Lightweight Road Crack Detection Algorithm Integrating Knowledge Distillation and Fractal Dimension

Zihao Zhao [a], Jun Cheng [b] and Yuanyuan Li [c] *

School of Railway Intelligent Engineering, Dalian Jiaotong University, HuangheRoad, Dalian ,116028, Liaoning, China

[a]1078758085@qq.com; [b]c_jun6@163.com; * [c]171979166@qq.com

Abstract. Accurate road crack detection plays a vital role in traffic safety and infrastructure maintenance. However, mainstream deep learning-based detection models suffer from high computational costs and are difficult to deploy on resource-constrained edge devices. This paper proposes FDKDNet, a lightweight road crack detection network that integrates knowledge distillation and fractal dimension analysis. The ResNet101-based teacher model is enhanced with asymmetric convolutions to effectively capture slender crack features. The ResNet18-based student model utilizes Gaussian process-guided Bayesian optimization to adaptively optimize the distillation temperature and realize efficient knowledge transfer. Furthermore, a novel fractal feature module is designed to quantify the morphological complexity of cracks within anchor boxes. Experimental results on the Crack500 dataset show that FDKDNet achieves 90.3% detection accuracy with only one-fourth the parameters of the teacher model, and outperforms existing state-of-the-art methods. The proposed method provides an efficient and practical solution for road crack detection on resource-limited platforms.

Keywords: Fractal dimension; Knowledge distillation; Residual network; Road crack detection

1 Introduction

Road cracks threaten driving safety and require timely detection. While deep learning has advanced crack detection, existing models [1-5] are computationally heavy and difficult to deploy on edge devices. Knowledge distillation (KD) compresses models by transferring knowledge from a teacher to a student network [6-9]. However, traditional KD relies on a fixed temperature parameter, which cannot adapt to irregular crack shapes. Fractal dimension quantifies complex morphological features [10-12], but most methods compute it on full images, causing background interference and redundancy.

To tackle the aforementioned challenges, this paper presents FDKDNet, a lightweight road crack detection network that integrates adaptive knowledge distillation and fractal dimension enhancement. The core contributions of this work are summarized as follows:

N. Malik et al. (eds.), *Proceedings of the 6th International Conference on Public Management and Intelligent Society (PMIS 2026)*, Advances in Engineering Research 301,
https://doi.org/10.2991/978-94-6239-721-7_23

1. We design an enhanced teacher–student distillation framework. The teacher model adopts asymmetric convolution to better capture slender crack features, while a Gaussian process-based Bayesian optimization strategy is proposed to adaptively search for the optimal distillation temperature, realizing efficient and stable knowledge transfer from the teacher to the lightweight student model.
2. We propose a novel anchor-constrained fractal dimension module, which quantifies the morphological complexity of cracks only within candidate regions to suppress background interference and supplement geometric prior knowledge, significantly improving the detection performance for complex and fine cracks.
3. Extensive experiments on the Crack500 dataset show that the proposed FDKDNet achieves 90.3% detection accuracy with only 11.7M parameters, outperforming mainstream state-of-the-art methods and showing strong practicality for deployment on resource-constrained devices.

2 Related Work

2.1 Knowledge Distillation

Knowledge distillation is a typical model compression approach that transfers structural and semantic knowledge from a large teacher model to a small student model. The core idea is to use soft labels produced by the teacher to guide the training of the student, so that the student can retain most of the teacher's representation ability with much fewer parameters. In detection tasks, traditional knowledge distillation mainly focuses on feature imitation and logit matching. However, most existing methods use a fixed distillation temperature, which cannot adaptively adjust knowledge transfer intensity for targets with irregular shapes like road cracks. As a result, the student model often fails to fully acquire the teacher's ability in recognizing slender and tiny targets, leading to obvious performance gaps.

2.2 Fractal Dimension in Image Analysis

Fractal dimension is a mathematical tool used to measure the complexity and irregularity of geometric structures. It can effectively describe the self-similarity and texture complexity of objects that are difficult to characterize with traditional Euclidean geometry. In crack detection, fractal features can reflect the morphological complexity of crack regions, helping distinguish real cracks from interference such as stains, shadows and joints. However, conventional fractal calculation is usually performed on the whole image, which introduces heavy background noise and redundant computation. Such methods are not compatible with lightweight networks and cannot support real-time detection on edge devices.

3 Method

3.1 The Proposed Model FDKDNet

The overall framework of the proposed FDKDNet is illustrated in Fig. 1. Following a teacher-student knowledge distillation architecture, the model contains four core components: an enhanced teacher model, a lightweight student model, an adaptive knowledge distillation module, and an anchor-constrained fractal dimension module. The teacher model extracts high-quality crack features and generates informative soft labels. The student model learns jointly from both the soft labels and the ground truth. The adaptive distillation module enables efficient and dynamic knowledge transfer, while the fractal module strengthens the feature representation for complex crack morphologies.

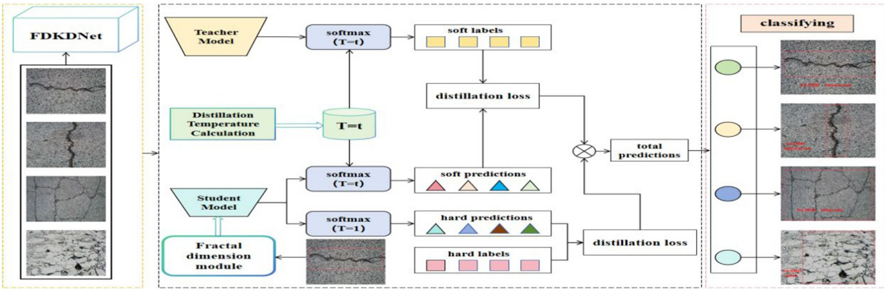

Fig. 1. Structure diagram of the FDKDNet model

The teacher model is built on ResNet101 and improved by asymmetric convolutions. The asymmetric convolution expands the receptive field and strengthens the ability to capture slender and linear crack features, providing high-quality supervision for the student model.The student model is constructed based on ResNet18, which has a shallow and compact structure to ensure low parameters and fast inference. Through residual connections, the student model maintains stable feature extraction ability. Under the supervision of the teacher model and the assistance of fractal features, the student model achieves strong detection performance with high efficiency.

3.2 Improved Distillation Temperature Coefficient

In experiments, the distillation temperature T is usually taken as an integer value based on people's experience, such as 1, 5, 10, etc. However, such a value selection lacks rigorous mathematical reasoning and may miss the optimal distillation temperature or a certain optimal non-integer temperature. Therefore, we have improved the selection of the distillation temperature.

Our improved method for selecting distillation temperature is:

(1) Define the objective function:Let the objective function be f(T), representing the accuracy of the student model on the validation set at the distillation temperature T. Our goal is to find the value of T that maximizes f(T). That is, $T^*=\text{argmax}_T\ f(T)$.

(2) Gaussian process modeling:The Gauss process is a probabilistic model used to describe the distribution of a function f(T). For any set of temperature values T_1, T_2,... T_n, corresponding function values $f(T_1)$, $f(T_2)$,... $f(T_n)$ follows a multivariate normal distribution:

$$\begin{bmatrix} f(T_1) \\ f(T_2) \\ \vdots \\ f(T_n) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu(T_1) \\ \mu(T_2) \\ \vdots \\ \mu(T_n) \end{bmatrix}, \begin{bmatrix} k(T_1,T_1) & k(T_1,T_2) & \cdots & k(T_1,T_n) \\ k(T_2,T_1) & k(T_2,T_2) & \cdots & k(T_2,T_n) \\ \vdots & \vdots & \ddots & \vdots \\ k(T_n,T_1) & k(T_n,T_2) & \cdots & k(T_n,T_n) \end{bmatrix} \right) \tag{1}$$

Among them, μ(T) is the mean function, usually set as 0 and 1, and k(Ti,Tj) is the covariance function:

$$k(T_i,T_j)=256exp\left(-\frac{(T_i-T_j)^2}{50}\right) \tag{2}$$

(3) Acquisition function

The acquisition function is used to select the next most promising temperature value for the experiment based on the current Gaussian process model. The collection function is Expected Improvement (EI):

$$EI(T) = E[\,max(f(T) - f(T^+), 0)] \tag{3}$$

Among them, $f(T^+)$ is the currently known maximum function value.

(4) Optimization process

① Initialization: Randomly select several initial temperature values within the temperature range T_1, T_2,... T_n ,conduct experiments respectively,obtain the corresponding function values $f(T_1)$, $f(T_2)$,... $f(T_n)$.

② Iterative optimization:Based on the existing temperature values and function values, update the Gaussian process model to obtain the new mean function μ(T) and covariance function $k(T_i,T_j)$. Using the acquisition function (EI), select the next most promising temperature value Tnew within the temperature range. Conduct the experiment to obtain $f(T_{new})$. Add the new temperature values and function values to the existing dataset, and repeat the above steps until the stopping conditions are met (such as reaching the maximum number of iterations or the improvement of the function values being less than a certain threshold).

3.3 Fractal Dimension Calculation

The box-counting method is employed to calculate the fractal dimension. To reduce background interference and redundant computation, fractal analysis is only performed within the predicted anchor boxes instead of the whole image.In this work, the box sizes are uniformly sampled from 2 to 32 pixels in exponential steps (i.e., 2, 4, 8, 16, 32), covering fine to coarse scales and ensuring stable estimation of fractal dimension. Multi-scale analysis is naturally realized by using boxes of different sizes, which captures the self-similarity of cracks across scales and enhances the robustness of morphological description.The fractal module only conducts simple counting and linear regression within small anchor regions, so its extra computational overhead is negligible (less than 3% increase in FLOPs). The lightweight design makes it compatible with real-time inference on resource-constrained devices.

The formula for calculating the fractal dimension by the box counting method is:

$$D = \lim_{\epsilon \to 0} \frac{\log(N(\epsilon))}{\log(1/\epsilon)} \tag{4}$$

In this formulation, D represents the fractal dimension, where N(ϵ) denotes the number of boxes of size ϵ required to cover the entire fractal object. As ϵ (the side length of the box) approaches zero, we obtain increasingly precise dimensional measurements.

4 Experiment and Analysis

4.1 Dataset

Two representative road crack datasets are used in our experiments: Crack500 and GAPs384.Crack500 contains a large number of pavement images with various types of cracks, covering common scenes in real road environments. GAPs384 focuses on complex and challenging pavement conditions, which is suitable for verifying the robustness of detection models.

4.2 Experimental Process

All Crack500 images are resized to 224×224×3 with data augmentation (random flipping, rotation, brightness adjustment). The dataset is split 7:2:1 for training/validation/test.The student is trained using Adam (lr=0.2, batch size 24, 49 epochs) with early stopping. For distillation temperature optimization, we use Bayesian optimization with T search range [1~20], starting with 5 random initial values and performing 15 iterations using Expected Improvement (EI). The optimal temperature is T* = 6.8, converging after about 12 iterations. All baselines (YOLOv7, YOLOv8, DETR, BiSwinTransformer, DeepCrack) are retrained under identical settings.

4.3 Evaluation Index

To quantitatively evaluate detection performance, four commonly used metrics are adopted: Accuracy, Precision, Recall, and F1-score. These indicators jointly reflect the model's ability to correctly identify cracks and avoid false detections. In addition, the number of model parameters is used to evaluate the lightweight level of the network.

4.4 Comparative Experiment

To verify the superiority of the proposed FDKDNet, we compare it with several mainstream crack detection methods, including YOLOv7, YOLOv8, DETR, BiSwinTransformer, DeepCrack, and the baseline ResNet18.Experimental results on the Crack500 dataset show that our method maintains a very lightweight parameter scale of only 11.7M while achieving higher accuracy and F1-score than all comparison algorithms. FDKDNet obtains 90.3% accuracy and 88.15% F1-score on Crack500.A slight performance gap remains in detecting thin and tiny cracks, as the teacher model captures finer textural details while the lightweight student model has limited representation capacity. Our adaptive distillation and fractal dimension module effectively narrow this gap by enhancing subtle crack feature learning.In summary, the proposed model achieves a better balance between lightweight characteristics and detection performance, making it more suitable for practical deployment on resource-constrained devices.The experimental results are shown in Table 1.It should be noted that a comprehensive comparison of inference speed (FPS) and memory usage is not provided in this study. However, parameter count serves as a primary indicator of model complexity and deployment cost. FDKDNet achieves the lowest parameter count (11.7M) among all compared methods.This strongly supports our claim that FDKDNet is lightweight and suitable for resource-constrained devices. Full FPS and memory benchmarks will be included in future work.

Table 1. Results of the Crack500 Comparative Experiment

Algorithm	Precision	Recall	Accuracy	F1 Score	Parameter
YOLOv7	86%	84.5%	88.4%	85.26%	39.8M
YOLOv8	87%	85%	89.1%	86%	28.6M
DETR	85%	83.4%	87.6%	84.2%	44.3M
BiSwinTransformer	88%	84.3%	89.4%	86.1%	44.8M
DeepCrack	87%	84.2%	88.36%	85.6%	36.6M
ResNet101 (teacher)	88%	85.6%	90.4%	86.8 %	44.5M
ResNet18(student)	86%	84.6%	88.2%	85.32%	11.7M
Ours	**84%**	**87.22%**	**90.3%**	**88.15%**	**11.7M**

4.5 Ablation Experiment

Ablation experiments on the Crack500 dataset verify the effectiveness of each module in FDKDNet. The baseline is the original ResNet18, and three modules are evaluated: adaptive distillation temperature, fractal dimension module, and asymmetric convolution.Results demonstrate that each module steadily boosts performance with clear synergistic effects. The full model achieves 90.3% accuracy and 88.15% F1-score, outperforming the baseline by 2.1% in accuracy.To ensure the improvements are statistically reliable, we perform paired t-tests under a 95% confidence level. Results confirm that all performance gains are significant ($p < 0.05$) and not caused by random fluctuations.The fractal dimension module contributes the most significant gain, while adaptive temperature and asymmetric convolution further stabilize detection. Experimental results are shown in Table 2.

Table 2. Results of Crack500 Ablation Experiment

Improve the distillation temperature	Fractal dimension module	Asymmetric Convolution	Accuracy	F1 score
		√	88.2%	86.5%
√			88.6%	86.7%
	√		89%	87%
√	√		89.7%	87.8%
√		√	89.1%	87.6%
	√	√	89.5%	87.65%
√	√	√	90.3%	88.15%

5 Conclusion

This paper proposes FDKDNet, a lightweight crack detection model integrating asymmetric convolution, adaptive distillation, and fractal dimension enhancement. Experiments show it reduces parameters while improving accuracy. Future work includes richer datasets, attention mechanisms, and interdisciplinary research.

Disclosure of Interests

The authors declare there are no financial interests, commercial affiliations, or other potential conflicts of interest that have influenced the objectivity of this research or the writing of this paper.

References

1. Zhang L, Yang F, Zhang Y D, et al. Road crack detection using deep con-volutional neural network[C]//2016 IEEE International Conference on Image Pro-cessing (ICIP). Phoenix, AZ, USA: IEEE, 2016.
2. Qu Z, Mei J, Liu L, et al. Crack detection of concrete pavement with cross-entropy loss functionand improved VGG16 network model[J]. IEEE Access, 2020.
3. Zhang K, Cheng H D. CrackGAN: Pavement crack detection usi-ng partially accurate ground truths based on generative adversarial learning[J]. IEEE Transactions on Intelligent Transportation Systems, 2020.
4. Pen Y F, Ji Y. Road Crack Detection Algorithm Based on Improved YOL-Ov8[C]//2023 International Conference on Artificial Intelligence and Computer Applications (ICAICA). Dalian, China: IEEE, 2023.
5. Zou, Q.; Zhang, Z.; Li, Q.; Qi, X.; Wang, Q.; Wang, S. DeepCrack: Learning hierarchical convolutional features for crack detection. IEEE Trans. Image Process. 2019.

6. Hinton G, Vinyals O, Dean J. Distilling the knowledge in a neural networ-k[J].arXiv preprint arXiv:1503.02531, 2015.
7. Zheng Z, Ye R, Wang P, et al. Localization distillation for dense object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2022.
8. Zheng Z, Ye R, Wang Y, et al. Localization Distillation for Dense Object Detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024.
9. Yang Z, et al. Focal and Global Knowledge Distillation for Detectors[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022.
10. Zhu D, Tang A, Wan C, Zeng Y. Investigation on the flexural toughness evaluation method and surface cracks fractal characteristics of polypropylene fiber reinforced cement-based composites. J. Build. Eng. 2021.
11. An Q, Chen X, Wang H, Yang H, Yang Y, Huang W, Wang L. Segmentation of Concrete Cracks by Using Fractal Dimension and UHK-Net. Fractal Fract. 2022.
12. J. Yan, Y. Sun, S. Cai, X. Hu. An improved boxcounting method to estimate fractal dimension of images[J]. Journal of Applied Analysis & Computation. 2016.