DynaEdge-Net: Dynamic Feature Gating and Edge Enhancement for Precise Road Crack Segmentation

Lingjie Cao[1], Hailing Sun[2], Sunwen Du[2,*]
[1]College of Civil Engineering, Taiyuan University of Technology, Taiyuan, 030024, China
[2]College of Geological and Surveying Engineering, Taiyuan University of Technology, Taiyuan, 030024, China
clj18135560334@163.com; shldezhanghu@outlook.com; [*]dusunwen@tyut.edu.cn

***Abstract*—Automatic detection of road cracks is essential for long-term safety maintenance of roads, bridges, and other infrastructure. Although existing deep learning-based crack segmentation methods have improved detection accuracy, challenges remain in terms of high computational complexity and inadequate capture of fine cracks and edge details. To address these issues, this study proposes an enhanced UNet-based architecture, termed DynaEdge-Net. In the encoder and decoder stages, a residual detail enhancement block (RDEB) and a cascaded group attention (CGA) module are incorporated to strengthen edge feature representation and focus on critical regions, respectively. In the skip connections, a group-wise dynamic gating (GDG) module is introduced to adaptively suppress background noise and optimise feature transmission. During decoding, a dynamic upsampling (DySample) strategy replaces conventional interpolation, enabling high-fidelity reconstruction of crack structures. Experimental results show that DynaEdge-Net achieves IoU, F1-score, boundary F1-score, and recall rates of 82.34 %, 90.83 %, 83.27 %, and 90.12 %, respectively, outperforming several state-of-the-art road segmentation algorithms. The proposed method not only improves the continuity and accuracy of crack extraction, but also demonstrates strong robustness and generalisation capability, providing a reliable solution for intelligent inspection and maintenance of transportation infrastructure.**

***Index Terms*—Road crack; Semantic segmentation; GDG; Edge enhancement.**

I. INTRODUCTION

Roads, as a crucial component of transportation infrastructure, have structural integrity that directly affects the safety and efficiency of transportation systems. Pavement cracks are the most common and destructive type of surface distress, caused by multiple factors such as traffic loads, environmental influences (e.g., temperature variations and water erosion), and material ageing. Cracks not only affect pavement aesthetics but also accelerate structural deterioration. If not detected and repaired in a timely manner, they may cause severe damage, posing safety hazards and increasing maintenance costs throughout the pavement life cycle [1]–[3]. Therefore, efficient and accurate automatic crack detection is essential to ensure long-term serviceability of roads [4].

Early crack detection methods relied mainly on manual inspections and handcrafted image processing techniques, such as Canny edge detection [5], Sobel operators [6], and local thresholding [7]. Although simple to implement, these methods are highly sensitive to changes in illumination and complex textures, making it difficult to achieve stable performance in real-world road environments. With the development of computer vision and deep learning, end-to-end segmentation networks have gradually become the mainstream research approach. Their powerful feature extraction and pattern recognition capabilities significantly improve automation and detection accuracy. For example, fully convolutional networks (FCN) [8] first introduced a fully convolutional structure for pixel-level prediction, laying the foundation for deep learning-based crack segmentation [9], [10]. SegNet improved boundary localisation through an encoder-decoder architecture [11], while U-Net effectively fused shallow details with deep semantic information via skip connections, significantly enhancing crack and defect detection performance [12]. Based on U-Net, various improved approaches have been proposed: Liu, Cao, Wang, and Wang [13] successfully applied it to concrete crack detection, while Augustauskas and Lipnickas further improved pavement defect extraction accuracy by adjusting network depth, filter size, and incorporating ASPP [14] and attention gate mechanisms [15].

However, current deep learning-based methods still have notable limitations. To achieve high accuracy, many models stack deep and complex network structures [16], [17], resulting in high computational costs that are unsuitable for resource-constrained scenarios such as edge devices or in-vehicle systems. Moreover, the complex morphology and inconspicuous features of fine cracks often cause information loss, missed detections, and blurred boundaries. For example: Dense-DeepLabV3+ enhances small crack detection via dense connections, but still has blurred edges and discontinuities in low-contrast regions [18], [19]; ConvNeXt-

Manuscript received 12 April, 2025; accepted 11 June, 2025.
The work was supported by the National Natural Science Foundation of China (Grant No. U21A20107), the National Natural Science Foundation of China (Grant No. 52525401), and the Fundamental Research Program of Shanxi Province (202203021211156).Remove The work was supported by the National Natural Science Foundation of China (Grant No. U21A20107), the National Natural Science Foundation of China (Grant No. 52525401), and the Fundamental Research Program of Shanxi Province (202203021211156).

FPN [20] maintains good overall connectivity, but is prone to false detections due to high-contrast noise [21]; PIPNet [22], though lightweight and efficient, struggles to preserve the continuity and sharpness of slender cracks due to its heatmap regression strategy. These shortcomings significantly limit the applicability of such methods in practical road maintenance.

To address the aforementioned challenges, this study proposes a novel road crack segmentation network termed DynaEdge-Net. Built upon the U-Net architecture, this network integrates dynamic feature gating and edge enhancement mechanisms, aiming to achieve high-precision and efficient crack detection [23] while effectively mitigating the issue of detail loss in complex scenarios. The main contributions of this paper are summarised as follows:

1. This study designs a DynaEdge-Net based on the UNet framework, which addresses the problems of insufficient edge feature capture and redundant feature transmission in road crack segmentation through the collaborative integration of multiple modules.
2. Several innovative modules are introduced into the network structure to enhance the model's ability to capture crack details and suppress noise. Specifically, a residual detail enhancement block (RDEB) and a cascaded group attention (CGA) module are designed to strengthen crack edge features and focus on critical regions, respectively. In the skip connections, a group-wise dynamic gating (GDG) module is incorporated to optimise information transmission and suppress background noise via a dynamic gating mechanism. Additionally, DySample is utilised to replace traditional upsampling, enabling accurate content-aware reconstruction and effectively improving the segmentation precision of fine cracks and boundaries.
3. We integrated four public datasets to construct a diverse mixed dataset, upon which comprehensive ablation studies and comparative experiments were conducted. The results demonstrate that DynaEdge-Net outperforms state-of-the-art segmentation algorithms in several key metrics (such as Intersection over Union (IoU), F1-score, and boundary F1-score), verifying its high robustness and generalisation capability under different pavement conditions, environmental interferences, and crack morphologies.

II. PROPOSED METHOD

A. Model Architecture

The proposed DynaEdge-Net architecture, illustrated in Fig. 1, is developed in the UNet framework to address the challenges of insufficient edge feature capture and redundant feature transmission in road crack segmentation. By integrating multiple modules in a collaborative manner, the network achieves accurate and robust segmentation results. During the feature extraction stage, both the encoder and decoder at each level incorporate the RDEB and CGA modules. The RDEB combines a residual structure with multidirectional differential convolutions to enhance the extraction of crack edges and gradient details. The CGA employs a feature-splitting and cascading strategy to increase the diversity of attention heads and reduce channel redundancy, enabling the network to focus more effectively on critical crack regions. In skip connections, a GDG module is introduced, which leverages group-wise feature processing and dynamic gating to adaptively adjust the contribution of different channel groups, highlighting crack-related features while suppressing background noise and improving feature transmission efficiency. In the decoder, the DySample module replaces traditional interpolation for upsampling. By dynamically generating sampling offsets based on local feature content, DySample enables content-aware high-resolution reconstruction, effectively restoring crack edges and spatial structures while reducing blurring and lowering computational cost.

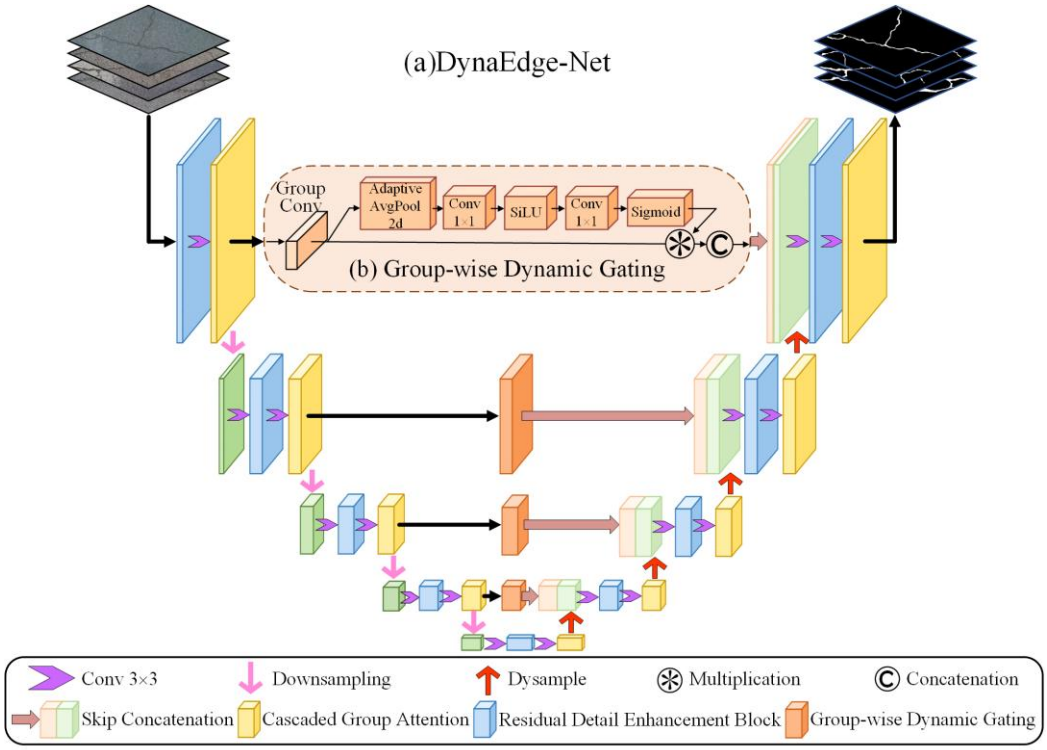


Fig. 1. Network architecture of DynaEdge-Net: (a) The overall structure; (b) GDG block.

B. Cascaded Group Attention (CGA)

The CGA module [24] is designed to address the computational redundancy inherent in multihead self-attention (MHSA). In CGA, the input feature map X_i is divided along the channel dimension into h independent segments X_{ij}, with each attention head processing only its corresponding segment. The attention computation for each head, as expressed in (1), is thereby explicitly decoupled across heads, which substantially reduces computational cost and parameter redundancy. Furthermore, CGA incorporates a cascading mechanism, in which the input of each head is augmented with the output of the preceding head, allowing the features to be progressively refined. The outputs of all attention heads are then concatenated and passed through a linear projection layer to restore the original feature dimension, as illustrated in Fig. 2

$$\tilde{X}_{ij} = Attn\left(X_{ij}W_{ij}^{Q}, X_{ij}W_{ij}^{K}, X_{ij}W_{ij}^{V}\right), \quad (1)$$

where $W_{ij}^{Q}, W_{ij}^{K}, W_{ij}^{V}$ are the projection matrices for the j-th head, mapping the segmented features into the Query (Q), Key (K), and Value (V) subspaces, respectively. By employing feature segmentation, CGA reduces the computational load and parameter count for each attention head, while the cascading operation fosters inter-head feature interaction. This design enhances the diversity of the attention maps and significantly improves computational efficiency without compromising model performance.

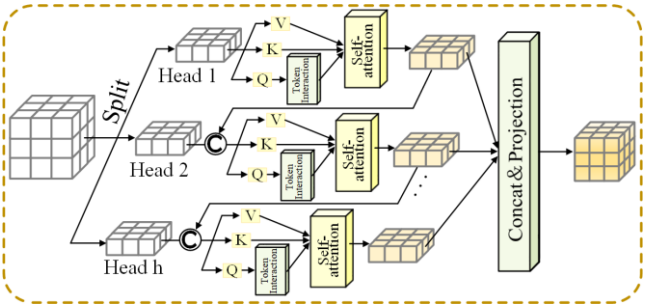


Fig. 2. The architecture of CGA.

C. Residual Detail Enhancement Block (RDEB)

RDEB [25] is a residual module designed to enhance the ability to extract edge details from images. Based on the classical residual structure, this module introduces a "detail enhancement convolution (DEConv)" mechanism to strengthen the expression of edge information. DEConv consists of five directional convolution kernels: vanilla convolution (VC), central difference convolution (CDC), angular difference convolution (ADC), horizontal difference convolution (HDC), and vertical difference convolution (VDC). Through parallel combination, the module can comprehensively perceive the morphological changes of cracks in different directions, thus improving the ability to model edges and angles. The input feature map first extracts local contextual information through two 3×3 vanilla convolutions. The encoded features are then fed into the DEConv module, which is responsible for perceiving feature changes in different directions to further enhance edge information. Subsequently, a dual residual mechanism is used to fuse detail-enhanced features with input features, to strengthen the transmission of information flow and alleviate the gradient vanishing problem in deep-layer training. RDEB avoids information loss and gradient vanishing through residual connections and strengthens details such as crack edges with multidirectional differential convolutions. This enables the model to accurately capture subtle structures like tiny cracks and blurred edges while retaining high-level semantics, making it suitable for pixel-level road crack segmentation (Fig. 3).

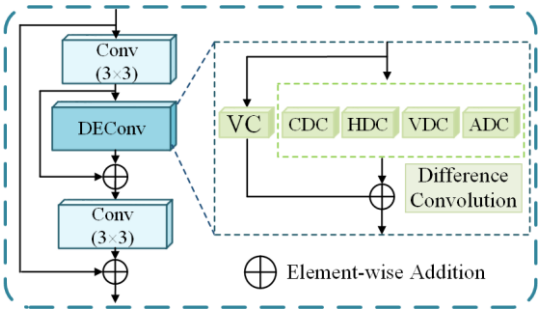


Fig. 3. The architecture of RDEB.

D. Group-wise Dynamic Gating (GDG)

GDG [26] is designed to dynamically adjust the weights of channel features, thereby enhancing the model's ability to express key features. The input features are divided into multiple groups by channels in GDG. Each group is subjected to a separate convolution operation, and adaptive average pooling combined with convolution is used to generate the corresponding gating weights. These weights are scaled to the range [0, 1] through the Sigmoid activation function, serving as adjustment factors for the output of each group. After the convolution output of each subgroup is multiplied by its corresponding gating weight, all weighted features are concatenated and fused to form the final output of the module. Its design enables the selective activation of each channel group on the feature map, thus improving the ability to express fine-grained structures (Fig. 1(b)).

E. Dynamic Upsampling (DySample)

Traditional upsampling methods, such as bilinear interpolation or transposed convolution, often suffer from information loss or artifact issues. In contrast, DySample [27], [28] is a mechanism that achieves efficient feature upsampling by dynamically learning the positions of sampling points, rather than relying on content-aware convolution kernels used by traditional dynamic upsamplers.

DySample adopts a "bilinear initialization" strategy to set initial sampling positions. Each pixel in the low-resolution feature map X is evenly assigned to s^2 initial positions, ensuring that the result is equivalent to bilinear interpolation under the zero-offset state. The input feature map is used to generate base offsets through a linear projection. Initially, offsets are constrained by a static range factor to avoid overlapping of adjacent sampling points. When introducing dynamic range factors, two linear layers are used to generate range modulation terms and base offsets, respectively, limiting the offset range to [0, 0.5], as shown in (2) $\mathcal{O}X$

$$\mathcal{O} = 0.5 \cdot \sigma\left(\text{linear}_1(X)\right) \cdot \text{linear}_2(X), \quad (2)$$

where $\mathcal{O}$ denotes the offset, $X \in \mathbb{R}^{C\times H\times W}$ represents the

input low-resolution feature map, $\text{linear}_1(X)$ and $\text{linear}_2(X)$ are two independent linear projection layers, which are used to generate the range modulation term and base offset, respectively, and $\sigma(\cdot)$ is the sigmoid function.

The feature map is divided into $\mathcal{G}$ independent groups along the channel dimension, and each group shares the sampling set $\mathcal{S} = \mathcal{G} + \mathcal{O}$ to reduce computational redundancy. Finally, the normalised sampling coordinates are processed using bilinear interpolation, and the high-resolution output feature map $X' \in \mathbb{R}^{C \times sH \times sW}$ is obtained through the grid_sample function, as shown in (3)

$$X' = \text{grid\_sample}(X, \mathcal{S}). \tag{3}$$

III. Experimental Results and Analysis

A. Dataset

This study integrates four authoritative public road crack datasets to establish an evaluation framework: CrackForest [29], [30] (mainly focussing on urban street scenes in Beijing, including 118 crack images affected by shadows and oil stains), DeepCrack [31], [32] (focussing on multiscale crack features on concrete pavements, with 537 pixel-level finely annotated images), AigleRN [21] (collected by multiple imaging devices, covering 185 samples affected by illumination variations and noise), and Crack500 [33], [34] (primarily featuring urban asphalt pavement scenes, including 500 high-resolution original images). Based on the above datasets, we constructed a mixed dataset. After unified cropping, a total of 11,298 crack images with a resolution of 448×448 pixels were obtained.

To fully verify the model's generalisation ability, a cross-dataset mixed training strategy was adopted, with the data split into training, validation, and test sets at a ratio of 7:2:1.

During training, data augmentation operations were introduced, including random rotation (±30 °), brightness jitter (±20 %), Gaussian noise (σ = 0.05), and random occlusion (maximum occlusion ratio of 15 %), to simulate complex interference factors in real detection environments.

Figure 4 shows some representative dataset samples. The columns present three typical crack morphologies in single crack, fractured cracks, and reticular cracks. Each crack morphology covers various scene conditions, including different pavement materials (e.g., cement, asphalt) and various environmental interferences (e.g., shadows, stains). This sample organisation method can simultaneously verify the model's adaptability to the diversity of crack morphologies and its robustness under complex scene interferences.

(a) (b) (c)

Fig. 4. Dataset: (a) Single crack; (b) Fractured crack; (c) Reticular crack.

B. Experimental Results and Analysis

1. Ablation experiments

To verify the effectiveness of each core component in DynaEdge-Net, ablation experiments were designed to analyse the model by disassembling it step by step. UNet was used as the baseline model, and five model variants were constructed by sequentially adding CGA, RDEB, GDG, and DySample modules. All models were trained in the same training and validation sets, and their performance in the test set was evaluated using IoU, Accuracy, Recall, F1, and FPS. The results are presented in Table I.

TABLE I. THE RESULTS OF THE ABLATION EXPERIMENT.

Model variant	IoU/%	Accuracy/%	Recall/%	F1/%	FPS
UNet	72.56	85.12	82.05	83.54	25.00
+CGA	75.21	87.33	84.12	85.71	21.50
+CGA + RDEB	77.89	89.01	86.34	87.62	18.20
+CGA + RDEB +GDG	79.65	90.23	87.89	89.06	17.00
DynaEdge-Net	82.34	91.56	90.12	90.83	15.00

As shown in Table I, with the gradual addition of core components, the model performance shows a continuous improvement trend. Meanwhile, the FPS gradually decreases (corresponding to the increase in the inference time), and the range of change is reasonable. This validates the rationality of each component in balancing accuracy improvement and efficiency.

Specifically, after introducing the CGA module into the baseline UNet model, the IoU increases from 72.56 % to 75.21 % (an improvement of 2.65 %) and the FPS drops from 25.00 to 21.50 (corresponding to increased inference time). This indicates that the cascaded group attention mechanism, through weighted optimisation of channel and spatial features, can effectively focus on crack regions and suppress background noise interference. It enhances feature expression ability, while only slightly increasing computational costs.

Based on this, after adding the RDEB module, the model's IoU increases further to 77.89 % (an improvement of 2.68 % compared with "UNet + CGA"), and the recall rate rises from 84.12 % to 86.34 % (an improvement of 2.22 %) . FPS drops to 18.20 (corresponding to a further increase in the inference time). This shows that the residual detail enhancement module, through multiscale feature fusion and residual connections, strengthens the ability to capture fine cracks

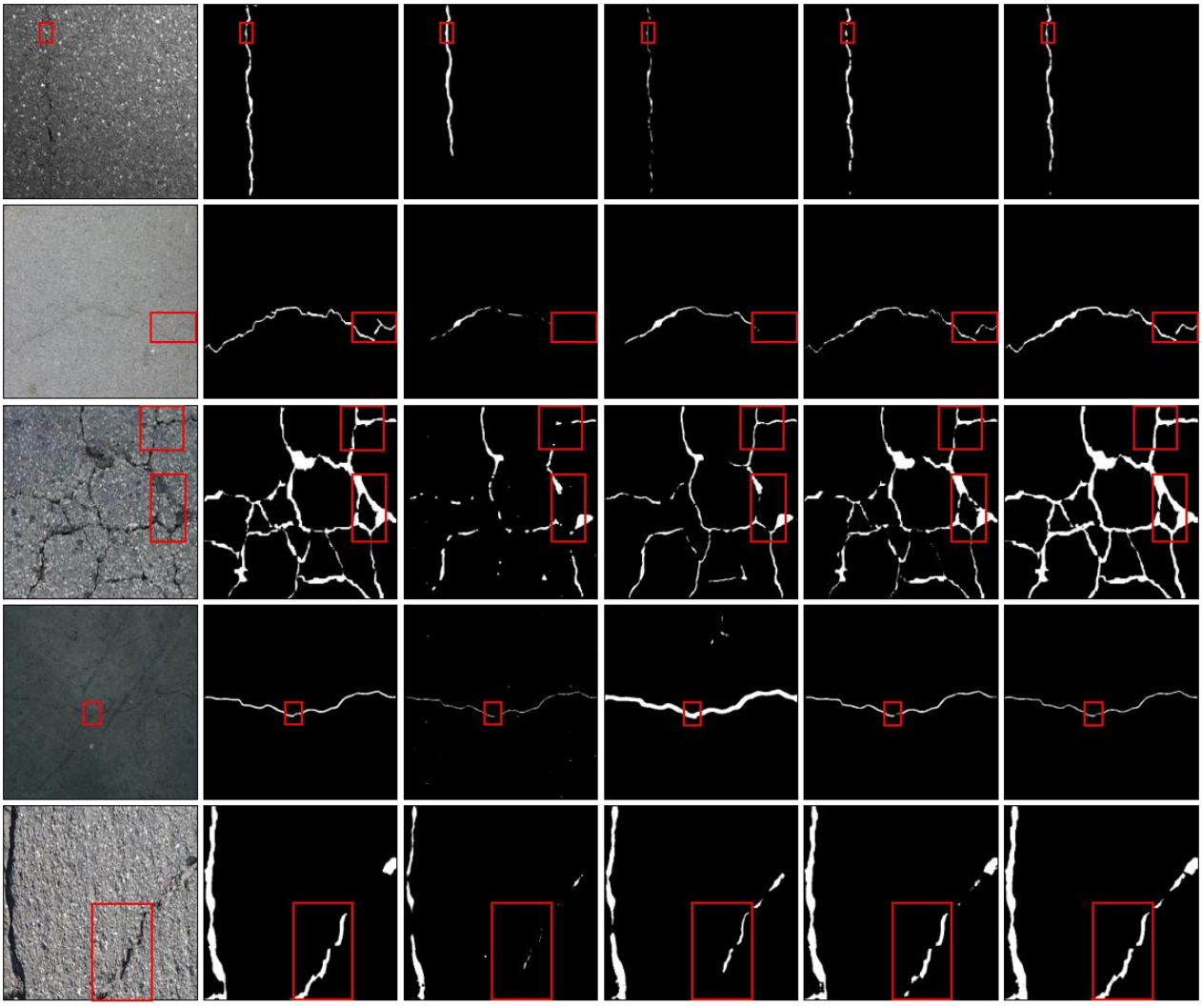
without introducing excessive redundant computations. It balances feature details and computational efficiency.

After adding the GDG module, the model IoU reaches 79.65 % (an improvement of 1.76 % compared with "UNet + CGA + RDEB"), and the accuracy increases from 89.01 % to 90.23 % (an improvement of 1.22 %). FPS drops to 17.00 (with a controllable increase in inference time), and the growth in the number of parameters remains within a reasonable range. This benefits from dynamic filtering and enhancement of feature channels by the group dynamic gating mechanism, which improves prediction accuracy while maintaining high computational efficiency. This further validates the effectiveness of the module design.

Finally, the complete DynaEdge-Net model, which incorporates the dynamic upsampling module, achieves the best results in all performance metrics: the IoU reaches 82.34 % (an improvement of 9.78 % compared with the baseline model), the FPS is 15.00 (with an acceptable increase in the inference time), and the growth in the number of parameters matches the extent of performance improvement. Dynamic upsampling, by adaptively adjusting sampling weights, accurately restores fine details of crack edges while balancing accuracy and efficiency.

2. *Scenario universality experiment*

In road crack segmentation tasks, scene factors such as pavement materials, environmental interferences (e.g., stains, shadows), and pavement conditions (e.g., wear, roughness) significantly affect the difficulty of segmentation. To verify the adaptability of DynaEdge-Net in complex scenes, five typical scenes were selected for scene generalisation experiments, namely asphalt pavement, concrete pavement, stained pavement, pavement with shadows, and worn rough pavement (including potholes). The segmentation results are shown in Fig. 5.



(a) (b) (c) (d) (e) (f)

Fig. 5. Visualization display of road crack detection results in different scenarios: (a) Input image; (b) Ground truth; (c) DeepLabV3+; (d) PIPNet; (e) ConvNeXt-FPN; (f) DynaEdge-Net.

In the asphalt pavement scene, the input image contains multiple fine and discontinuous cracks, and the ground truth clearly defines the complete shape of the cracks. DeepLabV3+ shows obvious missed detections for fine crack branches with a width of less than two pixels, with sawtooth distortion at the segmentation edges. PIPNet is disturbed by local textures, misclassifying some pavement particles as cracks and generating point noise. ConvNeXt-FPN maintains crack continuity but has edge offset at crack turns. In contrast, DynaEdge-Net, through the synergy of dynamic upsampling and the RDEB module, not only fully captures all fine cracks but also achieves high edge localisation accuracy, with significantly improved consistency with the ground truth boundaries.

In the concrete pavement scene, the input image has regular horizontal textures, and cracks are distributed obliquely, intersecting with background textures. DeepLabV3+ is strongly disturbed by textures, with approximately 30 % of pavement textures missegmented as cracks. PIPNet reduces false detections but results in discontinuous segmentation at the intersection of cracks and textures. ConvNeXt-FPN has limited ability to distinguish between cracks and textures, leading to blurred edges in segmentation results. DynaEdge-Net, relying on the feature screening mechanism of the GDG module, effectively suppresses interference from regular textures, accurately distinguishes cracks from the background, and achieves greater consistency with the ground truth.

In the stained pavement scene, the stained areas and cracks share similar grey features, increasing segmentation difficulty. DeepLabV3+ misclassifies some stain edges as crack extensions, causing expansion of the segmented area. PIPNet has missed detections in the transition area between stains and cracks. ConvNeXt-FPN shows a weak response to cracks in low-contrast areas. DynaEdge-Net, through the CGA module attention focussing on crack areas, successfully suppresses stain interference, accurately labels only crack areas, and its segmentation accuracy is significantly better than that of comparative models.

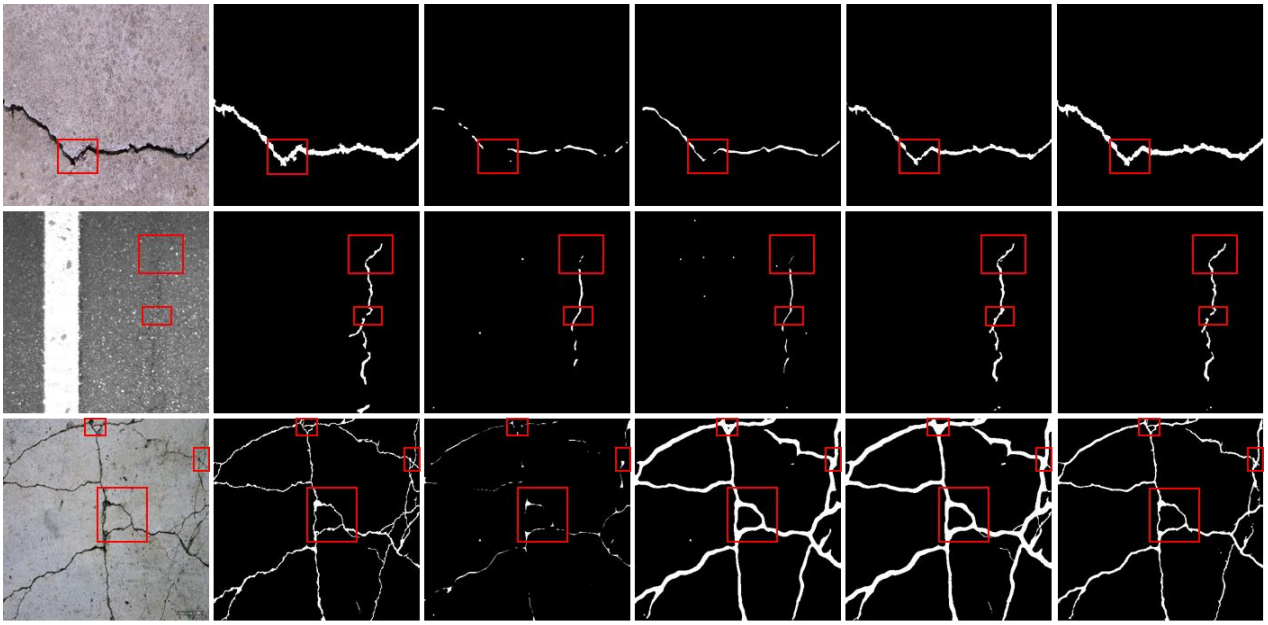
In the pavement scene with shadows, uneven illumination causes sharp local grey changes in cracks, and the grey value of cracks in shadow areas is lower than in normal areas. DeepLabV3+ has segmentation faults at shadow boundaries and more missed detections of cracks in shadow areas. PIPNet has insufficient sensitivity to low-grey cracks, resulting in discontinuous segmentation results. ConvNeXt-FPN detects cracks in shadow areas but with oversmoothed edges. DynaEdge-Net, through the multidirectional differential convolution of the RDEB module, enhances feature expression of low-contrast edges, effectively improves crack segmentation in shadow areas, and significantly reduces missed detections.

In the worn rough pavement scene, the pavement has numerous potholes and irregular textures, which are highly mixed with crack features. DeepLabV3+ misclassifies some potholes as cracks, resulting in a high false positive rate. PIPNet reduces pothole misclassification, but misses fine cracks adjacent to potholes. ConvNeXt-FPN insufficiently distinguishes structural differences between cracks and potholes, leading to low accuracy in the segmentation results. DynaEdge-Net, through the synergy of all core modules, accurately captures structural differences between cracks and potholes, effectively reduces the false positive rate, maintains complete crack segmentation, and shows strong anti-interference ability.

The comprehensive results of the five scenes show that DynaEdge-Net maintains stable segmentation performance under different pavement materials, environmental interferences, and structural noise conditions. This is due to the synergy of CGA, RDEB, GDG, and dynamic upsampling modules, which adaptively cope with challenges from complex scenes, verifying the model's scene generalisation.

3. *Experiment on the robustness of crack morphology*

The diversity of crack morphologies is another major challenge in road crack segmentation tasks. Different crack morphologies (e.g., single cracks, fractured cracks, and reticular cracks exhibit significant differences in feature expression, placing higher demands on the model's feature extraction and modelling capabilities. To verify the segmentation robustness of DynaEdge-Net for different crack morphologies, three typical crack morphologies were selected for visualisation experiments, namely single cracks, fractured cracks, and reticular cracks. The segmentation results are shown in Fig. 6.



(a) (b) (c) (d) (e) (f)

Fig. 6. Visual display of the detection results of road cracks with different crack shapes: (a) Input image; (b) Ground truth; (c) DeepLabV3+; (d) PIPNet; (e) ConvNeXt-FPN; (f) DynaEdge-Net.

In the single crack scenario, the crack in the input image shows a continuous linear distribution with uniform width but slightly blurred edges. The ground truth clearly presents the complete outline of the crack. For comparison, DeepLabV3+ can segment the approximate position of the crack, but suffers from severe discontinuity and fails to clearly segment the

correct edges of the crack. PIPNet, though overall better than DeepLabV3+, still has a certain degree of missed detections. ConvNeXt-FPN oversmooths the crack edges, losing details of the crack edges. In contrast, DynaEdge-Net, through the synergy of dynamic upsampling and the RDEB module, accurately captures the complete trend of the crack, with small edge localisation errors. It also fully retains the fine extensions at both ends of the crack, achieving the highest consistency with the ground truth.

In the fractured crack scenario, the crack in the input image is divided into multiple segments due to pavement particles or tiny depressions, with certain gaps between the segments. The ground truth clearly marks the relevance of each segment. Experimental results show that DeepLabV3+ has a weak ability to connect fractured regions, only roughly locating the crack, with obvious discontinuity and many missed detections in the segmentation results. PIPNet can depict some crack details but still has a large number of missed detections and overall noise. ConvNeXt-FPN can well identify crack edges and complete most crack segmentation, but for fractured cracks, it missegments noncrack parts as cracks, leading to segmentation errors. DynaEdge-Net, relying on the feature correlation mechanism of the CGA module and the dynamic screening capability of the GDG module, successfully identifies the overall structure of fractured cracks. It achieves complete crack segmentation by reasonably filling the gaps between segments, while avoiding false positives caused by overconnection. The structural consistency of its segmentation results with the ground truth is significantly improved.

In the reticular crack scenario, the cracks in the input image form a complex multidirectional interwoven structure, with numerous intersections and branches. The ground truth clearly distinguishes the intersection relationships of each crack. Observations show that DeepLabV3+ results in blurred segmentation in intersection regions, leading to a significant loss of structural information and containing a lot of noise. PIPNet, limited by the size of its receptive field, is insufficient in global modelling of reticular structures, with missed detections of some branch cracks in the segmentation results. ConvNeXt-FPN can identify most of the crack branches but has an unclear definition of the edge in the intersection regions, resulting in overlapping edges of different crack branches. DynaEdge-Net, through the selective activation of multidirectional features by the GDG module and the global attention adjustment by the CGA module, accurately segments the branches and intersection regions of each crack, clearly presenting the topological relationships of the reticular structure. Its recognition accuracy is significantly better than that of comparative models, with clear and independent edges of each branch crack and no overlapping.

Overall, DynaEdge-Net achieves the best performance in segmenting the three typical crack morphologies. The core lies in the synergy of its modules: the RDEB module strengthens the detailed features of the crack edges; the CGA module improves the correlation modelling of the overall crack structure; the GDG module realises dynamic screening of complex morphological features; and dynamic upsampling ensures the accuracy of edge localisation. This design enables the model to adapt to feature differences of different crack morphologies, showing strong morphological robustness and being suitable for various crack segmentation scenarios on actual roads.

To further quantitatively verify the segmentation performance of DynaEdge-Net, its performance on the test set is compared with that of three comparative segmentation models. The results are shown in Table II.

TABLE II. QUANTITATIVE COMPARISON OF SEGMENTATION PERFORMANCE FOR DIFFERENT MODELS.

Model	IoU/%	F1-score/%	Boundary F1/%	Recall/%
DeepLabV3+	76.23	85.17	72.45	82.69
PIPNet	78.56	86.89	74.32	84.15
ConvNeXt-FPN	80.12	88.03	76.59	85.78
DynaEdge-Net	82.34	90.83	83.27	90.12

As shown in Table II, DynaEdge-Net significantly outperforms other comparative models in all four evaluation metrics. Specifically, its IoU improves by 2.22 % compared to ConvNeXt-FPN (the second-best performer), indicating better global segmentation accuracy. The Boundary F1-score (BF1, an indicator quantifying the matching accuracy between crack edges and ground truth boundaries) increases by 6.68 %, reflecting its ability to precisely capture crack edges. The leading performance in F1 and Recall verifies the model's advantage in balancing precision and recall, especially its superior ability to detect crack pixels in complex scenarios.

This result is consistent with the visual analysis of the scene generalisation experiment, further confirming that DynaEdge-Net, through the synergy of CGA, RDEB, GDG, and dynamic upsampling modules, achieves higher precision than mainstream models.

IV. CONCLUSIONS

This study addresses the dual challenges of high computational complexity and loss of fine details in road crack segmentation tasks. It proposes DynaEdge-Net, a lightweight high-precision segmentation network based on GDG. The main conclusions are as follows:

1. The GDG mechanism enables on-demand scheduling of computational resources, and when working with RDEB, CGA, and DySample modules, it ensures segmentation accuracy while significantly reducing computational burden, with dynamic upsampling (DySample) further optimising edge structure recovery.
2. The ablation experiments show that DynaEdge-Net achieves an IoU of 82.34 %, which is a 9.78 percentage point improvement compared to the baseline model UNet (72.56 %). In the experiments on the integrated dataset, its IoU reaches 82.34 %, representing a 5.57 percentage point improvement over the comparative method ConvNeXt-FPN (80.12 %).
3. The method cuts computational complexity and parameters, allowing a feasible deployment on resource-constrained devices. However, performance lags in extreme noise (e.g., strong light reflection) and subpixel crack detection. Future work will optimise dynamic gating efficiency, boost extreme condition robustness, and extend to joint detection of multiple road damages (e.g., potholes, spalling).

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

[1] P.-H. Chen, J.-W. Hsieh, Y.-K. Hsieh, C.-W. Chang, and D.-Y. Huang, "Cross-scale overlapping patch-based attention network for road crack detection", *IEEE Transactions on Intelligent Transportation Systems*, vol. 26, no. 6, pp. 7587–7599, 2025. DOI: 10.1109/TITS.2025.3558279.

[2] X. Ren, S. Hu, Y. Hou, Y. Ke, Z. Chen, and Z. Wu, "A lightweight convolutional neural network for detecting road cracks", *Signal, Image and Video Processing*, vol. 18, no. 10, pp. 6729–6743, 2024. DOI: 10.1007/s11760-024-03347-2.

[3] X. Li, G. Tian, K. Li, H. Wang, and Q. Zhang, "Differential ECT probe design and investigation for detection of rolling contact fatigue cracks with different orientations", *IEEE Sensors Journal*, vol. 22, no. 12, pp. 11615–11625, 2022. DOI: 10.1109/JSEN.2022.3170598.

[4] J. Yuan, N. Wang, S. Cai, C. Jiang, and X. Li, "A lightweight crack detection model based on multibranch ghost module in complex scenes", *IEEE Sensors Journal*, vol. 23, no. 19, pp. 22754–22762, 2023. DOI: 10.1109/JSEN.2023.3300715.

[5] J. Canny, "A computational approach to edge detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6, pp. 679–698, 1986. DOI: 10.1109/TPAMI.1986.4767851.

[6] Y. Chien, "Pattern classification and scene analysis", *IEEE Transactions on Automatic Control*, vol. 19, no. 4, pp. 462–463, 1974. DOI: 10.1109/TAC.1974.1100577.

[7] P. Li, C. Wang, S. Li, and B. Feng, "Research on crack detection method of airport runway based on twice-threshold segmentation", in *Proc. of Fifth International Conference on Instrumentation and Measurement, Computer, Communication and Control (IMCCC)*, 2015, pp. 1716–1720. DOI: 10.1109/IMCCC.2015.364.

[8] S. Zhang, P. Rao, X. Chen, and Y. Li, "An FCN-based transfer-learning method for spatial infrared moving-target recognition", *Infrared Physics & Technology*, vol. 137, art. 105145, 2024. DOI: 10.1016/j.infrared.2024.105145.

[9] E. Shelhamer, J. Long, and T. Darrell, "Fully convolutional networks for semantic segmentation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 4, pp. 640–651, 2017. DOI: 10.1109/TPAMI.2016.2572683.

[10] W. Huangfu *et al.*, "Automated extraction of mining-induced ground fissures using deep learning and object-based image classification", *Earth Surface Processes and Landforms*, vol. 49, no. 7, pp. 2189–2204, 2024. DOI: 10.1002/esp.5824.

[11] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder-decoder architecture for image segmentation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017. DOI: 10.1109/TPAMI.2016.2644615.

[12] H. Zunair and A. Ben Hamza, "Sharp U-Net: Depthwise convolutional network for biomedical image segmentation", *Computers in Biology and Medicine*, vol. 136, art. 104699, 2021. DOI: 10.1016/j.compbiomed.2021.104699.

[13] Z. Liu, Y. Cao, Y. Wang, and W. Wang, "Computer vision-based concrete crack detection using U-net fully convolutional networks", *Automation in Construction*, vol. 104, pp. 129–139, 2019. DOI: 10.1016/j.autcon.2019.04.005.

[14] J. Niyogisubizo, K. Zhao, J. Meng, Y. Pan, R. Didi, and Y. Wei, "Attention-guided residual U-Net with SE connection and ASPP for watershed-based cell segmentation in microscopy images", *Journal of Computational Biology*, vol. 32, no. 2, pp. 225–237, 2025. DOI: 10.1089/cmb.2023.0446.

[15] R. Augustauskas and A. Lipnickas, "Improved pixel-level pavement-defect segmentation using a deep autoencoder", *Sensors*, vol. 20, no. 9, p. 2557, 2020. DOI: 10.3390/s20092557.

[16] Q. Chen, H. Li, and X. Zheng, "A deep neural network for operator learning enhanced by attention and gating mechanisms for long-time forecasting of tumor growth", *Engineering with Computers*, vol. 41, pp. 423–533, 2025. DOI: 10.1007/s00366-024-02003-0.

[17] X. Li, X. Xu, and H. Yang, "A road crack detection model integrating GLMANet and EFPN", *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 11, pp. 18211–18223, 2024. DOI: 10.1109/TITS.2024.3432995.

[18] C. Shi *et al.*, "Real-time ConvNext-Based U-Net with feature infusion for egg microcrack detection", *Agriculture*, vol. 14, no. 9, p. 1655, 2024. DOI: 10.3390/agriculture14091655.

[19] Y. Liu, J. Yao, X. Lu, R. Xie, and L. Li, "DeepCrack: A deep hierarchical feature learning architecture for crack segmentation", *Neurocomputing*, vol. 338, pp. 139–153, 2019. DOI: 10.1016/j.neucom.2019.01.036.

[20] Y.-L. He *et al.*, "Eccentricity fault diagnosis of permanent magnet synchronous generators based on 2D recursive fusion graph and CBAM-ConvNeXt-FPN", *Measurement Science and Technology*, vol. 36, no. 5, 2025. DOI: 10.1088/1361-6501/adcd8e.

[21] L. Shi *et al.*, "AHC-Net: A road crack segmentation network based on dual attention mechanism and multi-feature fusion", *Signal, Image and Video Processing*, vol. 18, pp. 5311–5322, 2024. DOI: 10.1007/s11760-024-03234-w.

[22] H. Liu *et al.*, "PIPNet: A deep convolutional neural network for multibaseline InSAR phase unwrapping based on pure integer programming", *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 18, pp. 15495–15507, 2025. DOI: 10.1109/JSTARS.2025.3579319.

[23] Y. Feng, J. Bi, Y. Huo, and J. Li, "PIPNet: Lightweight asphalt pavement crack image segmentation network", *Journal of Computer Applications*, vol. 44, no. 5, pp. 1520–1526, 2024. DOI: 10.11772/j.issn.1001-9081.2023050911.

[24] X. Liu, H. Peng, N. Zheng, Y. Yang, H. Hu, and Y. Yuan, "EfficientViT: Memory efficient vision transformer with cascaded group attention", in *Proc. of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023, pp. 14420–14430. DOI: 10.1109/CVPR52729.2023.01386.

[25] J. Wang, B. Shen, G. Li, J. Gao, and C. Chen, "A novel multi-scale feature enhancement U-shaped network for pixel-level road crack segmentation", *Electronics*, vol. 13, no. 22, p. 4503, 2024. DOI: 10.3390/electronics13224503.

[26] S. Wang *et al.*, "GH-UNet: Group-wise hybrid convolution-ViT for robust medical image segmentation", *npj Digital Medicine*, vol. 8, art. no. 426, 2025. DOI: 10.1038/s41746-025-01829-2.

[27] W. Liu, H. Lu, H. Fu, and Z. Cao, "Learning to upsample by learning to sample", in *Proc. of 2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, 2023, pp. 6004–6014. DOI: 10.1109/ICCV51070.2023.00554.

[28] M. A. Khan, H. Park, and J. Chae, "A lightweight convolutional neural network (CNN) architecture for traffic sign recognition in urban road networks", *Electronics*, vol. 12, no. 8, p. 1802, 2023. DOI: 10.3390/electronics12081802.

[29] Y. Shi, L. Cui, Z. Qi, F. Meng, and Z. Chen, "Automatic road crack detection using random structured forests", *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 12, pp. 3434–3445, 2016. DOI: 10.1109/TITS.2016.2552248.

[30] L. Cui, Z. Qi, Z. Chen, F. Meng, and Y. Shi, "Pavement distress detection using random decision forests", in *Proc. of the Second International Conference on Data Science*, 2015, pp. 95–102. DOI: 10.1007/978-3-319-24474-7_14.

[31] H. Fu, D. Meng, W. Li, and Y. Wang, "Bridge crack semantic segmentation based on improved Deeplabv3+", *Journal of Marine Science and Engineering*, vol. 9, no. 6, p. 671, 2021. DOI: 10.3390/jmse9060671.

[32] R. Amhaz, S. Chambon, J. Idier, and V. Baltazart, "Automatic crack detection on two-dimensional pavement images: An algorithm based on minimal path selection", *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 10, pp. 2718–2729, 2016. DOI: 10.1109/TITS.2015.2477675.

[33] L. Zhang, F. Yang, Y. D. Zhang, and Y. J. Zhu, "Road crack detection using deep convolutional neural network", in *Proc. of 2016 IEEE International Conference on Image Processing (ICIP)*, 2016, pp. 3708–3712. DOI: 10.1109/ICIP.2016.7533052.

[34] F. Yang, L. Zhang, S. Yu, D. Prokhorov, X. Mei, and H. Ling, "Feature pyramid and hierarchical boosting network for pavement crack detection", *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 4, pp. 1525–1535, 2020. DOI: 10.1109/TITS.2019.2910595.