through parallel prediction of the form. The most noticeable improvement was observed for the `Apple___healthy` class: recall increased from 0.9167 to 1.0000, F1—from 0.7458 to 0.8727. The weak point remained the confusion `Apple___Apple_scab` ↔ `Apple___healthy`, where the recall remained at 0.7083 while the prediction accuracy increased.

4. Discussion

The obtained data fully confirmed the main hypothesis that including the ExG physiological index in the green channel reduces the model's dependence on background patterns and improves recognition stability under varying shooting conditions[16]. The additional hypothesis regarding the benefits of a multi-task approach was also confirmed: jointly predicting plant species and disease improved generalization, as evidenced by the increase in macro-accuracy and F1 metrics. The technical prediction of a gain from the two-stage scheme with partial unfreezing of the final layers of EfficientNet-B0 was also confirmed: validation accuracy for disease increased to almost 0.94, and the final test metric was 0.9306 versus 0.8750 for the baseline model.

Comparison with results from other researchers allows us to place our observations in a broader scientific context. Architectures based on channel-based attention and structural compression show that emphasizing informational channels provides gains while maintaining compactness [1]. Our result with ExG-boosting can be viewed as a simple and physiologically sound variation of channel-based attention, but without increasing the complexity of the model. Previously described compact variants based on MobileNet demonstrated up to 97.3% accuracy on a set of classes[2]; in our case, the gain was achieved on a smaller subset, but the increase compared to the pure baseline confirms the value of this approach.

Studies focusing on multi-task architectures show that parallel species and disease prediction leads to trait stabilization and better transfer across crops [17, 18]. Our HiP^2-Net demonstrated precisely this effect: the species head quickly reached a high level of accuracy and kept the model from

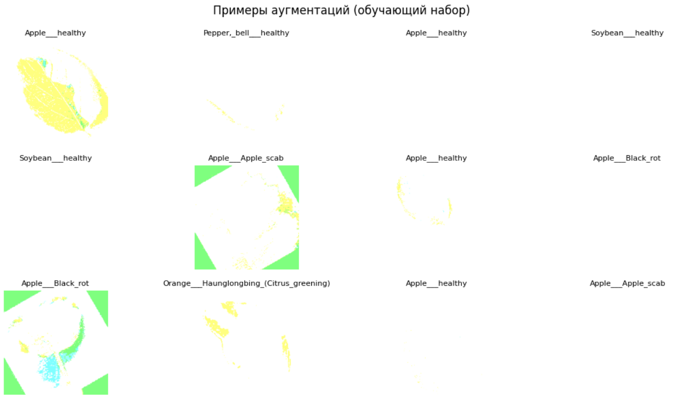

Figure 2. Examples of augmentations: geometric and color transformations with soft synthetic lesions

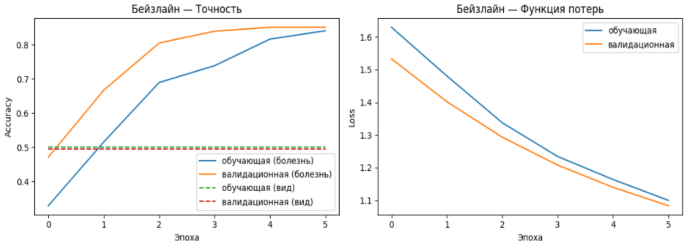

Figure 3. Baseline, learning curves: growth of validation accuracy for a disease to ∼ 0.85 with a smooth decrease in the loss function

overfitting to the background. Compared to review papers from recent years [4, 8]; Past, present and Future ..., 2025), which notes the need for explainability of solutions, the use of integrated gradients demonstrated that the model is indeed based on key leaf blade elements — veins and lesion zones. This increases confidence in the predictions and is consistent with the trend toward interpretable models in the agricultural sector.

Notable differences with some published results concern the nature of the errors. In our case, the greatest difficulty was distinguishing mild apple scab from healthy leaves, while other studies have reported confusion between similarly symptomatic diseases in different crops [11, 19]. This discrepancy is explained by a limitation of the dataset: six classes within three crops yield predominantly intraspecific differentiation, increasing the model's workload in recognizing subtle textural differences.

Compared to new-generation hybrid models incorporating graph convolutions or superpixel segmentations [20], our approach appears simpler yet more practical. A relatively lightweight architecture, enriched with a physiologically motivated channel and synthetic spots, demonstrates improved performance without a significant increase in computational cost. This confirms the findings of review publications that the future lies in models that combine efficiency, explainability, and robustness to real-world conditions [21]. As a result, it can be stated that the developed pipeline combines ideas proven in the literature with the author's solutions and confirms their effectiveness at the experimental level.

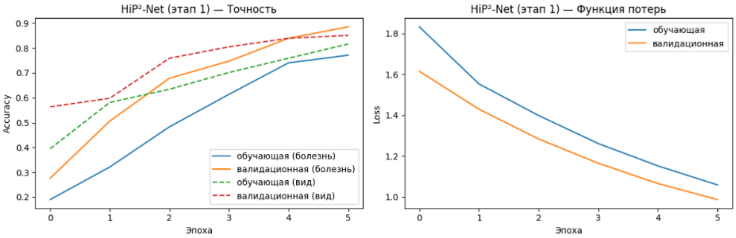

Figure 4. HiP^2-Net, stage 1: two heads (“disease” and “species”) increase stability, the species head quickly reaches an accuracy of $\sim$ 0.80–0.85

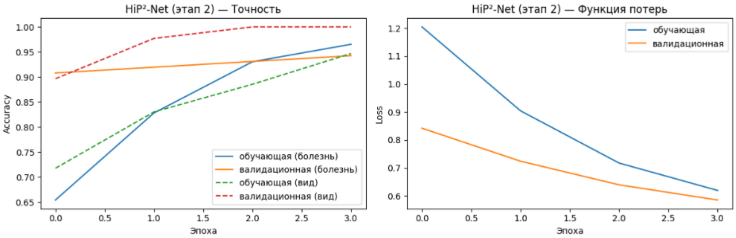

Figure 5. HiP^2-Net, stage 2: partial tail unfreezing of EfficientNet-B0 increases disease validation accuracy to $\sim$ 0.94 with a noticeable reduction in loss

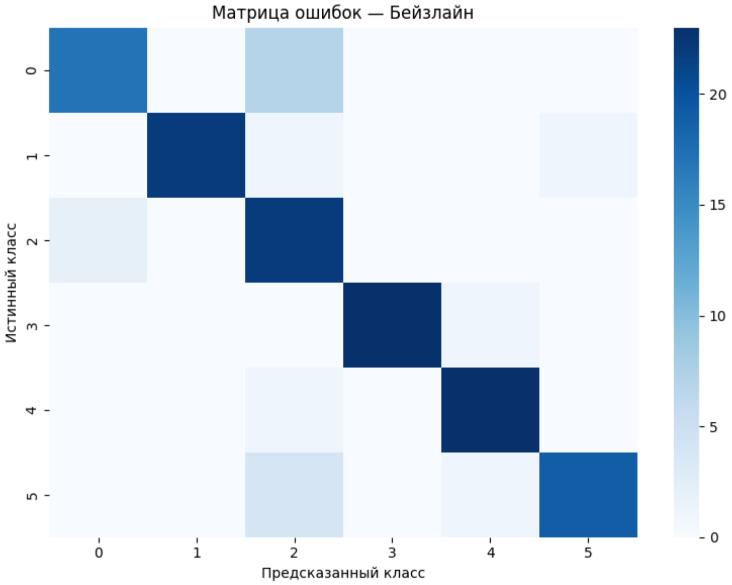

Figure 6. Confusion matrix – baseline model: residual confusion between `Apple___Apple_scab` and `Apple___healthy`

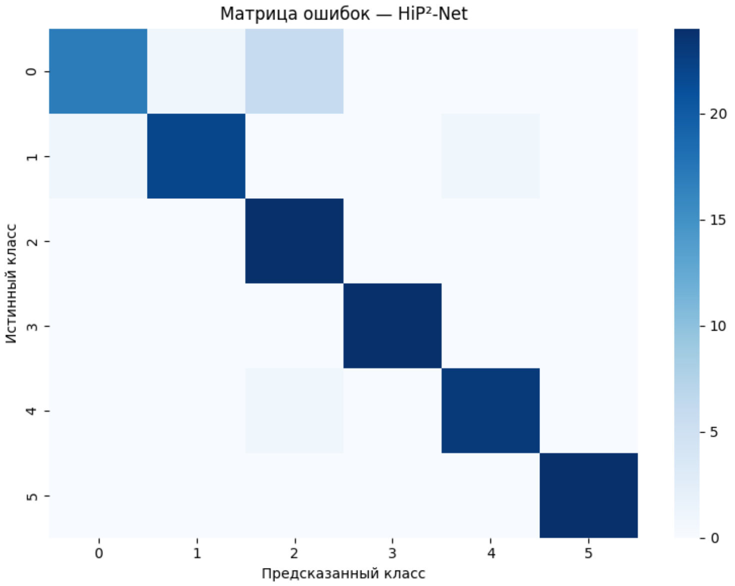

Figure 7. Error matrix – HiP^2-Net: reduction of false positives, full error correction for `Soybean___healthy` and citrus “greens”

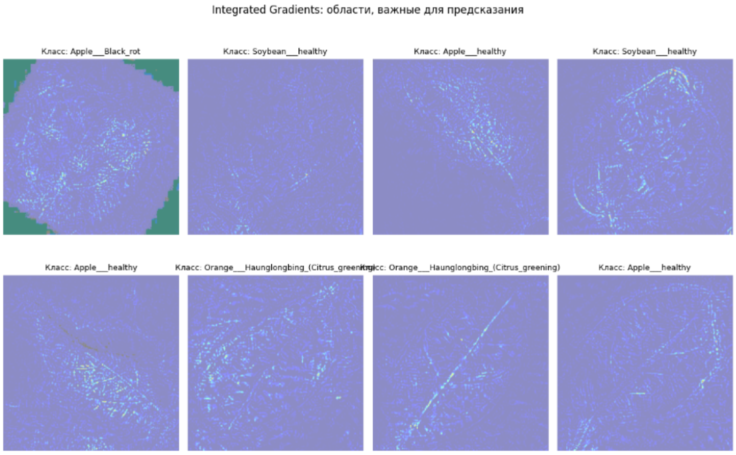

Figure 8. Integrated gradients: the model is based on veins and lesion boundaries, which corresponds to visual criteria for leaf diagnostics

5. Conclusion

The experiment confirmed all stated hypotheses and answered the research questions. Embedding the ExG physiological index in the green channel reduced the dependence on background and illumination, resulting in more reliable performance on the test. A multi-task setup with parallel view prediction served as a soft regularization and improved the network’s generalization ability.

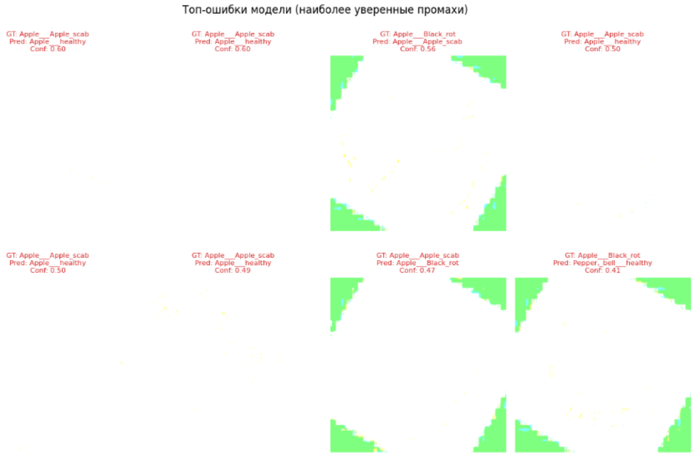

Figure 9. Top mistakes: definite misses are concentrated around mild apple scab, similar in texture to a healthy surface

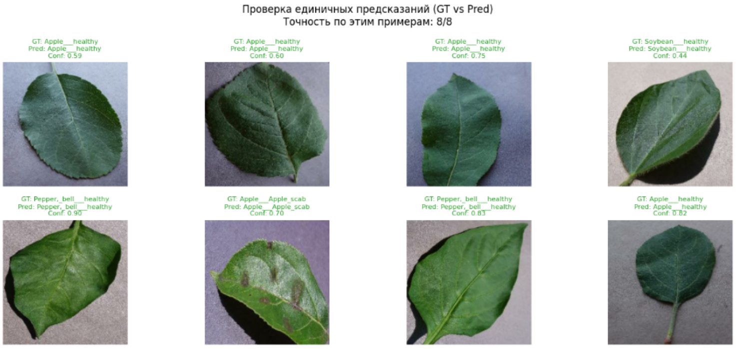

Figure 10. Testing of single examples (GT vs Pred): demonstration of correct and difficult cases on the test

A two-stage scheme with partial unfreezing of the final convolutional blocks provided an additional performance boost with comparable training time. The final accuracy of HiP^2-Net on the test portion was 0.9306 versus 0.8750 for the baseline EfficientNet-B0; simple test augmentation added a few tenths of a percent to 0.9375. Integrated gradient maps demonstrated alignment of the model's attention with veins and focal boundaries, increasing confidence in the predictions and confirming the physiological validity of the chosen pipeline.

The scientific significance of this work lies in the combination of three practical ideas that are easily replicated in practical settings: gentle physiological enhancement of the green channel, a multitasking "species + disease" head, and a short retraining regimen with partial unfreezing. This set of solutions demonstrates how simple engineering interventions improve classifier robustness without complicating the architecture or significantly increasing computational costs. For agroinformatics, these results are important because they encourage the development of more explainable and portable diagnostic systems suitable for mobile devices and field scenarios.

Practical prospects include expansion to a wider range of crops and diagnoses, integration into mobile agricultural consulting apps, and cloud-based crop monitoring services. The approach is expected to evolve in several directions. Expanding datasets beyond laboratory settings and incorporating field images will improve external validity. Self-learning and domain-specific adaptation will help reduce labeling requirements when implementing on new farms. Quantization, channel threshold pruning, and structural thinning will facilitate deployment on edge devices. Combining classification with leaf blade segmentation and lesion area assessment will create the basis for applied phytopathological analytics. Finally, probability calibration and uncertainty management will enable the implementation of a "human-in-the-loop" mode, where an agronomist confirms questionable cases and guides active learning.

Author Contributions: Conceptualization, A.S. Muthana and E.V. Lyapuntsova; methodology, A.S. Muthana; software, A.S. Muthana; validation, A.S. Muthana and E.V. Lyapuntsova; formal analysis, A.S. Muthana; investigation, A.S. Muthana; resources, E.V. Lyapuntsova; data curation, A.S. Muthana; writing—original draft preparation, A.S. Muthana; writing—review and editing, A.S. Muthana and E.V. Lyapuntsova; visualization, A.S. Muthana; supervision, E.V. Lyapuntsova; project administration, E.V. Lyapuntsova; funding acquisition, E.V. Lyapuntsova. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: All datasets utilized in this study are publicly available and have been appropriately cited within the article.

Conflicts of Interest: The authors declare no potential conflict of interest.

Declaration on Generative AI: The authors have not employed any Generative AI tools.

References

1. Chen, R., Qi, H., Liang, Y. & Yang, M. Identification of plant leaf diseases by deep learning based on channel attention and channel pruning. *Frontiers in Plant Science* **13** (2022).
2. El Fatimi, E. H. Leaf diseases detection using deep learning methods. *arXiv preprint.* 31 December 2024 (2024).
3. Gupta, A., Garg, P., Thakur, D. & Palit, R. *Plant Disease Detection Using Deep Learning* in *Proceedings of the Fifth Congress on Intelligent Systems (CIS 2024)* **1277** (2025), 373–384.
4. Russel, N. S. & Selvaraj, A. Leaf species and disease classification using multiscale parallel deep CNN architecture. *Neural Computing and Applications* **34,** 19217–19237 (2022).
5. Shoaib, M., Shah, B., El-Sappagh, S., Ali, A., *et al.* An advanced deep learning models-based plant disease detection: A review of recent research. *Plant Bioinformatics* **14** (2023).
6. Yang, B., Wang, Z., Guo, J., Guo, L., *et al.* Identifying plant disease and severity from leaves: A deep multitask learning framework using triple-branch Swin Transformer and deep supervision. *Computers and Electronics in Agriculture* **209,** 107809 (2023).
7. Sundhar, S., Sharma, R., Maheshwari, P., Kumar, S. R. & Kumar, T. S. Enhancing Leaf Disease Classification Using GAT-GCN Hybrid Model. *arXiv preprint.* 7 April 2025 (2025).
8. Romiyal, G., Selvarajah, T., Roshan, G., Kayathiri, M., *et al.* Past, present and future of deep plant leaf disease recognition: A survey. *Computers and Electronics in Agriculture* **234,** 110128. doi:10.1016/j.compag.2025.110128 (2025).
9. Tunio, M. H., Li, J., Zeng, X., Ahmed, A., *et al.* Advancing plant disease classification: A robust and generalized approach with transformer-fused convolution and Wasserstein domain adaptation. *Computers and Electronics in Agriculture* **226,** 109574 (2024).
10. Lu, F., Shangguan, H., Yuan, Y., Yan, Z., *et al.* LeafConvNeXt: Enhancing plant disease classification for the future of unmanned farming. *Computers and Electronics in Agriculture* **233,** 110165 (2025).

11. Zhang, E., Zhang, N. & Lv, C. A lightweight dual-attention network for tomato leaf disease identification. *Frontiers in Plant Science* **15,** 1420584 (2024).
12. Yao, J., Tran, S. N., Garg, S. & Sawyer, S. Deep Learning for Plant Identification and Disease Classification from Leaf Images: Multi-prediction Approaches. *arXiv preprint.* 25 October 2023 (2023).
13. Indira, K. & Mallika, H. Classification of Plant Leaf Disease Using Deep Learning. *Journal of The Institution of Engineers (India): Series B* **105,** 609–620 (2024).
14. Deng, H., Luo, D., Zhou, Z., Hou, J., *et al.* Leaf disease recognition based on channel information attention network. *Multimedia Tools and Applications* **83,** 6601–6619 (2024).
15. Khan, M. A., Huss, Khan, M. S., *et al.* Deep learning-based segmentation and classification of leaf images for detection of tomato plant disease. *Frontiers in Plant Science* **13,** 1031748 (2022).
16. Chen, W., Chen, J., Duan, R., Fang, Y., *et al.* MS-DNet: A mobile neural network for plant disease identification. *Computers and Electronics in Agriculture* **199,** 107175 (2022).
17. Quan, S., Wang, J., Jia, Z., Yang, M. & Xu, Q. MS-Net: a novel lightweight and precise model for plant disease identification. *Frontiers in Plant Science* **14,** 1276728 (2023).
18. Shafik, W., Tufail, A., Liyanage, C. D. S. & Apong, R. A. A. H. M. Using transfer learning-based plant disease classification and detection for sustainable agriculture. *BMC Plant Biology* **24,** 136 (2024).
19. Zhang, Z. & Wang, H. MAFDE-DN4: Improved few-shot plant disease classification method based on meta-learning. *Computers and Electronics in Agriculture* **225,** 109540 (2024).
20. Li, H., Zhang, Z., Zhou, P., *et al.* An Effective Image Classification Method for Plant Diseases with Improved Channel Attention (aECAnet). *Symmetry* **16,** 451 (2024).
21. Zhang, Y., Ren, P., Fu, X., *et al.* Improving plant disease classification using realistic data augmentation. *Multimedia Tools and Applications* **83,** 37703–37723 (2024).

Information about the authors

Ali Salem Muthana—PhD Student at National University of Science and Technology MISIS (NUST MISIS) (e-mail: m2112648@edu.misis.ru, phone: +79198121751, ORCID: 0000-0003-4304-7469)

Elena Vyacheslavovna Lyapuntsova—Professor at National University of Science and Technology MISIS (NUST MISIS) (e-mail: lev77@me.com, phone: +79150247700, ORCID: 0000-0002-3420-3805)

УДК 519.872, 519.217
PACS 07.05.Tp, 02.60.Pn, 02.70.Bf
DOI: 10.22363/2658-4670-2025-33-4-361-373 EDN: HYUCMC

Распознавание болезней листьев с помощью методов глубокого обучения

А. С. Мутхана, Е. В. Ляпунцова

Национальный исследовательский технологический университет “МИСиС”, Ленинский проспект, д. 4, Москва, 119049, Российская Федерация

Аннотация. Цифровизация растениеводства выдвинула распознавание болезней по изображениям листьев в число приоритетных задач. В работе представлена компактная и воспроизводимая система, пригодная для быстрого развёртывания в облачной среде и последующей адаптации. Подход сочетает многозадачное обучение (одновременное предсказание вида растения и болезни), физиологически мотивированную обработку каналов и устойчивые к ошибкам процедуры подготовки данных. Эксперименты выполнены на наборе New Plant Diseases Dataset (Augmented). Для ускорения выбраны шесть наиболее представленных классов; по каждому использовано до 120 изображений. Данные масштабировались до 192×192 и дополнялись геометрическими и цветовыми преобразованиями, а также мягкими синтетическими пятнами поражения. Индекс зелени ExG внедрялся в зелёный канал входного изображения. Архитектурной основой служила EfficientNet-B0: предложенная HiP^2-Net имела две классификационные головы для болезни и вида. Обучение проводилось в два коротких этапа с частичной разморозкой хвоста базовой сети на втором этапе. Оценивание включало стандартные метрики, матрицы ошибок, тестовую аугментацию при выводе и анализ карт интегрированных градиентов для объяснимости. На сформированном подмножестве многозадачная HiP^2-Net стабильно превосходила замороженную базовую модель по доле верных ответов и сводным метрикам. Синтетические пятна снижали чувствительность к фону и помогали распознавать слабые поражения, а внедрение ExG улучшало выделение тканей листа при переменном освещении. Карты интегрированных градиентов показывали фокус на прожилках и очагах некроза, что укрепляло доверие к предсказаниям и облегчало экспертную интерпретацию. Предлагаемая схема соединяет практичность облачного запуска и простые приёмы, опирающиеся на физиологию листа. Рекомендуется использовать постановку «вид+болезнь», включать ExG в предобработку и добавлять мягкие синтетические пятна: эти шаги повышают устойчивость к освещению, фону и геометрическим вариациям и упрощают перенос на новые коллекции изображений.

Ключевые слова: распознавание болезней растений, изображения листьев, глубокое обучение, перенос обучения, многозадачная классификация, интерпретируемость, лёгкие модели