

Robust Blueberry Leaf Disease Detection Using Transformer-Based and Local Cosine Feature Method

Vinh Dinh Nguyen[1(✉)], Ngoc Phuong Ngo[2], and Kha Hoang Nguyen[1]

[1] Department of Information Technology, FPT University, Can Tho Campus, Can Tho City, Vietnam
vinhnd29@fe.edu.vn, khanhce171115@fpt.edu.vn

[2] Department of Plant Physiology-Biochemistry, College of Agriculture, Can Tho University, Can Tho City, Vietnam
npngoc@ctu.edu.vn

Abstract. Detecting blueberry leaf diseases early is essential for safeguarding crop quality and yield. Traditional methods for this task often involve feature preprocessing combined with machine learning techniques such as K-NN, decision trees, or support vector machines. Recently, there has been a growing interest in deep learning approaches to enhance the performance of existing leaf disease detection systems. While convolutional deep learning methods have shown promising results under standard evaluation conditions, their performance tends to degrade in the presence of noise, such as Gaussian noise. Given the successful application of transformers in the vision domain, this study aims to investigate the use of transformers combined with a novel local descriptor to improve plant leaf disease detection performance under challenging conditions. Experimental results demonstrate that the proposed system achieves stable and superior performance even in noisy environments, outperforming existing methods, including YOLOv5 and YOLOv8, by 2.9% and 4.4% in terms of accuracy, respectively.

Keywords: leaf disease detection · machine learning · deep learning · blueberry leaf disease · local pattern

1 Introduction

The precise identification of plant diseases has significant value in maintaining sustainable and effective agricultural practices, while also avoiding unnecessary depletion of financial and other resources. Although certain plant ailments may lack observable indicators, necessitating the application of advanced analytical techniques, the majority exhibit visible symptoms. Currently, the widely accepted approach involves skilled plant pathologists diagnosing diseases by visually inspecting infected plant leaves. To effectively diagnose plant diseases, a

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
O. Gervasi et al. (Eds.): ICCSA 2025, LNCS 15648, pp. 320–331, 2025.
https://doi.org/10.1007/978-3-031-97000-9_20

plant pathologist needs keen observational abilities to recognize different symptoms. However, the extensive range of plant species, fluctuations in disease progression caused by climate changes and the rapid spread of diseases to new regions where they have not previously been observed can challenge even experienced pathologists, causing them to struggle to diagnose certain diseases [2]. Numerous techniques have been suggested for detecting and classifying plant leaf diseases, encompassing both unsupervised methods like k-means, c-means, and LDA, as well as supervised approaches such as SVM, Naïve Bayes, ANN, Decision Tree, and k-NN. In addition, alternative methods such as fuzzy logic or feature-based learning have been explored. Recently, deep learning models have gained traction for improving the performance of plant leaf disease detection. These include AlexNet (2012), ZFNet (2013), VGG and GoogleNet (2014), ResNet and UNet (2015), YOLO and R-CNN (2016), Fractale Net (2017), and EfficientNet (2018) [2]. The adoption of deep convolutional neural networks in these approaches significantly boosts the accuracy of plant leaf disease detection.

The use of local binary pattern (LBP) has been demonstrated to be advantageous in improving the performance of texture classification applications [6]. Numerous studies have investigated LBP to enhance performance across various application domains such as texture classification, stereo matching, and object detection [3,15]. Several researchers have been investigated YOLO for object detection and classification [22]. Recently, transformers have emerged as successful tools in natural language processing applications, leveraging their capabilities in encoding text sequences through positional encoding, self-attention, and cross-attention mechanisms [20]. Furthermore, there has been a surge in research investigating the benefits of transformers for computer vision tasks, exemplified by studies on vision transformers [7]. Plant leaf dataset acquired in uncontrolled environments introduces greater system complexity, yet it holds immense significance in agricultural development and is increasingly relevant in contemporary research. In addition, similarity in infected areas often results in the extraction of inappropriate features, leading to erroneous classifications based on irrelevant feature matching. Therefore, this research marks the first attempt to introduce an efficient architecture combining a novel LBP descriptor and transformer concept for blueberry leaf disease detection. Experiment on two datasets, including Plant Leaf Disease [17] and our custom dataset were captured at CanTho, Vietnam [14] demonstrated that the proposed method obtained a good performance even under difficult testing conditions. We summarize the main contributions of this research as follows:

- A novel local descriptor, named as CS-LBP, which evaluates the cosine similarity between neighboring regions.
- A robust feature for identifying plant leaf diseases by combining the attributes of CS-LBP with linear projection embedding and positional encoding.
- Incorporating the proposed feature, comprising CS-LBP and linear embedding features, into Yolov8 [10] to ensure stable detection of blueberry leaf diseases.

The rest of this study is organized as follows. Section 2 presents and discusses the related research on plant leaf disease detection. Section 3 introduce the proposed method. The results obtained in the research are given and discussed in Sect. 4. Final, the research is concluded with Sect. 5.

2 Related Work

According to a comprehensive report in [2], artificial intelligence methods are being employed to improve agricultural production through the monitoring of plant diseases. Several research papers have been published on this topic, with some focusing on specific approaches and others concentrating on particular diseases. However, a comprehensive summary of plant disease detection, classification, and diagnosis is lacking. Therefore, this research aims to highlight various techniques employed by different researchers in this field as shown in Table 1.

Traditional methods for detecting plant leaf diseases typically involve various preprocessing techniques, such as color conversion (e.g., RGB to HSV, Lab), and image enhancement algorithms, including median filtering, Laplacian filtering, histogram equalization, or augmentation techniques. In 2019, Rath et al. introduced a method for detecting and classifying rice leaf diseases using a radial basis function neural network [16]. Their system achieved an accuracy of 95%, precision of 97%, and recall of 95%. Recognizing that utilizing multiple features can enhance detection system performance, Shrivastava et al. explored 14 different color models and extracted four characteristics from each color channel, resulting in a total of 172 features [18]. These features were then inputted into an SVM for classification, achieving an accuracy of 94.65%. Instead of relying on a single method for detecting leaf diseases in wheat and corn, Kusumo et al. combined several models simultaneously, including support vector machines, decision trees, random forests, and naive Bayes, along with preprocessing methods such as SIFT and SURF [12]. This combined approach achieved an accuracy of 87%. Rather than inputting 14 distinct color models into the SVM, Kaur et al. utilized only three features comprising RGB color, texture, and shape [11]. This combination yielded an accuracy of 84% with the Soybean dataset. Recently, Feretinos et al. demonstrated that the performance of detection systems can be enhanced by employing deep learning approaches, specifically using transfer learning with AlexNet and VGG-16 [4]. For detecting leaf diseases in grapes, Javidam et al. employed the Gray Level Co-Occurrence Matrix (GLCM) with the support vector machine algorithm [8]. Das et al. proposed another approach for detecting tomato leaf diseases by preprocessing input features with mean and entropy before inputting them into SVM and K-NN for classification, achieving an accuracy of 87.60% on the Tomato dataset [5]. Chen et al. also aimed to improve the performance of plant leaf disease detection by employing transfer learning techniques with VGG16 and InceptionNet [13]. More recently, Atila et al. introduced a method for detecting leaf diseases in tomatoes using EfficientNet [1]. From this brief review, it is evident that most existing methods perform well under normal testing conditions; however, their performance tends

to degrade under challenging conditions such as Gaussian noise. Therefore, there is a need to investigate methods to enhance the performance of existing systems by introducing new local descriptors along with linear embedding and positional encoding.

3 Proposed Method

The proposed method comprises three primary stages, as illustrated in Fig. 1. First, we introduce a novel local descriptor called CS-LBP to establish a stable feature representation. Second, to preserve the structure of the image, we adopt a similar framework to the vision transformer [7]. Here, the original image and the output of the CS-LBP are divided into fixed-size patches, each of which is linearly embedded and supplemented with position embeddings. Finally, the linearly embedded features and position embeddings are fed into YOLOv8 [10] for the detection and classification of blueberry leaf diseases.

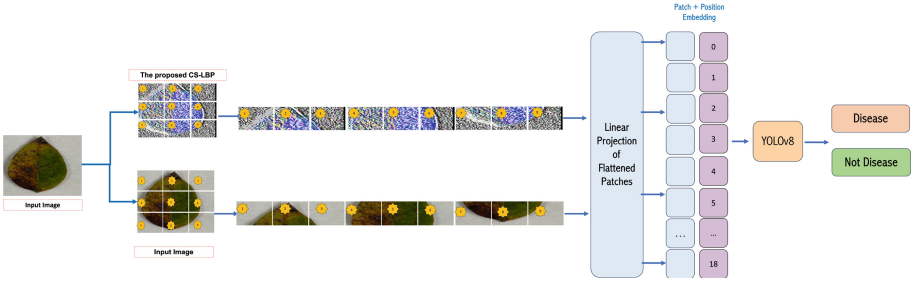

Fig. 1. The proposed LBP and Vision Transformer Architecture for Blueberry Leaf Disease Detection.

3.1 The Proposed CS-LBP

The Local Binary Pattern (LBP) is a technique that doesn't rely on specific parameters and instead focuses on grasping the local patterns within an image. It achieves this by comparing the intensity of surrounding pixels to that of the central pixel, generating binary values based on whether the surrounding pixel intensities surpass that of the center. These binary values are then combined to form a numerical representation. Figure 2(b) illustrates the outcomes of the standard LBP method, utilizing a 3×3 window with eight neighboring pixels ($N = 8$), and a neighborhood radius of 1 ($R = 1$). However, LBP's effectiveness diminishes notably in uniformly textured regions and when subjected to Gaussian noise. Its primary vulnerability lies in the thresholding process between the central pixel and its neighbors. Even slight variations in the central pixel's

Table 1. Survey of existing plant leaf disease detection and classification

Year	Authors	Method	Discussion
2019	Rath and Meher [2]	They uses radial basis function neural network for classification	Rice dataset: Real Field Image. Accuracy of 95%, Precision of 97% and Recall of 95%
2021	Shrivastava and Pradhan [18]	They investigated 14 distinct color models and derived four characteristics from each color channel, resulting in a total of 172 features. The output feature is then input to SVM for classification step	Rice dataset: Rice Real Field Image. Accuracy of 94.65%
2019	Kusumo et al. [12]	They used combination of various techniques such as support vector matching, decision tree, random forest and naïve bayes along with the pre-processing method including SIFT and SURF	Wheat and Corn dataset: Plant Village. Accuracy of 87%
2018	Kaur et al. [11]	SVM was used three input features: color, texture and shape	Soyabean dataset: Plant Village. Accuracy of 84%
2018	Feretinos [4]	The transfer learning is employed with the Alexet and VGG 16	Mix dataset: Plant Village. Accuracy of 99.53%.
2023	Javidam et al. [8]	Gray Level Co-Occurrence Matrix (GLCM) was used with support vector machine algorithm	Seft Grape dataset. Accuracy of 98.97%.
2020	Das et al. [5]	The input image was preprocessing by using Mean and Entropy, the SVM and KNN was then used for classification step.	Tomato dataset. Accuracy of 87.60%.
2020	Chen et al. [13]	Transfer learning the deep CNN was applied in the identification of plant disease (VGGNet16 and InceptionNet)	Self dataset. Accuracy of 92%.
2020	Wspanialy Moussan [21]	They employed ResNet and Unet for detecting plant leaf disease	Tomato dataset: Plant Village. Accuracy of 98%.
2021	Atila et al. [1]	They Investigate EfficientNet-B4-5 for detecting tomato leaf disease	Tomato dataset: Plant Village. Accuracy of 98.42%

value can significantly alter the resulting LBP encoding sequence. The traditional LBP fails to extract the local feature when the noise occur as shown in Fig. 2(b). Therefore, we introduce an approach to handle noise by considering the relationship of local neighbor region as shown in Fig. 2(a). Given a center pixel Lc=(x,y) in the image, the proposed CS-LBP is computed as follows:

$$CS_LBP_{N,R}\left(L_c\right) = \sum_{n=1}^{N} F\left(\psi\left(L_c\right), L_n\right) \times 2^n \tag{1}$$

where R and N are the radius and number of neighbor pixel in the local region. The $F\left(\psi\left(L_c\right), L_n\right)$ is calculated as follows:

$$F\left(\psi\left(L_c\right)\right) = \begin{cases} 1 & \text{if } I\left(L_c\right) \geq I\left(L_n\right) \\ 0 & \text{if } I\left(L_c\right) < I\left(L_n\right) \end{cases} \tag{2}$$

where $I\left(L_n\right)$ is the pixel intensity value of of n-th pixel. $\psi\left(L_c\right)$ is designed to capture the relationship between left local region of the current computing CS_LBP as illustrated in Fig. 2(a):

$$\psi\left(L_c\right) = \frac{\underset{V_1}{\rightarrow}\left[L_0^1, L_1^1, ..., L_n^1\right] \cdot \underset{V_2}{\rightarrow}\left[L_0^1, L_1^2, ..., L_n^2\right]}{\left\|\underset{V_1}{\rightarrow}\left[L_0^1, L_1^1, ..., L_1^n\right]\right\| \left\|\underset{V_2}{\rightarrow}\left[L_0^2, L_1^2, ..., L_n^2\right]\right\|} \times I\left(L_c\right) \tag{3}$$

where $\underset{V_1}{\rightarrow}\left[L_0^1, L_1^1, ..., L_n^1\right]$ is the vector of pixel intensity values of the considering CS-LBP region and $\underset{V_2}{\rightarrow}\left[L_0^2, L_1^2, ..., L_n^2\right]$ is the is the vector of pixel intensity values of the left local neighbor regions as shown in Fig. 2 (a).

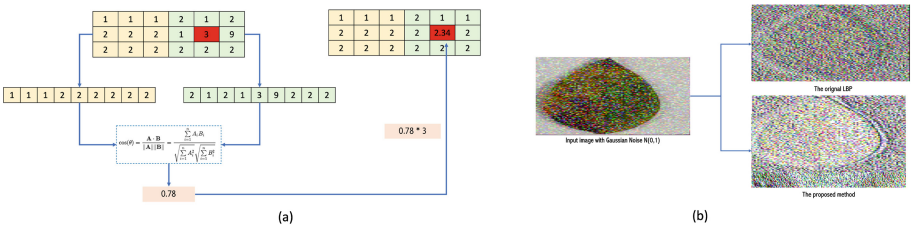

Fig. 2. (a) The proposed CS-LBP. (b) The outcomes from the proposed approach compared with those from the original LBP technique.

3.2 Robust Feature By Using Linear Projection and Positional Encoding Feature

Motivation by the advantages of existing vision transformer algorithm, we introduce an approaches to combine the original image and CS-LBP features by using linear project and positional encoding techniques based on the previous research [20]. The conventional Transformer model takes in a one-dimensional sequence consisting of token embeddings as its input. Similar to [20], the RGB image and CS-LBP feature are reshaped from (Height,Width,Channel) into a sequence of flatten 2D patches N×P2×C as show.
For each patch, $x_p^i \in R^{P\times P\times C} \rightarrow x_p^i \in R^{1\times P^2\times C}$ and $x_p^i \in R^{1\times P^2\times C}.W \in R^{P^2C\times D} = x_p^i W = z_1^i \in R^{1\times D}$.

For all patches at a time, N is total tokens (N=P*P): $x_p^i \in R^{N\times P^2\times C}.W \in R^{P^2C\times D} = xW = z_1 \in R^{N\times D}$, where N, calculated as H×W/P2, represents the resulting number of patches.

The output feature from the linear projection step are then fused with the positional encoding to provide a final output feature for the Yolov8 [10].

4 Experimental Results and Discussion

4.1 Dataset

There are two datasets, including Plant Leaf Disease [17] and our custom dataset were captured at CanTho, Vietnam [14]. The Plant Leaf Disease dataset [17]

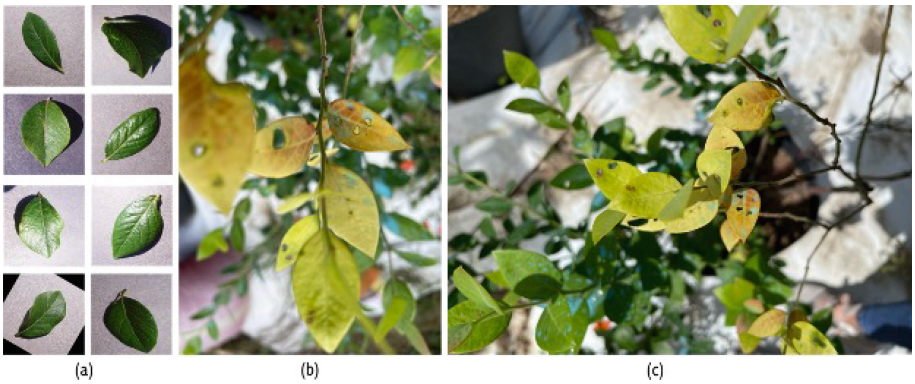

Fig. 3. Samples blueberry leaf datasets from Plant Leaf Disease [17] and our dataset were captured at CanTho, Vietnam [14].

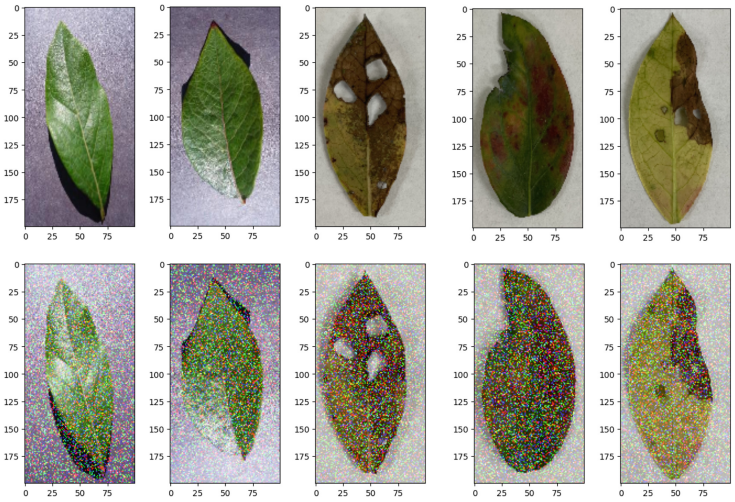

Fig. 4. Samples image are added Gaussian Noise with N(0,1). First row are shown the original images. Second are shown the noise image.

underwent offline augmentation using the original dataset available on GitHub. It consists of approximately 87,000 RGB images showcasing both healthy and diseased crop leaves, categorized into 38 distinct classes. The dataset is divided into training and validation sets in an 80/20 ratio while preserving the directory structure. Moreover, a separate directory is created later, containing 33 test images for predictive analysis. In this research, our experiments are centered around particular classes extracted from the Plant Diseases Dataset [17]. Specifically, we delve into the blueberry healthy class, comprising 1817 images as shown in Fig. 3(a). Additionally, we also evaluate the performance on our

dataset showcasing blueberry diseases, as depicted in Fig. 3(b, c). This dataset was gathered at Can Tho University, Vietnam [14].

To create a comprehensive dataset for evaluation. First, the augmentation algorithm [16] was used to create 10,000 samples, where 7,000 samples for training, 2,000 sample for validation, and 1000 samples for testing from the Plant Leaf Disease [14] and Can Tho University Dataset [10]. In addition, to created the difficult of the testing dataset, we add Gaussian noise into the original image, as described in Fig. 4.

4.2 Metric for Evaluation and Algorithms for Comparison

This research used four metrics to evaluate the performance of the proposed method included precision Φ and recall Γ , F1 score, and accuracy Υ as described below:

$$\Phi = \frac{\sum TP}{\sum TP + FP} \tag{4}$$

$$\Gamma = \frac{\sum TP}{\sum TP + FN} \tag{5}$$

$$F1 = \frac{2 \times \Phi \times \Gamma}{\Phi + \Gamma} \tag{6}$$

$$\Upsilon = \frac{TP + TN}{TP + TN + FN + FP} \tag{7}$$

To assess the effectiveness of the proposed method compared to existing approaches, we conducted an experiment that included several established methods, namely YOLOv8 [10], YOLOv5 [9], and EfficientNet [19].

4.3 Detail Implementation

Google Colab (a 2.30 GHz CPU, a 16 GB Tesla GPU, and 26 GB of DRAM) was used to train our proposed method. For consistency, we employed the same parameter settings as YOLOv5+LBP, YOLOv8, and EfficientNet during both the training and testing phases.

4.4 Results and Discussions

We first verify the performance of the proposed system and compare it to the method under normal condition cases, where no noise is added to the input image. Figure 5 shows the results of the proposed system and compared methods using four evaluation metrics: precision, recall, F1-score, and accuracy. Yolov8 obtained precision, recall, F1-score, and accuracy of 95.2%, 90.9%, 93.02%, and 97%, respectively. Yolov5+LBP obtained precision, recall, F1-score, and accuracy of 92.68%, 88.37%, 90.47%, and 96%, respectively. Yolov8+LBP obtained

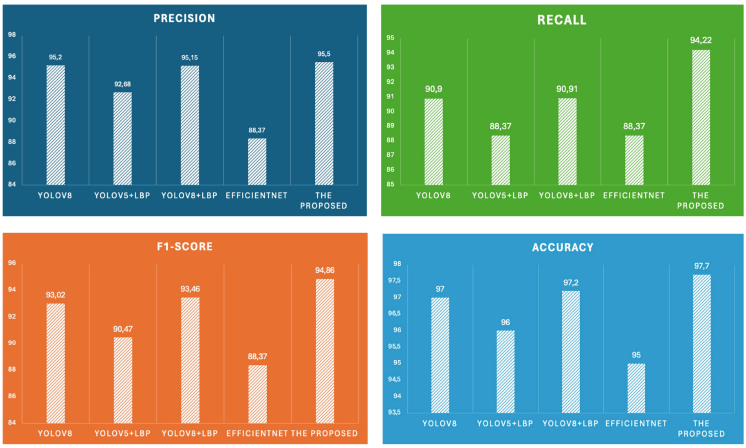

Fig. 5. Experimental results of the proposed system and existing methods under normal conditions.



Fig. 6. Experimental results of the proposed system and existing methods under difficult conditions (Gaussian noise is incorporated).

precision, recall, F1-score, and accuracy of 95.15%, 90.91%, 93.46%, and 97.2%, respectively. EfficientNet obtained precision, recall, F1-score, and accuracy of 88.37%, 88.37%, 88.37%, and 95%, respectively. The proposed method obtained precision, recall, F1-score, and accuracy of 95.5%, 94.22%, 94.86%, and 97.7%, respectively. Observing the results, we found that the performance of EfficientNet is lowest in terms of precision, recall, and accuracy, while the proposed method obtained the best performance. For normal conditions, most compared methods obtained comparable results to the proposed method because they successfully capture and learn the input structure. Our method provides good performances

due to the benefits of the proposed CS-LBP, along with the constraints of linear embedding.

However, the results of the existing algorithms significantly drops when we add noise to the input under difficult conditions. Figure 6 shows the performance of the our method and existing methods under Gaussian noise with N(0,1). The performance of Yolov8 was decreased by 7.4%, 10.9%, 9.3%, and 4% in terms of precision, recall, F1-score, and accuracy, respectively. Analyzing Yolov8, we found that this happens because it was only designed for handling normal input conditions. We further conducted an ablation study by fusing Yolov8 and LBP, as shown in Fig. 6. The precision, recall, F1-score, and accuracy of Yolov8+LBP are better than the original Yolov8 due to the benefits of LBP. The performance of Yolov8+LBP decreased by only 4.54%, 5.13%, 5.33%, and 2.4% in terms of precision, recall, F1-score, and accuracy, respectively. For a deeper study of the behavior of another approach, Yolo5+LBP, we realized that its performance was also decreased under difficult testing; however, the decreasing rate is lower compared to YOLOv8 without using LBP. The performance of Yolov5+LBP decreased by only 3.98%, 2.01%, 3.12%, and 1.5% in terms of precision, recall, F1-score, and accuracy, respectively. We also conducted more experiments using EfficientNet under difficult conditions. The performance of EfficientNet was significantly decreased by 5.8%, 10.11%, 8.01%, and 3.8% in terms of precision, recall, F1-score, and accuracy, respectively. Thus, we found that, somehow, LBP might help to improve the performance of the compared methods under difficult conditions. However, as discussed earlier, LBP still fails to handle huge noise. In addition, Yolov8, Yolov5, and EfficientNet are designed based on convolutional approaches; therefore, they fail to reverse the constraint of the structure of the input object. This is the main reason why our proposed method still obtained the best performance under difficult conditions. The performance of the proposed method was only decreased by 0.48%, 0.89%, 0.69%, and 0.3% in terms of precision, recall, F1-score, and accuracy, respectively, under difficult conditions, because CS-LBP successfully extracts and handles the noise case by considering the distribution of the local region with the help of cosine similarity metric. In addition, the idea of linear project embedding and positional encoder also helps to reverse the order constraint of the object.

5 Conclusion

This study has introduced an effective approach to enhance the performance of plant leaf disease detection by incorporating the proposed CS-LBP along with linear embedding and positional encoding to generate a robust feature representation. The developed system achieved consistent and competitive performance compared to state-of-the-art methods under both normal and noisy conditions. However, a limitation of the proposed method lies in its processing time, attributed to the growing number of input features for training and inference. As a future direction, we intend to explore multi-modal fusion approaches to further enhance the performance of the proposed method.

References

1. Ümit, A., Uçar, M., Akyol, K., Uçar, E.: Plant leaf disease classification using efficientnet deep learning model. Ecol. Inf. **61**, 101182 (2021). https://doi.org/10.1016/j.ecoinf.2020.101182
2. Bhargava, A., Shukla, A., Goswami, O.P., Alsharif, M.H., Uthansakul, P., Uthansakul, M.: Plant leaf disease detection, classification, and diagnosis using computer vision and artificial intelligence: a review. IEEE Access **12**, 37443–37469 (2024). https://doi.org/10.1109/ACCESS.2024.3373001
3. Dinh, V.Q., Nguyen, P.H., Nguyen, V.D.: Feature engineering and deep learning for stereo matching under adverse driving conditions. IEEE Trans. Intell. Transp. Syst. **23**(7), 7855–7865 (2022). https://doi.org/10.1109/TITS.2021.3073557
4. Ferentinos, K.P.: Deep learning models for plant disease detection and diagnosis. Comput. Electron. Agricult. **145**, 311–318 (2018). https://doi.org/10.1016/j.compag.2018.01.009
5. Francisco, M., Ribeiro, F., Metrôlho, J., Dionísio, R.: Algorithms and models for automatic detection and classification of diseases and pests in agricultural crops: a systematic review. Appl. Sci. **13**(8) (2023). https://doi.org/10.3390/app13084720
6. Hadid, A.: The local binary pattern approach and its applications to face analysis. In: 2008 First Workshops on Image Processing Theory, Tools and Applications, pp. 1–9 (2008). https://doi.org/10.1109/IPTA.2008.4743795
7. Han, K., Wang, Y., Chen, H., Chen, X., Guo, J., Liu, Z., Tang, Y., Xiao, A., Xu, C., Xu, Y., Yang, Z., Zhang, Y., Tao, D.: A survey on vision transformer. IEEE Trans. Pattern Anal. Mach. Intell. **45**(1), 87–110 (2023). https://doi.org/10.1109/TPAMI.2022.3152247
8. Javidan, S.M., Banakar, A., Vakilian, K.A., Ampatzidis, Y.: Diagnosis of grape leaf diseases using automatic k-means clustering and machine learning. Smart Agricult. Technol. **3**, 100081 (2023). https://doi.org/10.1016/j.atech.2022.100081
9. Jocher, G.: Ultralytics YOLOV5 (2020). https://doi.org/10.5281/zenodo.3908559
10. Jocher, G., Chaurasia, A., Qiu, J.: Ultralytics YOLOV8 (2023). https://github.com/ultralytics/ultralytics
11. Kaur, S., Pandey, S., Goel, S.: Semi-automatic leaf disease detection and classification system for soybean culture. IET Image Process. **12**, 1038–1048 (2018). https://api.semanticscholar.org/CorpusID:49238886
12. Kusumo, B.S., Heryana, A., Mahendra, O., Pardede, H.F.: Machine learning-based for automatic detection of corn-plant diseases using image processing. In: 2018 International Conference on Computer, Control, Informatics and its Applications (IC3INA), pp. 93–97 (2018). https://doi.org/10.1109/IC3INA.2018.8629507
13. Mohanty, S.P., Hughes, D.P., Salathé, M.: Using deep learning for image-based plant disease detection. Front. Plant Sci. **7** (2016). https://api.semanticscholar.org/CorpusID:2528492
14. Nguyen, V.D., Ngo, N.P., Debnath, N.C.: Leaf disease detection in blueberry using efficient semi-supervised learning approach. In: Hassanien, A.E., et al. (eds.) The 3rd International Conference on Artificial Intelligence and Computer Vision (AICV2023), March 5–7, 2023, pp. 188–196. Springer, Cham (2023)
15. Nguyen, V.D., Nguyen, P.H., Debnath, N.C.: Local binary pattern and census, which one is better in stereo matching. In: 2020 7th NAFOSTED Conference on Information and Computer Science (NICS), pp. 244–249 (2020). https://doi.org/10.1109/NICS51282.2020.9335907

16. Rath, A.K., Meher, J.K.: Disease detection in infected plant leaf by computational method. Arch. Phytopathol. Plant Prot. **52**(19–20), 1348–1358 (2019). https://doi.org/10.1080/03235408.2019.1708546
17. Ruth, J.A., Uma, R., Meenakshi, A., Ramkumar, P.: Meta-heuristic based deep learning model for leaf diseases detection. Neural Process. Lett. **54**, 5693–5709 (2022). https://api.semanticscholar.org/CorpusID:249834375
18. Shrivastava, V.K., Pradhan, M.K.: Rice plant disease classification using color features: a machine learning paradigm. J. Plant Pathol. **103**(1), 17–26 (2020). https://doi.org/10.1007/s42161-020-00683-3
19. Tan, M., Le, Q.V.: Efficientnet: rethinking model scaling for convolutional neural networks. arXiv preprint arXiv:1905.11946 (2019)
20. Vaswani, A., et al.: Attention is all you need. In: Proceedings of the 31st International Conference on Neural Information Processing Systems, pp. 6000–6010. NIPS'17, Curran Associates Inc., Red Hook, NY, USA (2017)
21. Wspanialy, P., Moussa, M.: A detection and severity estimation system for generic diseases of tomato greenhouse plants. Comput. Electron. Agricult. **178**, 105701 (2020). https://doi.org/10.1016/j.compag.2020.105701
22. Zhao, Z., He, P.: YOLO-mamba: object detection method for infrared aerial images. Sig. Image Video Process. (2024). https://api.semanticscholar.org/CorpusID:272118442