Research on Improved Pavement Distress Detection Algorithm of Yolov12

Nan Deng, Quan Song*

College of Information and Engineering, Swan College of Central South University of Forestry and Technology, Changsha 410211, China

Abstract: Pavement health significantly affects traffic safety, regional economic growth, and residents' living standards. To address the inefficiency and inaccuracy of traditional manual inspection methods, this paper proposes a pavement distress detection algorithm based on an improved YOLOv12 model. By incorporating DySample upsampling and the NWD (Normalized Wasserstein Distance) Loss function, the algorithm enhances detection accuracy and robustness in complex scenarios. Validated on the UAV-PDD2023 dataset, the improved YOLOv12-ND model achieves 83.9% mAP50 and 58.4% mAP50-95, outperforming mainstream algorithms like YOLOv11, YOLOv10, and YOLOv9, especially in detecting repairs and longitudinal cracks. This study offers a scientific basis for pavement maintenance decisions and advances intelligent, precise pavement detection technology, crucial for building a safe and efficient transportation system.

Keywords: pavement distress detection; YOLOv12; DySample upsampling; NWD Loss; deep learning

1. Introduction

Pavements are vital for urban development, transportation, and economic and social activities, with their condition directly influencing traffic safety, economic growth, and quality of life. As socio-economic development and urbanization accelerate, pavement networks are expanding and becoming more complex, leading to increased damage that challenges sustainability [1]. Pavement distress, including various cracks, repair areas, and potholes, stems from design flaws, construction issues, vehicle loads, and environmental factors like temperature changes and rainwater erosion. Delayed repairs can shorten pavement lifespan, raise costs, and threaten driving safety.

Traditional manual detection methods are inefficient, subjective, and prone to errors, disrupting traffic and limited to daytime operations, making it hard to detect nighttime or hidden issues. Emerging technologies such as drones, computer vision, and deep learning, particularly CNNs [2], offer new avenges for pavement inspection. Drones efficiently collect comprehensive pavement images with minimal traffic interference, while computer vision and deep learning enable accurate identification and classification of pavement diseases by automatically learning complex image features [3-4].

Recent advancements like the MN-YOLOv5 algorithm [5], SVRDD dataset [6], and MGB-YOLO model [7] mark a shift towards intelligent pavement inspection, improving accuracy and supporting real-time detection. However, challenges persist, including image quality variations under different weather conditions and material differences across pavement types, necessitating improved model generalization. By introducing the UAV-PDD2023 dataset [8], DySample upsampling [9], and NWD loss function [10], extensive experimentation optimizes model structures, enhancing adaptability and robustness in complex scenarios. This exploration provides a scientific basis for pavement maintenance decisions and advances pavement detection technology towards intelligence and precision, laying a solid foundation for a safe and efficient transportation system.

2. Materials and Methods

2.1 NWD Loss

NWD Loss, based on Wasserstein distance and normalized for scale insensitivity, accurately measures differences between small target bounding boxes and provides effective gradients even without overlap, significantly improving detection accuracy and model convergence, especially in dense small target scenarios. NWD calculates the Wasserstein distance between object distributions on the feature map to measure their similarity as a loss function.

In our method, we introduce a new loss function by incorporating NWD loss into the localization loss. The model uses both CIoU and NWD losses, with a 3:7 ratio (CIoU:NWD), determined as optimal through preliminary experiments.

2.2 DySample upsampling

DySample, a lightweight and efficient dynamic upsampling method, replaces it due to its dynamic sampling mechanism, point sampling design, and lightweight nature. DySample adaptively adjusts upsampling based on input content, enhancing adaptability. Its point sampling avoids complex convolutions, making it resource-efficient and easy to implement. With fewer parameters, lower computational load, and reduced memory usage, DySample excels in dense prediction tasks, preserving feature details and improving detection accuracy, especially for small objects and complex scenes. The DySample operator's core is the sample point generator, which creates an offset combined with the original position to form the sample set, alleviating overlapping offset issues.

For the input feature map, a sample set is generated by the sampling point generator, and then the grid sample function is utilized to resample the input features, resulting in a new feature map.

3. Comparison of the Results of Different Models

The results are presented in Table 1.On the UAV-PDD2023 dataset, YOLOv12-ND achieves superior performance with 83.9% mAP50 and 58.4% mAP50-95, excelling particularly in categories like Repair (93.7%) and Longitudinal crack (89.0%). Older models such as YOLOv9 lag behind (68.7% mAP50) due to architectural constraints, while YOLOv11 and YOLOv10, though competitive in some areas, do not match YOLOv12-ND's overall accuracy. Both DySample upsampling and NWD loss individually boost performance, but their combination delivers the best results, effectively balancing computational efficiency and feature representation.

Table 1. Performance comparison of YOLO models

Model	Alligator crack	Longitudinal crack	Oblique crack	Pothole	Repair	Transverse crack	mAP50	mAP50-95)
Yolov12	83.8	86	71.7	54.2	89	84	78.1	49.6
Yolov12-DySample	88.7	88.1	78.2	57.3	91.6	86.8	81.8	54.7
Yolov12-NWD	85.3	88.8	76.8	55.2	88.5	87	80.3	56
Yolov12-ND	90.6	89	77.7	65.1	93.7	87.5	83.9	58.4
Yolov11	85.6	86.8	73.8	55.8	93.3	86.3	80.3	53.3
Yolov10	82.5	84.3	72.5	55	86.9	84.4	77.6	53.1
Yolov9	71.6	74.7	62.9	36.1	88.8	78.1	68.7	40.7

4. Discussion

This article proposes an improved YOLOv12 model-based road damage detection algorithm to address the inefficiency and inaccuracy of traditional detection methods. The YOLOv12-dl model, integrating DySample dynamic upsampling and NWD Loss function, achieves excellent performance on the UAV-PDD2023 dataset, with mAP50 and mAP50-95 reaching 83.9% and 58.4%, respectively, outperforming other mainstream algorithms. DySample enhances feature detail capture, especially for small targets and complex scenes, while NWD Loss improves detection accuracy and convergence speed by addressing the limitations of traditional IoU Loss for small target displacements.

This research provides an efficient and precise tool for road maintenance and offers new insights for intelligent pavement detection. Future work will focus on optimizing the model to enhance its generalization across diverse weather conditions and road types, supporting practical road maintenance and contributing to a safe and efficient transportation system.

Acknowledgments

This paper was supported by Research on Multimodal Sentiment Computation for Social Media Images and Texts.

References

[1] Cao W, Liu Q, He Z. Review of pavement defect detection methods[J]. IEEE access, 2020, 8: 14531-14544.
[2] Cao W, Liu Q, He Z. Review of pavement defect detection methods[J]. IEEE access, 2020, 8: 14531-14544.
[3] Du Y, Pan N, Xu Z, et al. Pavement distress detection and classification based on YOLO network[J]. International Journal of Pavement Engineering, 2021, 22(13): 1659-1672.
[4] Kheradmandi N, Mehranfar V. A critical review and comparative study on image segmentation-based techniques for pavement crack detection[J]. Construction and Building Materials, 2022, 321: 126162.
[5] Guo, G., Zhang, Z. Road damage detection algorithm for improved YOLOv5. Sci Rep 12, 15523 (2022). https://doi.org/10.1038/s41598-022-19674-8
[6] Ren, M., Zhang, X., Zhi, X. et al. An annotated street view image dataset for automated road damage detection. Sci Data

11, 407 (2024). https://doi.org/10.1038/s41597-024-03263-7
[7] Pang, D., Guan, Z., Luo, T. et al. Real-time detection of road manhole covers with a deep learning model. Sci Rep 13, 16479 (2023). https://doi.org/10.1038/s41598-023-43173-z
[8] Yan H, Zhang J. UAV-PDD2023: A benchmark dataset for pavement distress detection based on UAV images[J]. Data in Brief, 2023, 51: 109692.
[9] Liu W, Lu H, Fu H, et al. Learning to upsample by learning to sample[C]//Proceedings of the IEEE/CVF international conference on computer vision. 2023: 6027-6037.
[10] Wang J, Xu C, Yang W, et al. A normalized Gaussian Wasserstein distance for tiny object detection[J]. arXiv preprint arXiv:2110.13389, 2021.
[11] Tian Y, Ye Q, Doermann D. Yolov12: Attention-centric real-time object detectors[J]. arXiv preprint arXiv:2502.12524, 2025.
[12] Dao T, Fu D, Ermon S, et al. Flashattention: Fast and memory-efficient exact attention with io-awareness[J]. Advances in neural information processing systems, 2022, 35: 16344-16359.