RESEARCH ARTICLE

Expand to portion: Pixel-level challenges in wheat infection and pest segmentation

Leilei Deng[1, 2], Hua Li[1, *], Lifeng Sun[3]

[1]College of Computer Science and Technology, Changchun University of Science and Technology, Jilin, Jilin, China. [2]College of Information Technology, Jilin Agricultural University, Jilin, Jilin, China. [3]Department of Computer Science and Technology, Tsinghua University, Beijing, China.

Received: December 22, 2025; **accepted:** January 16, 2026.

Accurate segmentation of wheat foliar diseases and pest damage is crucial for effective crop management and disease control. However, pest damage typically comprises only a small fraction of the labeled pixels. This extreme pixel-level imbalance poses a significant challenge to segmentation performance, potentially leading to overfitting to common classes and underlearning of rare classes, thus degrading overall performance. This research proposed a random projected copy-and-paste (RPCP) augmentation technique to address the pixel imbalance problem. The rare pest damage patches were extracted from annotated training images and applied random geometric transformations to simulate variations. The transformed patches were then passed into appropriate regions while avoiding overlaps with lesions or existing damaged regions. In addition, a random projection filter was applied to the pasted regions, refining local features and ensuring a natural blend with the new background. The results showed that the proposed method substantially improved segmentation performance on the pest damage class, while maintaining or even slightly enhancing accuracy on other categories. The results highlighted the effectiveness of targeted augmentation in mitigating extreme pixel imbalance, offering a straightforward effective solution for agricultural segmentation problems.

Keywords: semantic segmentation; long-tail distribution; wheat disease; infection; pest; detection.

[*]**Corresponding author:** Hua Li, College of Computer Science and Technology, Changchun University of Science and Technology, Jilin, Jilin, China. Email: lihua@cust.edu.cn.

Introduction

Wheat is one of the most widely cultivated crops and a key source of dietary calories worldwide [1, 2]. However, its yield and grain quality are often threatened by a variety of diseases and pests, leading to substantial economic losses and posing serious challenges to global food security [3, 4]. Early and accurate detection of these threats is essential for effective crop protection, timely intervention, and sustainable management practices. Recent advances in deep learning have enabled automated perception and analysis in agricultural vision tasks, providing an efficient and scalable alternative to traditional manual inspection [5, 6].

Among various computer vision approaches, semantic segmentation has emerged as a powerful tool for pixel-wise recognition of disease and pest symptoms, enabling fine-grained characterization of lesion morphology and spatial distribution [7-9]. Despite its potential, applying semantic segmentation to wheat foliar disease datasets remains challenging due to large intraclass variation and severe class imbalance. Visual symptoms can differ significantly in size and appearance, complicating pixel-wise recognition. Rare classes such as pest damage often occupy only a tiny fraction of annotated pixels. This extreme pixel-level imbalance results in biased optimization, leading models to overfit on dominant classes and neglect rare classes. Even state-of-the-art models such as SegFormer achieve high accuracy on common classes like healthy tissue and septoria tritici blotch (STB) lesions [10], their performance on pest damage regions remains substantially lower [11].

To address these challenges, this research proposed a targeted data augmentation strategy of random projected copy-and-paste (RPCP), which explicitly increased the representation of rare classes for training by using a public wheat foliar disease segmentation dataset containing 3 classes of healthy leaves, STB lesions, and pest damage [12]. The generality of the proposed approach was assessed by comparing with multiple representative segmentation models. This research proposed a model-agnostic strategy that could be seamlessly integrated into diverse existing training pipelines without requiring extra annotations or introducing architectural changes. Through extensive evaluations on multiple representative models, the proposed method consistently improved rare-class accuracy without compromising common class performance, highlighting its strong generalization capability.

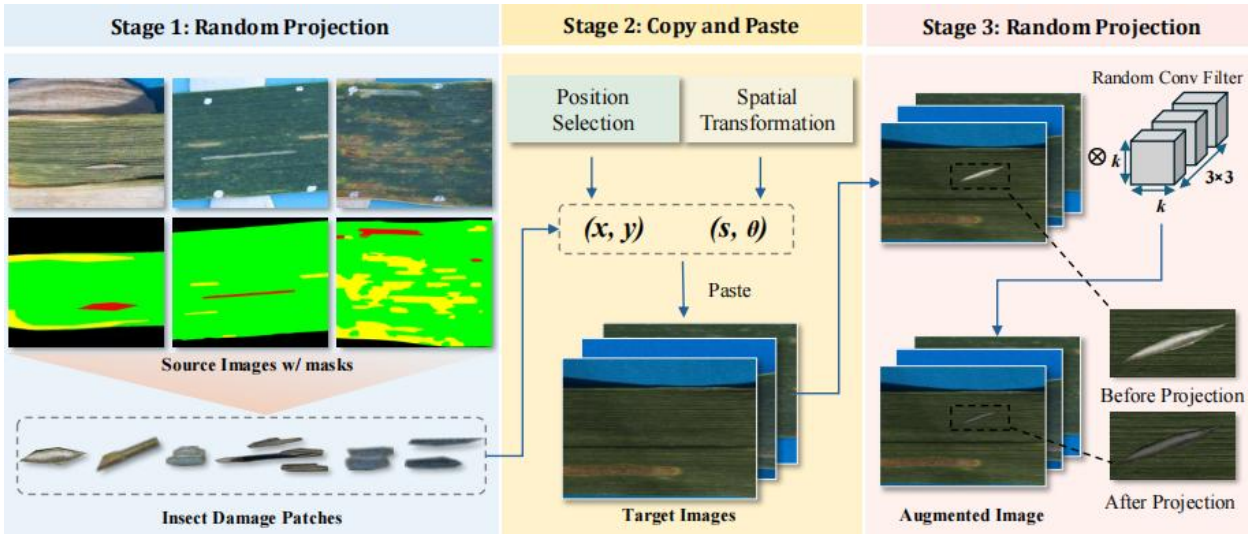
Materials and methods

Research subjects

This study investigated semantic segmentation for a specific type of wheat foliar disease, septoria tritici blotch (STB), and pest damage. Given an RGB image $I \in RH \times W \times 3$, the objective of this study was to predict a per-pixel label map $Y \in \{0, 1, ..., C\}$, where each label corresponded to one of C semantic classes such as healthy leaf, disease lesions, and pest damage. This task exhibited a pronounced class imbalance because pest damage regions were extremely scarce and characterized by irregular and localized patterns. Such rare class pixels were typically small and provided limited training information, leading to sub-optimal segmentation performance. To mitigate this issue, the study proposed a random projected copy-and-paste (RPCP) augmentation strategy to enhance the representation of pest damage regions.

Development of RPCP method

To address the significant class imbalance in wheat foliar disease segmentation, particularly the under representation of regions with pest damage, a targeted data augmentation pipeline was developed, which enriched training images with additional rare-class instances through a two-stage process of category aware patch extraction and spatially constrained pasting. In the category aware patch extraction stage, annotated images were scanned to identify regions belonging to underrepresented classes, which were then cropped into patches with corresponding binary masks. In the spatially constrained pasting stage, each patch was independently transformed by random rotation and scaling and subsequently pasted into contextually appropriate locations. The candidate paste regions were restricted to valid areas like healthy leaf regions, while avoiding overlapping with existing rare class regions to preserve structural integrity. The pasting operation was parameterized by (x, y, s, θ), where (x, y) specified the placement position, s controlled the scaling factor, and θ denoted the rotation angle. While class-aware augmentation directly improved class balance, the copy–paste operation introduced artifacts and texture inconsistencies, potentially harming model generalization. To address these issues, a localized refinement method was introduced, which operated only on pasted regions, altering their local appearance while preserving semantic structure. The random convolution was an effective data augmentation approach, which could distort local textures while preserving the



overall shape [13-19]. This study applied random projection to the pasted regions in augmented images, enabling the creation of visually diverse appearances while maintaining label consistency. Given the augmented image $I' \in RH \times W \times C$ and its patch mask $YP \in \{0, 1\}\ H \times W$, a random projection filter $\Theta \in Rh \times w \times C \times C$ was generated, where H, W, and C were the height, width, and channels of I′, respectively, while h and w were the height and width of Θ. The weights of Θ were randomly sampled from a Gaussian distribution $N(0, \sigma 2)$ with the hyper-parameter σ controlling the perturbation magnitude. The overall proposed RPCP framework was shown in Figure 1.

Experimental dataset

The experiments were conducted using STB dataset introduced by Zenkl *et al*. [13], which was a high-resolution image dataset specifically curated for semantic segmentation tasks of wheat foliar disease and pest damage. The dataset comprised 422 RGB images of wheat leaves captured under diverse lighting conditions with a resolution of 1,024 × 1,024 pixels. For the semantic segmentation task, each image was densely annotated with pixel-level masks for three semantic classes including healthy leaf area, necrotic tissue, and pest damage. Notably, the STB dataset exhibited a significant class imbalance, where healthy leaf regions dominated the pixel distribution, while pest damage appeared sparsely. The skewed distribution introduced a strong bias toward majority classes, often resulting in poor performance on rare categories.

Baseline comparison

A set of representative semantic segmentation models were selected as the baselines including both convolutional and transformer-based architectures. PSPNet [20], CCNet [21], and the DeepLabV3 series [22, 23] were adopted as the representative CNN-based models with multiscale context modules and encoder–decoder designs, while SegFormer [24] and SAN [25] were employed for transformer-based approaches. In addition, advanced architectures such as SegNeXt [26] and ConvNeXt [27] were included, which combined the efficiency of CNN architectures with transformer-inspired contextual modeling to enhance segmentation performance.

Model implementation

All models were trained by using the MMSegmentation framework built upon PyTorch [28-31]. AdamW optimizer (https://docs.pytorch.org/docs/stable/generated/torch.optim.AdamW.html) was adopted with learning rates from $3e^{-5}$ to $1e^{-4}$ and weight decay values from {0.0001, 0.0005, 0.001}. Experiments were conducted by

Table 1. Overall model performance.

Method	Class 1		Class 2		Class 3		Average	
	IoU	Acc	IoU	Acc	IoU	Acc	mIoU	mAcc
DeepLabV3	97.60	98.28	81.96	87.08	57.79	72.19	79.12	85.85
w/RPCP	97.78	99.18	82.17	87.76	61.76	77.14	80.57	88.03
Δ	(+0.18)	(+0.90)	(+0.21)	(+0.68)	(+3.97)	(+4.95)	(+1.45)	(+2.18)
DeepLabV3+	97.29	98.45	81.93	88.29	58.89	75.61	79.37	87.45
w/RPCP	97.75	99.03	84.47	89.85	62.32	78.99	81.51	89.29
Δ	(+0.46)	(+0.58)	(+2.54)	(+1.56)	(+3.43)	(+3.38)	(+2.14)	(+1.84)
PSPNet	97.48	99.02	82.75	88.71	60.91	74.32	80.38	87.35
w/RPCP	97.64	98.89	83.78	90.41	64.38	78.96	81.93	89.42
Δ	(+0.16)	(-0.13)	(+1.03)	(+1.70)	(+3.47)	(+4.64)	(+1.55)	(+2.07)
CCNet	97.60	98.11	84.08	91.62	60.22	75.93	80.63	88.55
w/RPCP	97.79	98.92	84.98	91.10	64.50	82.08	82.42	90.70
Δ	(+0.19)	(+0.81)	(+0.90)	(-0.52)	(+4.28)	(+6.15)	(+1.79)	(+2.15)
SAN	97.53	98.85	83.14	90.10	62.73	76.18	81.13	88.38
w/RPCP	97.57	98.90	83.20	90.17	65.86	77.32	82.21	88.80
Δ	(+0.04)	(+0.05)	(+0.06)	(+0.07)	(+3.13)	(+1.14)	(+1.08)	(+0.42)
SegFormer	97.38	98.01	82.73	90.12	68.16	79.87	82.76	89.33
w/RPCP	97.98	99.12	85.82	91.67	72.36	82.57	85.39	91.12
Δ	(+0.60)	(+1.11)	(+3.09)	(+1.55)	(+4.20)	(+2.70)	(+2.63)	(+1.79)
ConvNeXt	98.00	98.85	85.18	91.57	70.67	84.31	84.62	91.58
w/RPCP	98.00	**99.43**	85.46	89.09	74.30	84.58	85.92	91.03
Δ	(+0.00)	(+0.58)	(+0.28)	(-2.48)	(+3.63)	(+0.27)	(+1.30)	(-0.55)
SegNeXt	97.93	99.06	85.32	91.45	72.81	81.56	85.35	90.69
w/RPCP	98.01	98.97	**86.04**	**92.87**	**75.62**	84.40	**86.56**	**92.08**
Δ	(+0.08)	(-0.09)	(+0.72)	(+1.42)	(+2.81)	(+2.84)	(+1.21)	(+1.39)

Notes: Green numbers indicated improvement, red numbers indicated decrease. Δ indicated the change relative to the baseline.

using batch sizes {4, 8, 16}. During the training process, input images were first randomly resized within a scale ratio range of [0.5, 2.0], then randomly cropped to 512 × 512 pixels followed by horizontal flipping with a probability of 0.5 and color jittering to enhance robustness to illumination changes.

Results and discussion

The segmentation performance of all baseline models and their RPCP-enhanced variants demonstrated that, for healthy leaf (Class 1), all methods achieved very high intersection over union (IoU) and accuracy (Acc) values with differences generally below 1%, indicating that this class was well-represented and easy to segment. Lesion region (Class 2) exhibited moderate variation across models, while all IoU values remained above 81%. In contrast, pest damage (Class 3) showed the largest performance gaps, reflecting the difficulty of segmenting underrepresented features. Among the baselines, SegNeXt achieved the strongest Class 3 performance with 72.81% of IoU and 81.56% of Acc followed closely by ConvNeXt with 70.67% of IoU and 84.31% of Acc (Table 1). The results showed that introducing RPCP consistently improved Class 3 results across a wide range of backbones, highlighting the robustness of the proposed augmentation. The largest IoU gains were observed in CCNet as +4.28% and SegFormer as +4.20%, while DeepLabV3+ (+3.43%) and PSPNet (+3.47%) also exhibited substantial improvements. Moreover, the results on common classes indicated that RPCP did not sacrifice performance where training data was already abundant. Overall, RPCP yielded clear benefits in both average loU (mIoU) and ACC (mAcc) for most models. SegFormer achieved an mIoU increase from 86.26% to 88.89% (+2.63%) and an mAcc increase from 87.54% to 89.33% (+1.79%). Although a slight decline was observed in the overall performance of ConvNeXt, the general trend confirmed that RPCP was effective in enhancing rare class segmentation while maintaining or improving accuracy for dominant classes.

The comparison for infected leaf detection showed that RPCP method for detecting damaged parts of leaf focused on all parts of leaf that could not contribute to the photosynthesis process, the main function of the leaf on the plant, whereas the method developed by Nazare *et al*. only considered the damaged leaf as the destroyed one [32]. According to anatomical concept of the plant, the proposed method of this study came up with an accurate leaf disease detection rate of 26.25% compared to the method of Nazare that took the tested leaf as a healthy one. The proposed method was confirmed with accurate and efficient detection.

This research proposed a rare-class–oriented augmentation framework (RPCP) for wheat leaf disease segmentation. By combining category-aware copy-paste with random-projection refinement, RPCP generated realistic augmented images that enhanced the representation of rare classes. Extensive experiments across diverse segmentation methods demonstrated that RPCP consistently improved rare class performance while maintaining accuracy on common classes. These results highlighted RPCP as a scalable and model-agnostic augmentation strategy for robust plant pathology segmentation.

Acknowledgements

This research was supported by the Scientific Research Projects of the Educational Commission of Jilin Province (Grant No. JJKH20250566KJ).

References

1. Fones H, Gurr S. 2015. The impact of septoria tritici blotch disease on wheat: An EU perspective. Fung Genet Biol. 79:3-7.
2. Shoaib M, Hussain T, Shah B, Ullah I, Shah SM, Ali F, *et al*. 2022. Deep learning-based segmentation and classification of leaf images for detection of tomato plant disease. Front Plant Sci. 13:1031748.
3. Chou HP, Chang SC, Pan JY, Wei W, Juan DC. Remix: Rebalanced mixup. In: European Conference on Computer Vision. pp. 95–110. Springer. 2020.
4. Chu P, Bian X, Liu S, Ling H. 2020. Feature space augmentation for long-tailed data. In: European Conference on Computer Vision. pp. 694–710. Springer. 2020.
5. Erenstein O, Jaleta M, Mottaleb KA, Sonder K, Donovan J, Braun HJ. Global trends in wheat production, consumption and trade. In: Wheat improvement: Food security in a changing climate. pp. 47–66. Springer International Publishing Cham. 2022.
6. Fan J, Gao B, Jin H, Jiang L. 2022. Ucc: Uncertainty guided crosshead co-training for semi-supervised semantic segmentation. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 9947–9956.
7. Wei T, Chen Z, Yu X. 2024. Snap and diagnose: An advanced multimodal retrieval system for identifying plant diseases in the wild. Proceedings of the 6th ACM International Conference on Multimedia in Asia. pp. 1–3.
8. Fang HS, Sun J, Wang R, Gou M, Li YL, Lu C. 2019. Instaboost: Boosting instance segmentation *via* probability map guided copy-pasting. Proceedings of the IEEE/CVF international conference on computer vision. pp. 682–691.
9. Ferentinos KP. 2018. Deep learning models for plant disease detection and diagnosis. Comput Electron Agric. 145:311–318.
10. Gupta A, Dollar P, Girshick R. 2019. Lvis: A dataset for large vocabulary instance segmentation. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 5356–5364.
11. Ghiasi G, Cui Y, Srinivas A, Qian R, Lin TY, Cubuk ED, *et al*. 2021. Simple copy-paste is a strong data augmentation method for instance segmentation. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 2918–2928.
12. Guo MH, Lu CZ, Hou Q, Liu Z, Cheng MM, Hu SM. 2022. Segnext: Rethinking convolutional attention design for semantic segmentation. Advan Neur Info Proc Syst. 35:1140–1156.
13. Zenkl R, McDonald BA, Walter A, Anderegg J. 2025. Towards high throughput in-field detection and quantification of wheat foliar diseases using deep learning. Comput Electron Agric. 232:109854.
14. He R, Yang J, Qi X. 2021. Re-distributing biased pseudo labels for semi-supervised semantic segmentation: A baseline investigation. Proceedings of the IEEE/CVF international conference on computer vision. pp. 6930–6940.
15. Hu H, Wei F, Hu H, Ye Q, Cui J, Wang L. 2021. Semi-supervised semantic segmentation *via* adaptive equalization learning. Advan Neur Info Proc Syst. 34:22106–22118.
16. Huang Z, Wang X, Huang L, Huang C, Wei Y, Liu W. 2019. CCNET: Criss-cross attention for semantic segmentation. Proceedings

of the IEEE/CVF international conference on computer vision. pp. 603–612.
17. Li J, Qiao Y, Liu S, Zhang J, Yang Z, Wang M. 2022. An improved yolov5-based vegetable disease detection method. Comput Electron Agric. 202:107345.
18. Aldakheel EA, Zakariah M, Alabdalall AH. 2024. Detection and identification of plant leaf diseases using yolov4. Front Plant Sci. 15:1355941.
19. Liu Z, Mao H, Wu CY, Feichtenhofer C, Darrell T, Xie S. 2022. A convnet for the 2020s. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 11976–11986.
20. Zhao H, Shi J, Qi X, Wang X, Jia J. 2017. Pyramid scene parsing network. Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 2881–2890.
21. Polly R, Devi EA. 2024. Semantic segmentation for plant leaf disease classification and damage detection: A deep learning approach. Smart Agric Technol. 9:100526.
22. Saad MH, Salman AE. 2024. A plant disease classification using one-shot learning technique with field images. Multimed Tools Appl. 83(20):58935-58960.
23. Samuel D, Chechik G. 2021. Distributional robustness loss for long-tail learning. Proceedings of the IEEE/CVF international conference on computer vision. pp. 9495–9504.
24. Savary S, Willocquet L, Pethybridge SJ, Esker P, McRoberts N, Nelson A. 2019. The global burden of pathogens and pests on major food crops. Nat Ecol Evol. 3(3):430-439.
25. Xu M, Zhang Z, Wei F, Hu H, Bai X. 2023. Side adapter network for open-vocabulary semantic segmentation. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 2945–2954.
26. Shorten C, Khoshgoftaar TM. 2019. A survey on image data augmentation for deep learning. J Big Data. 6(1):1–48.
27. Singh AK, Rao A, Chattopadhyay P, Maurya R, Singh L. 2024. Effective plant disease diagnosis using vision transformer trained with leafy-generative adversarial network-generated images. Expert Syst Appl. 254:124387.
28. Ullah N, Khan JA, Almakdi S, Alshehri MS, Al Qathrady M, El-Rashidy N, *et al*. 2023. An effective approach for plant leaf diseases classification based on a novel deepplantnet deep learning model. Front Plant Sci. 14:1212747.
29. Vinh NX, Erfani S, Paisitkriangkrai S, Bailey J, Leckie C, Ramamohanarao K. 2016. Training robust models using random projection. 2016 23rd International conference on pattern recognition (ICPR). pp. 531–536.
30. Wang J, Zhang W, Zang Y, Cao Y, Pang J, Gong T, *et al*. 2021. Seesaw loss for long-tailed instance segmentation. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 9695–9704.
31. Wang Z, Zenkl R, Greche L, De Solan B, Samatan LB, Ouahid S, *et al*. 2025. The global wheat full semantic organ segmentation (GWFSS) dataset. Plant Phenomics. Plant Phenomics. 7(3):100084.
32. Nazare-JR AC, Menotti D, Neves JMR, Sediyma T. 2010. Automatic detection of the damaged leaf area in digital images of soybean. IWSSIP 2010.