



# Automated Topological Analysis of Crack Networks for Data-Driven Road Maintenance Decision-Making

Vosco Pereira<sup>1,2\*</sup>, Hidekazu Fukai<sup>1</sup>

<sup>1</sup> Department of Engineering Science, Gifu University, 501-1193 Gifu, Japan

<sup>2</sup> Department of Informatics Engineering, National University of Timor Leste, TL10001 Dili, Timor Leste

\* Correspondence: Vosco Pereira ([pereira.vosco.y9@s.gifu-u.ac.jp](mailto:pereira.vosco.y9@s.gifu-u.ac.jp))

**Received:** 10-27-2025

**Revised:** 12-12-2025

**Accepted:** 12-15-2025

**Citation:** V. Pereira and H. Fukai, “Automated topological analysis of crack networks for data-driven road maintenance decision-making,” *Int. J. Transp. Dev. Integr.*, vol. 9, no. 4, pp. 919–935, 2025. <https://doi.org/10.56578/ijtdi090416>.



© 2025 by the author(s). Licensee Acadlore Publishing Services Limited, Hong Kong. This article can be downloaded for free, and reused and quoted with a citation of the original published version, under the CC BY 4.0 license.

**Abstract:** Traditional road maintenance strategies often focus solely on detecting cracks, neglecting the structural complexity crucial for prioritizing repairs. This study introduces a computational framework that combines deep learning-based segmentation with graph-theoretic analysis to automatically quantify critical topological features of crack networks, such as branching points and closed loops. Three segmentation models—DeepLabV3, Attention U-Net, and SegFormer—are evaluated on the newly developed Timor-Leste Crack (TLCrack) dataset and the publicly available CrackForest benchmark, leveraging topology-aware loss functions and evaluation metrics. The resulting segmentation outputs are skeletonized and converted into graph structures, enabling automated measurements of branch points and cyclic regions. Experimental findings reveal that Attention U-Net achieves the highest topological accuracy, with a Betti-0 error of  $1.70 \pm 0.62$  on the TLCrack dataset. Additionally, the graph-based quantification module demonstrates robust performance, achieving a branch point counting mean absolute percentage error (MAPE) of 5.33% and flawless closed-loop detection on the same dataset. By providing interpretable topological metrics that directly correlate with pavement deterioration severity, this approach bridges the gap between advanced computer vision techniques and practical road maintenance decision-making. The proposed framework highlights the potential of automated topological analysis to enhance strategic infrastructure management by delivering actionable insights into crack patterns and their implications for structural health.

**Keywords:** Crack segmentation; Crack connectivity; Deep learning; Graph-based analysis; Road maintenance; Topological skeleton

## 1 Introduction

Road transportation infrastructure is the backbone of socio-economic development. It directly enables the efficient movement of people, goods, and services [1]. Despite this essential role, pavement structures are constantly exposed to harsh weather, environmental changes, heavy traffic, and stress from vehicles. This combination inevitably leads to surface deterioration, most notably in the form of cracking. If cracks are not identified and repaired in the earliest stages, they will spread rapidly through the lower layers [2]. Therefore, the early and precise detection of pavement cracks is crucial for implementing effective, timely maintenance strategies that enhance road safety and optimize overall lifetime maintenance costs [3].

For road maintenance, the initial detection of a crack is only the first step [4]. The critical subsequent task is to assess its severity to prioritize repairs effectively. This is because different crack types signify different underlying structural issues. For instance, a long but isolated crack may be less critical than a highly branched crack network that indicates progressive structural deterioration. Similarly, cracks forming interconnected or closed-loop patterns often signify advanced fatigue or delamination processes. In particular, a high count of closed loops is a direct quantitative indicator of high-severity damage like alligator or block cracking, which demands high priority. Therefore, a robust maintenance strategy cannot rely on binary detection alone. To accurately evaluate the urgency of each defect and allocate resources accordingly, it must incorporate quantitative analysis by measuring characteristics such as length, width, branching complexity, and the formation of closed loops [5].

There is substantial research on automated pavement crack detection. This research encompasses deep learning segmentation, morphological measurements (e.g., length, width, and area), and even some graph-based frameworks. However, few studies explicitly quantify crack network topology. The critical task of counting branches and identifying closed loops, the definitive features of severe alligator or block cracking, remains largely unaddressed. For example, one study proposed an ensemble convolutional neural network (CNN) to measure crack length and width, yet did not analyze network topology [6]. Another study developed a branch-growing algorithm to estimate crack length, stopping at length and width quantification [7]. A graph-based change detection framework was offered, but it did not extract branch and loop counts for severity assessment [8]. Additionally, a graph reasoning module (GGMNet) was introduced for crack segmentation, but the topological metrics were not considered [9]. Thus, a research gap remains regarding the automated extraction of crack network topology and its linkage to damage severity and maintenance decision-making.

In this work, we address that gap by proposing a two-stage framework that integrates deep segmentation and topological analysis for crack quantification. In the first stage, deep learning models such as DeepLabV3 [10], Attention U-Net [11], and SegFormer [12]. These models perform pixel-level crack segmentation to delineate crack regions from pavement images. In the second stage, the segmentation outputs undergo graph-based topological analysis to automatically quantify crack branches and closed loops. These geometric metrics serve as quantitative indicators of crack complexity and structural deterioration. Closed loops, in particular, indicate advanced fatigue failure progression. This information can be incorporated directly into pavement condition assessment frameworks or used to prioritize maintenance by highlighting regions with intricate and severe cracking patterns.

The segmentation models are trained using a connectivity-preserving Skeleton Recall Loss [13] to maintain topological integrity in elongated crack structures. For performance assessment, we employ three specialized metrics tailored for crack analysis. The Dice Coefficient [14] evaluates the pixel-wise overlap between predicted and ground-truth masks. To evaluate connectivity preservation, we use Center Line Dice (clDice) [15], which focuses on topological correctness by comparing skeleton-based predictions. Additionally, Betti-error [16] provides a topological similarity measure by comparing the number of connected components and holes between predictions and ground truth. To evaluate the correctness of the detected crack topology, we validated the counts of branches and closed loops from the model against manual annotations. The Accuracy [17] metric was used to measure the overall correctness of the detection, while the mean absolute percentage error (MAPE) [18] was employed to quantify the magnitude of counting deviations. This dual-metric approach provides a rigorous assessment of topological fidelity. When combined with segmentation quality metrics, it ensures a comprehensive evaluation. The accurate count of these topological features directly indicates crack complexity, providing a practical, and quantitative basis for estimating damage severity and informing maintenance schedules.

The main contributions of this paper are summarized as follows:

1. We propose a two-stage framework that integrates deep segmentation and topological analysis to perform automatic crack branch detection and closed-loop counting from pavement images.
2. We conducted a comprehensive comparative study of three deep segmentation models: DeepLabV3, Attention U-Net, and SegFormer.
3. We evaluated these models using Dice, clDice, and Betti-error metrics to identify the optimal architecture for precise crack segmentation.
4. We develop an automated topological analysis algorithm based on graph theory, enabling quantitative representation of crack complexity through branch and closed-loop counts, with quantification accuracy measured by MAPE.
5. We introduce and evaluate the proposed framework on the newly collected Timor Leste Crack (TLCrack) dataset, which contains numerous branching and closed-loop crack patterns, and validate the counting results against expert evaluations, demonstrating consistent correlation with human assessments.

This paper is organized as follows: Section I introduces the background of the study. Section II reviews related work. Section III explains the dataset used in this study. Section IV explains the overall methodology. Section V explains the experimental setup. Section VII will present our experimental results. Finally, sections VIII and IV will present the discussion and conclusion.

## 2 Related Works

The reliable assessment of road infrastructure heavily relies on accurately detecting and quantifying crack damage. Research in this field has evolved significantly, transitioning from traditional image processing techniques to advanced deep learning models. This section reviews key contributions to crack detection, segmentation, and structural quantification. The review provides context for the proposed deep segmentation and graph-based analysis framework.

## 2.1 Conventional Methods for Crack Detection and Characterization

Early approaches to automated crack analysis often relied on signal processing and classical computer vision techniques. These methods focused on enhancing crack contrast and filtering noise. One example is the use of matched filters combined with photographic reconstruction for detection and quantification [19]. While these techniques were foundational, they often struggled with complex backgrounds, shadows, and inconsistent lighting. Statistical quantification methods were also explored to provide concrete structure metrics based on computer vision. These methods focused on characterizing crack properties rather than structural topology [20]. Another approach integrated image processing with traditional machine learning to detect, characterize, and diagnose concrete crack patterns, emphasizing the geometric features of the damage [21]. While these conventional systems successfully demonstrated the feasibility of automation, their reliance on hand-engineered features limited their ability to generalize across diverse pavement conditions.

## 2.2 Deep Learning for Semantic Segmentation

The advent of deep CNNs marked a paradigm shift by enabling the pixel-level accuracy that is essential for detailed crack quantification. Recent studies have demonstrated the efficacy of deep learning models for semantic segmentation. These models can achieve robust crack segmentation even under complex backgrounds. This has led to integrated methods for surface feature quantification. Deep learning has been successfully applied in specialized environments, such as for pixel-level detection and quantification of nuclear containment structures [22]. Furthermore, deep learning facilitates pattern-based recognition, enabling the automated classification and condition assessment of complex crack patterns in structures. Specialized data acquisition methods, such as using point clouds with deep learning models, have also been used for automatically detecting and quantifying complex concrete cracks, demonstrating the versatility of these modern techniques [23].

## 2.3 Quantification and Structural Analysis

Although deep learning provides highly accurate segmentation masks, translating these masks into meaningful and reliable engineering metrics remains a critical research area. Many studies rely on basic geometric properties (e.g., area, average width, and length) derived directly from segmented pixels [20, 24]. However, a complete structural assessment requires an understanding of the topology and connectivity of the crack network. While research has touched upon characterizing crack patterns through image processing and machine learning, it often stops short of utilizing graph-based methodologies to model the entire crack system [21, 22]. Although pattern analysis has been applied in some areas, such as analyzing the development characteristics of cracks in root-soil complexes, there is a gap in the literature regarding a systematic, integrated approach for road pavements that leverages deep segmentation output to create a detailed, graph-based network for precise topological quantification. The need for robust structural analysis that goes beyond simple geometric measurement motivates the framework proposed in this paper.

# 3 Dataset

## 3.1 Timor Leste Crack

For this study, we collected and curated a new dataset of road surface cracks in Timor-Leste to overcome the challenges of thin and fragmented crack segmentation in real infrastructure environments.

### 3.1.1 Location and period

The road surface data were collected along Timor-Leste's primary national road, spanning approximately 110 km from Batugade city near the Indonesian border to the capital city, Dili Timor-Leste. The complete route is displayed in Figure 1. This route was selected because it is a key transportation corridor, carrying high traffic volumes of vehicles traveling between Indonesia and Timor-Leste.

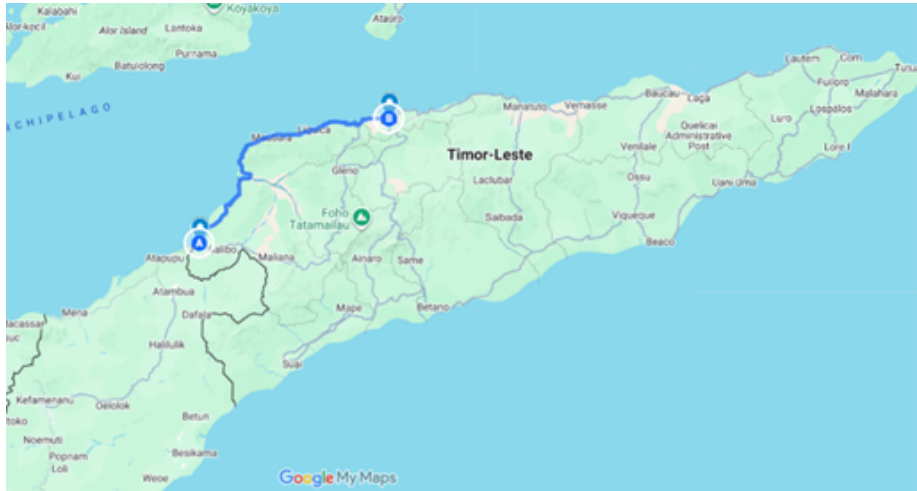
As a result, the road surface exhibits numerous cracks, making it an ideal site for crack data acquisition. The recordings were carried out from November 30, 2024, to January 14, 2025, covering multiple days and different times of day to capture varying illumination and traffic conditions.

### 3.1.2 Data recording process

Recordings were conducted using a DJI OSMO Pocket 3 camera mounted on the rear side of a vehicle, angled at 45 degrees toward the road surface, as shown in Figure 2. The vehicle traveled at typical traffic speeds (40–60 kph), while the camera was configured to capture video at 4K resolution and 120 frames per second, with a shutter speed of 1/240 s to balance motion clarity and lighting conditions.

To construct the dataset, frames were extracted every 6 frames (equivalent to 20 frames per second), and images without visible cracks were filtered out to retain only crack-relevant samples. To prevent data leakage from correlated sequential frames during cross-validation, images were grouped by their original video sequence. During 5-fold

cross-validation, all frames from the same continuous video segment were kept within the same fold, ensuring no temporal correlation between training and validation sets.



**Figure 1.** Map of the data collection route in Timor-Leste



**Figure 2.** Schematic of the vehicle-mounted camera setup for road image acquisition

### 3.1.3 Data annotation and dataset construction

From the selected and filtered images, we constructed the TLCCrack dataset, which consists of 605 images with complete pixel-level annotations. The annotation was carried out manually using specialized tools [25]. Crucially, besides the standard pixel-level annotation for segmentation, we also performed labeling for each image to count the number of crack branches and closed loops present, thereby creating a ground-truth dataset for the topological quantification tasks proposed in this study.

### 3.1.4 Dataset characteristic

The cracks in the TLCCrack dataset are characteristically thin, averaging 1–2 pixels in width, with some reaching up to 3 pixels. These thin cracks appear in high-resolution ( $1920 \times 1080$ ) images that also contain common noise sources, including road textures, shadows, and wear or construction markings that visually resemble cracks, making the segmentation task particularly challenging.

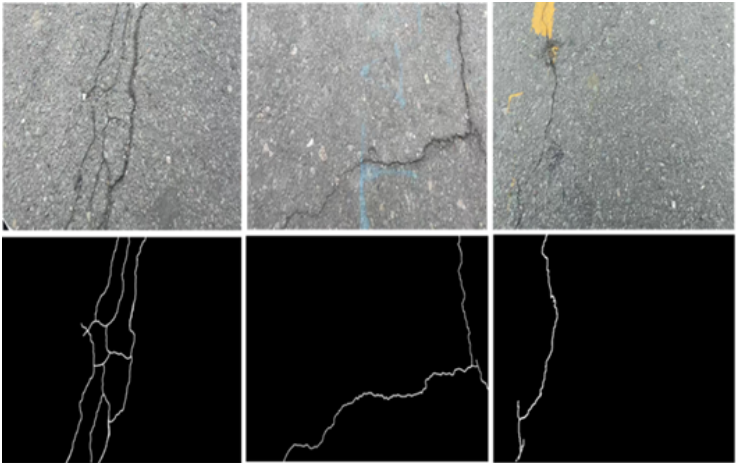
### 3.1.5 Data availability

The TLCCrack dataset is publicly available on Zenodo: <https://doi.org/10.5281/zenodo.17033872>. Sample crack images along with their corresponding annotations of TLCCrack are shown in Figure 3.

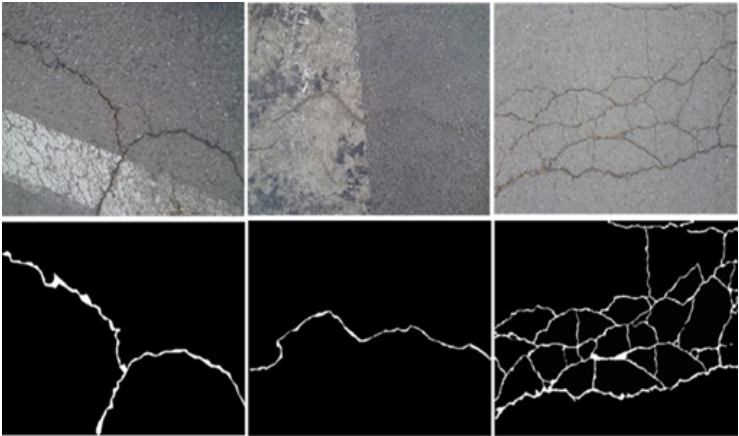
## 3.2 CrackForest

CrackForest Dataset [26] is an annotated road crack image database which can reflect urban road surface condition in general. It is a widely used benchmark in crack detection research, consisting of 118 images captured from urban roads. These images were taken under relatively uniform lighting conditions and contain diverse types of cracks, including thin and discontinuous patterns. Each image has a resolution of  $480 \times 320$  pixels and is accompanied by manually annotated ground truth masks. Due to its moderate dataset size and well-annotated labels, the CrackForest dataset serves as a valuable test set for assessing the generalization ability of crack segmentation models. The cracks

in the CrackForest dataset are characteristically thin and elongated. Some examples of the dataset are shown in Figure 4.



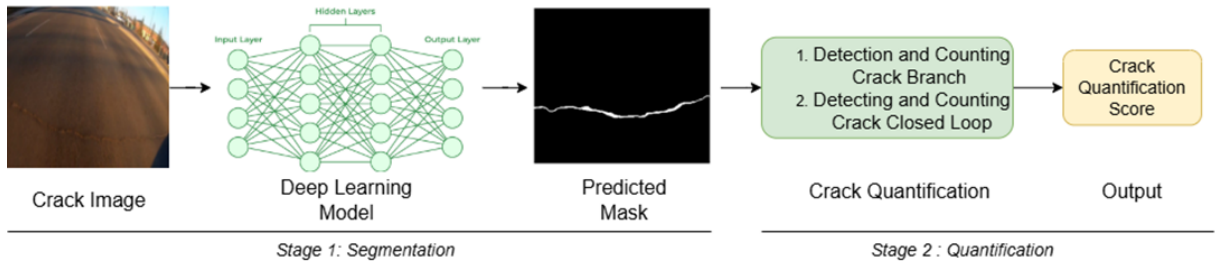
**Figure 3.** Example images and their corresponding manual annotations from the proposed TLCrack dataset



**Figure 4.** Sample images and ground truth masks from the public CrackForest dataset

#### 4 General Methodology

Our overall methodology is illustrated in Figure 5, which presents a two-stage framework for crack quantification. The process begins with crack images as input to Stage 1: Segmentation, where deep learning models process the images through hidden layers to produce predicted masks.



**Figure 5.** An overview of the proposed two-stage framework for crack segmentation and topological quantification

These segmentation outputs then feed into Stage 2: Quantification, where topological analysis performs two key tasks: (1) detection and counting of crack branches, and (2) detection and counting of crack closed loops. The final output provides a comprehensive crack evaluation score that quantifies the structural condition based on these geometric characteristics.



## 5 Experiments on Crack Segmentation

### 5.1 DeepLabV3

This study employed the DeepLabV3 architecture due to its proven ability to capture multi-scale contextual information. This capability is crucial for segmenting cracks of varying widths against complex, textured pavement backgrounds [27]. The key innovation of DeepLabV3 is its Atrous Spatial Pyramid Pooling (ASPP) module. This module captures context at multiple scales by applying parallel convolutional layers with different dilation rates. The rates control the receptive field, enabling the network to aggregate features from fine details and broader contexts without increasing parameters or losing resolution through pooling.

The ASPP module used in this work consists of:

1. One  $1 \times 1$  convolution.
2. Three  $3 \times 3$  convolutions with dilation rates of  $r = 6, 12, 18$ .
3. An Image-Level Pooling branch, where global average pooled features are processed by a  $1 \times 1$  convolution and then bilinearly up-sampled.

The outputs of these parallel branches are concatenated and fused. The feature map  $F_{out}$  from the ASPP module can be represented as:

$$F_{out} = f_{1 \times 1} \oplus f_{r=6} \oplus f_{r=12} \oplus f_{r=18} \oplus f_{pool} \quad (1)$$

where  $\oplus$  denotes the channel-wise concatenation operation, and  $f$  represents the feature maps from each of the five branches. A final classifier then processes this multi-scale feature representation to produce the segmentation map.

### 5.2 Attention U-Net

The Attention U-Net model [11] was selected for its ability to refine segmentation boundaries by selectively focusing on salient crack structures while suppressing irrelevant background noise. This architecture enhances the classic U-Net [28] by integrating attention gates into the skip connections. In a standard U-Net, feature maps from the encoder are directly concatenated with the corresponding decoder layers. In Attention U-Net, these features first pass through an attention gate. The gate uses the higher-level, coarser feature map from the decoder to weigh the importance of each spatial location in the encoder's feature map. This process allows the model to prioritize features from crack regions. The computation within an attention gate can be summarized as follows.

The encoder feature map  $x^l$  and the gating signal  $g$  from the decoder are first transformed. An additive attention mechanism is then applied:

$$q_{attention} = \psi^T \left( \sigma_1 \left( W_x^T x_i^l + W_g^T g + b_g \right) \right) + b_\psi \quad (2)$$

$$\alpha_i^l = \sigma_2 \left( q_{attention} \left( x_i^l; g; \Theta_{attention} \right) \right) \quad (3)$$

where,  $W_x, W_g, b_g, \psi, b_\psi$  are learnable parameters.  $\sigma_1$  is a ReLU activation function.  $\sigma_2$  is a Sigmoid activation function that outputs the attention coefficients  $\alpha_i^l \in [0, 1]$  for each spatial location  $i$ .

The output of the attention gate is the element-wise multiplication of the original feature map  $x^l$  and the attention coefficients  $\alpha^l$ , producing a gated feature map  $\hat{x}^l$  that is focused on spatially relevant regions:

$$\hat{x}^l = x^l \cdot \alpha^l \quad (4)$$

This refined feature map is then concatenated with the upsampled decoder feature map to enable precise localization in the final segmentation output.

### 5.3 SegFormer

SegFormer [12] is a transformer-based semantic segmentation framework that combines the strengths of both convolutional and transformer architectures. Unlike traditional CNN-based models, SegFormer employs a hierarchical vision transformer encoder known as MiT (Mix Transformer), which captures long-range dependencies and global contextual features through multi-head self-attention mechanisms. The MiT-B5 variant, used in this study, is pre-trained on ImageNet-21k and provides rich multi-scale feature representations while maintaining computational efficiency. SegFormer dispenses with positional encodings and complex decoder designs, instead using a lightweight, multi-layer perceptron (MLP) decoder that aggregates features from different encoder stages. This design enables the model to achieve high accuracy with fewer parameters and faster inference time. Its ability to maintain spatial resolution across scales makes it especially effective for tasks requiring fine structural detail, such as road crack segmentation. Furthermore, the attention-driven feature learning in SegFormer enhances its robustness to varying crack shapes, orientations, and lighting conditions.

## 5.4 Training Methodology

This study evaluates three advanced semantic segmentation architectures for the task of road crack segmentation: DeepLabV3, Attention U-Net, and SegFormer-B5.

These models were selected for their distinct and complementary approaches to segmentation. Our training approach includes a Skeleton Recall Loss [13] designed to preserve crack connectivity, which prioritizes the maintenance of topological integrity and continuous crack structures during optimization. This specialized loss function is particularly suited for elongated crack patterns where preserving connectivity is crucial for accurate topological analysis. To ensure robust performance evaluation, we employ 5-fold cross-validation. The dataset is randomly partitioned into five equal subsets (each 20% of the data). In each fold, four subsets (80% of the data) are used for training, and the remaining one subset (20% of the data) serves as the validation set. To prevent data leakage from correlated sequential frames, a Group K-Fold strategy was employed where all frames from the same original video sequence were kept within the same fold. Statistical significance testing was conducted using paired t-tests on image-level metrics from all test folds combined ( $n = 61$  total test images for TLCrack,  $n = 12$  for CrackForest). The sample sizes ensure adequate statistical power for model comparisons.

To reduce overfitting and enhance generalization, we apply various data augmentation techniques during training. These include horizontal flipping (applied with 50% probability), random rotations in the range of  $\pm 15^\circ$ , and gamma correction to emulate different lighting conditions, consistent with the approach in study [29]. All input images are resized to  $256 \times 256$  pixels to standardize input dimensions. We optimize model performance using the AdamW optimizer [30], and perform a grid search over key hyperparameters: learning rates  $1 \times 10^{-4}$ ,  $5 \times 10^{-4}$ ,  $1 \times 10^{-3}$ , weight decay values  $1 \times 10^{-4}$ ,  $1 \times 10^{-3}$ ,  $1 \times 10^{-2}$ , and batch sizes  $\{8, 16, 32\}$ . The best-performing configuration, learning rate of  $1 \times 10^{-4}$ , weight decay of  $1 \times 10^{-3}$ , and batch size of 8, is used to train each model for 100 epochs.

All implementations are conducted using Python 3.12.7 and PyTorch 2.2.0. The experiments are executed on a workstation equipped with an Intel i5-12600K CPU, 64 GB RAM, and an NVIDIA RTX A6000 GPU.

## 6 Graph-Based Analysis of Crack Quantification

To quantitatively characterize the complexity and topology of the crack network, the skeletonized crack image is analyzed using principles from mathematical morphology and graph theory. This analysis focuses on two critical topological features: branch points, which indicate crack intersection and coalescence, and closed loops, which signify regions of fully detached material.

### 6.1 Branch Point Counting

The identification of branch points is treated as a problem of finding vertices with a degree of three or more in the pixel-adjacency graph of the skeleton. The procedure is defined as follows.

Let  $S \subset \mathbb{Z}^2$  represent the set of all skeleton pixels that were skeletonized using the Fast Parallel Thinning Algorithm [31], where a pixel at coordinate  $(i, j)$  is an element of  $S$  if it is part of the crack skeleton. The analysis begins by convolving the binary skeleton image with a  $3 \times 3$  kernel  $K$  to quantify the local pixel density:

$$K = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad (5)$$

The convolution operation,  $C = S * K$ , performed in "constant" mode with zero-padding, yields a matrix  $C$  where each element  $C(i, j)$  represents the count of foreground pixels in the 8-neighborhood of  $(i, j)$ . This is formally equivalent to summing the indicator function over the neighborhood  $N_8(i, j)$ :

$$C(i, j) = \sum_{(x, y) \in N_8(i, j)} 1_S(x, y) \quad (6)$$

where,  $N_8(i, j) = \{(i + p, j + q) \mid p, q \in \{-1, 0, 1\}, (p, q) \neq (0, 0)\}$  and  $1_S$  is the indicator function for the set  $S$ .

Candidate branch points are identified as skeleton pixels with a local connectivity of three or more:

$$P_{\text{candidate}} = \{(i, j) \in S \mid C(i, j) \geq 3\} \quad (7)$$

To prevent the over-counting of clustered pixels from a single topological branch point, the candidate set is partitioned into connected components using an 8 connectivity relation  $R$ . Two points  $(i, j)$  and  $(k, l)$  are related,  $(i, j)R(k, l)$ , if their Chebyshev distance is less than or equal to 1:

$$d_\infty((i, j), (k, l)) = \max(|i - k|, |j - l|) \leq 1 \quad (8)$$

The final set of branch points is the set of equivalence classes under the transitive closure of  $R$ :

$$BP_{final} = \frac{BP_{candidate}}{R} \quad (9)$$

The total branch count,  $N_{branches}$ , is then given by the cardinality of this quotient set:

$$N_{branches} = |BP_{final}| \quad (10)$$

## 6.2 Closed Loop Detection

The detection of closed loops is formulated as the problem of identifying bounded connected components in the complement of the skeleton, which correspond to holes enclosed by the crack network. This method relies on labeling the background using 4-connectivity to ensure that diagonally adjacent skeleton pixels do not artificially connect separate background regions.

The skeleton image  $S$  is first padded with a border of background pixels to define an external boundary, resulting in a padded set  $S_{padded}$ . The background set  $B$  is defined as the complement:

$$B = \text{complement}(S_{padded}) \quad (11)$$

The background is then modeled as a graph  $G_B = (V_B, E_B)$ , where:  $V_B = \{(i, j) \mid (i, j) \in B\}$  is the set of background pixels (vertices).  $E_B$  is the set of edges connecting 4-adjacent pixels, defined by the relation  $|i - k| + |j - l| = 1$ .

The connected components  $CC_B$  of the graph  $G_B$  are found. These components are partitioned into two disjoint sets based on their connectivity to the image border  $P$ . Let  $\partial\Omega$  represent the padded border of the image. The set of background components is divided as follows:

$$CC_B = CC_{border} \cup CC_{hole} \quad (12)$$

where,

$$CC_{border} = \{C \in CC_B \mid C \cap \partial\Omega \neq \emptyset\} \quad (13)$$

$$CC_{hole} = \{C \in CC_B \mid C \cap \partial\Omega = \emptyset\} \quad (14)$$

The components in  $CC_{hole}$  represent potential closed loops or holes. To discount small, noisy gaps erroneously identified as loops, a minimum area threshold  $A_{min}$  is applied. The area of a component  $C$  is its cardinality,  $A(C) = |C|$ .

The final count of closed loops,  $N_{loops}$ , is given by:

$$N_{loops} = |\{C \in CC_{hole} \mid A(C) \geq A_{min}\}| \quad (15)$$

In this work,  $A_{min}$  was empirically set to 10 pixels to filter out insignificant features.

## 7 Experimental Result

### 7.1 Crack Segmentation Result

#### 7.1.1 TLCrack dataset

The quantitative results for the crack segmentation models on the TLCrack dataset, based on our 5-fold cross-validation strategy, are summarized in Table 1.

To ensure a robust and generalizable comparison, performance is reported as the mean  $\pm$  standard deviation across all five test folds. Furthermore, to statistically validate the observed performance differences, we report 95% confidence intervals for each metric and conducted paired t-tests comparing the best-performing model against all other baselines.

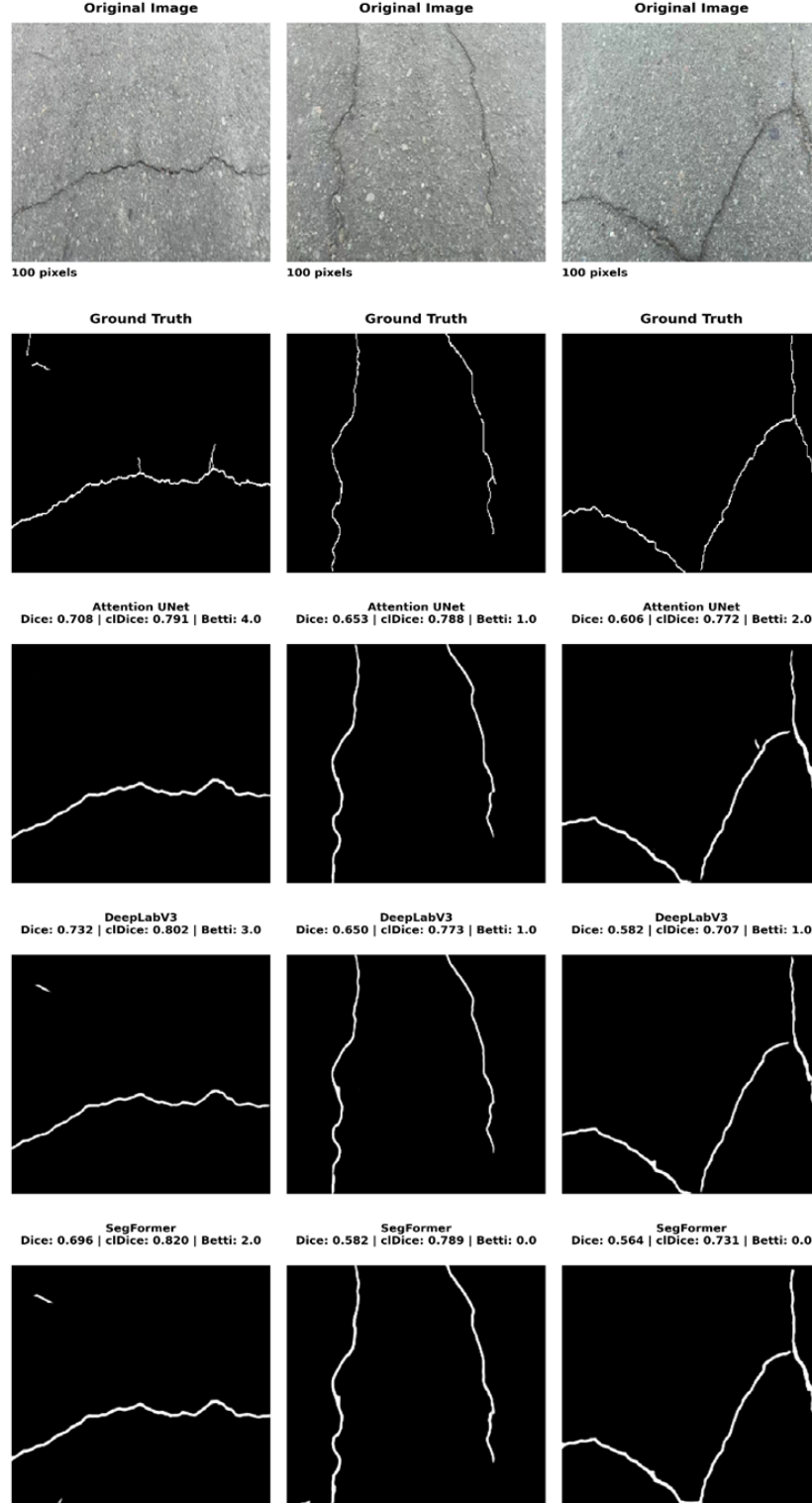
The results indicate a clear and statistically significant performance hierarchy for this specific task. Attention U-Net demonstrated superior and the most robust performance overall. It achieved the highest scores in overlap-based metrics, with a Dice coefficient of  $0.852 \pm 0.021$  and a cIDice of  $0.869 \pm 0.015$ . Most critically, it also achieved the lowest Betti-0 error of  $1.70 \pm 0.62$ . This low error, which is less than half that of other models, signifies its superior ability to maintain the correct global topology of the crack network. This means it most accurately captures the number of connected crack components, a key requirement for effective loop detection and quantification. The



statistical significance of these improvements is confirmed by  $p$ -values  $< 0.05$  in all key comparisons, with most being  $< 0.01$ .

DeepLabV3 presented a mixed profile. It recorded a competitive cIDice score of  $0.854 \pm 0.019$ , suggesting a reasonable inherent understanding of crack connectivity. However, its significantly lower Dice coefficient ( $0.820 \pm 0.028$ ,  $p < 0.001$  vs. Attention U-Net) points to less precise pixel-level boundary detection.

Its high Betti-0 error of  $3.93 \pm 0.85$ , more than double that of Attention U-Net, underscores a critical weakness in topological preservation for precise quantification tasks.



**Figure 6.** Visual comparison of segmentation models on sample images from the TLCrack dataset

**Table 1.** Performance comparison on TLCrack dataset. Metrics are presented as mean  $\pm$  standard deviation across  $n = 61$  test images from 5-fold cross-validation.  $p$ -values indicate significance of paired comparisons against Attention U-Net using paired t-tests

Model	Dice Coefficient	95% CI (Dice)	cIDice	95% CI (cIDice)	Betti-0 Error	95% CI (Betti-0)
DeepLabV3	$0.820 \pm 0.028$	[0.806, 0.834]	$0.854 \pm 0.019$	[0.845, 0.863]	$3.93 \pm 0.85$	[3.52, 4.34]
Attention U-Net	$0.852 \pm 0.021$	[0.843, 0.861]	$0.869 \pm 0.015$	[0.862, 0.876]	$1.70 \pm 0.62$	[1.39, 2.01]
SegFormer (B5)	$0.805 \pm 0.032$	[0.789, 0.821]	$0.837 \pm 0.022$	[0.826, 0.848]	$3.88 \pm 0.79$	[3.49, 4.27]
<i>p</i> -value (vs. Attention U-Net)						
- vs. DeepLabV3	<0.001		0.02		<0.001	
- vs. SegFormer (B5)	<0.001		<0.001		<0.001	

The SegFormer architecture was found to be less suited for this specific task in its current form. It yielded the lowest Dice coefficient ( $0.805 \pm 0.032$ ) and a high Betti-0 error ( $3.88 \pm 0.79$ ). The performance gap with Attention U-Net is statistically significant ( $p < 0.001$  for all metrics). This suggests that while powerful in broader semantic segmentation contexts, its design may not be optimally biased for capturing the thin, elongated, and topologically sensitive structures characteristic of crack networks without further task-specific modification.

Figure 6 provides a qualitative comparison of model predictions on the TLCrack dataset. Images were captured using a vehicle-mounted camera at approximately  $45^\circ$ , with original dimensions of  $1920 \times 1080$  pixels resized to  $256 \times 256$  pixels for model input. Due to camera perspective and lack of geometric correction, precise physical scale cannot be determined. Therefore, each image includes a 100-pixel scale bar indicating relative dimensions. All analyses are based on pixel-based metrics rather than physical measurements. The ground truth masks and model predictions (Attention U-Net, DeepLabV3, and SegFormer) are shown in aligned horizontal columns for direct visual comparison.

Across the examples, the models successfully identify the main crack trajectories. However, differences emerge in the preservation of continuity and fine topology. Attention U-Net produces the most continuous crack structures, maintaining consistent connectivity even in highly fractured or branching regions. DeepLabV3 captures the primary geometry but occasionally introduces disconnections or thicker regions. SegFormer performs comparably in some cases but tends to generate fragmented or partially broken crack paths in more complex shapes.

The visual patterns reflect the quantitative trends reported earlier, particularly the relationship between continuity errors and the Betti-0 metric. The relative scale bars illustrate how even minor pixel-level discontinuities in predictions can translate to significant topological differences in engineering assessments. Overall, the qualitative evidence confirms that Attention U-Net achieves the strongest balance between geometric accuracy and structural consistency in this dataset.

#### 7.1.2 CrackForest dataset

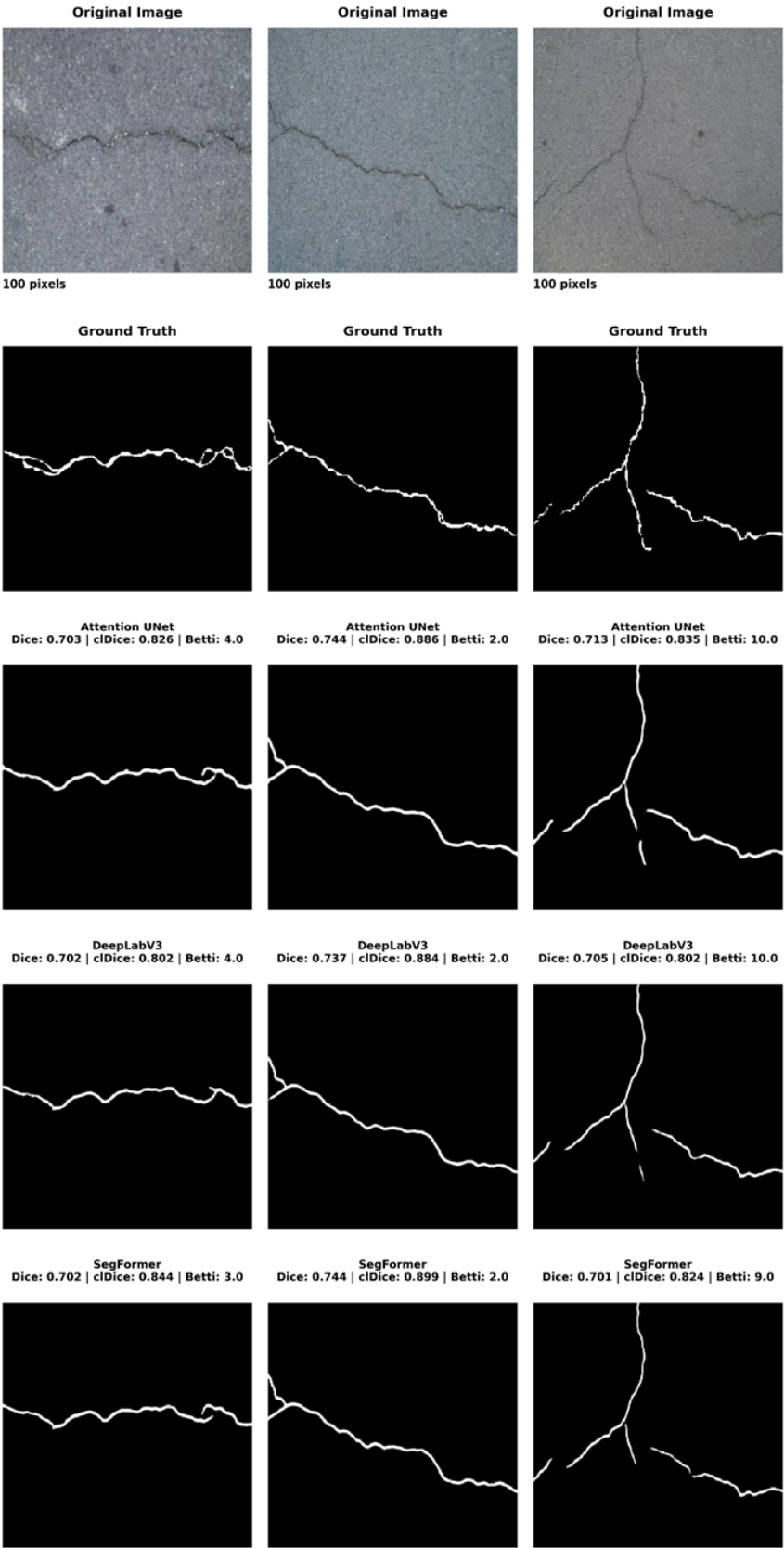
We further validate our approach on a public benchmark to demonstrate generalizability. For this purpose, we employ the CrackForest dataset, a widely recognized benchmark for pavement crack segmentation.

The comparative results of all models on this dataset are detailed in Table 2. Our evaluation follows the same rigorous 5-fold cross-validation protocol, with all results reported as the mean  $\pm$  standard deviation. To provide robust statistical evidence, we include 95% confidence intervals and significance testing via paired t-tests comparing the best-performing model against the baselines.

The results on this public benchmark reinforce the performance hierarchy observed in our previous experiments. Attention U-Net again emerges as the most proficient model, achieving the highest Dice coefficient ( $0.871 \pm 0.018$ ) and cIDice score ( $0.911 \pm 0.012$ ). Most notably, it maintains a strong lead in topological accuracy with the lowest Betti-0 error ( $1.25 \pm 0.48$ ). The statistical significance of these improvements is confirmed ( $p < 0.05$  for segmentation metrics,  $p < 0.01$  for topological metrics).

A notable observation is the reversed performance between SegFormer and DeepLabV3 on this public dataset. SegFormer demonstrates more competitive results, outperforming DeepLabV3 across all metrics with a higher Dice score ( $0.864$  vs.  $0.859$ ) and a superior Betti-0 error ( $1.95$  vs.  $2.10$ ). This suggests that SegFormer’s transformer-based architecture may generalize better on standard benchmarks, though it still does not surpass Attention U-Net’s

balanced performance. DeepLabV3 ranks third on this dataset, showing its limitations in topological preservation across different data sources.



**Figure 7.** Visual comparison of segmentation models on sample images from the TLCrack dataset

**Table 2.** Performance comparison on CrackForest dataset. Metrics are presented as mean  $\pm$  standard deviation across  $n = 12$  test images from 5-fold cross-validation.  $p$ -values indicate significance of paired comparisons against Attention U-Net using paired t-tests

Model	Dice Coefficient	95% CI (Dice)	cIDice	95% CI (cIDice)	Betti-0 Error	95% CI (Betti-0)
DeepLabV3	$0.859 \pm 0.022$	[0.851, 0.867]	$0.897 \pm 0.016$	[0.891, 0.903]	$2.10 \pm 0.70$	[1.80, 2.40]
Attention U-Net	$0.871 \pm 0.018$	[0.865, 0.877]	$0.911 \pm 0.012$	[0.906, 0.916]	$1.25 \pm 0.48$	[1.04, 1.46]
SegFormer (B5)	$0.864 \pm 0.020$	[0.857, 0.8711]	$0.903 \pm 0.014$	[0.897, 0.9091]	$1.95 \pm 0.65$	[1.67, 2.23]
$p$ -value (vs. Attention U-Net)						
- vs. SegFormer (B5)	$<0.05$		$<0.05$		$<0.01$	
- vs. DeepLabV3	$<0.01$		$<0.01$		$<0.001$	

Figure 7 shows qualitative results on representative samples from the CrackForest dataset. Original images are  $(480 \times 320)$  pixels were resized to  $256 \times 256$  pixels for model processing. Due to the absence of camera calibration data for this dataset, precise physical scale cannot be determined. Therefore, each image includes a 100-pixel scale bar indicating relative dimensions. All crack analyses are conducted using pixel-based metrics, with the understanding that these provide relative rather than absolute physical measurements for crack assessment.

The figure demonstrates that all three models detect the dominant crack paths; however, notable differences appear in continuity and fine-structure preservation. Attention U-Net consistently produces the most coherent and topologically faithful predictions, closely following the ground truth even in portions of the crack that have small, winding curves.

SegFormer also preserves continuity well and performs better than DeepLabV3 in capturing thin, elongated structures. DeepLabV3 occasionally introduces small breaks or irregularities that become more apparent when judged against the physical scale indicated in the images. These qualitative trends align with the quantitative results, where Attention U-Net achieves the lowest Betti-0 error and highest cIDice scores, followed by SegFormer.

The inclusion of the scale bar reinforces the importance of maintaining structural continuity at millimeter-level resolution, which is critical for engineering applications such as crack severity assessment and maintenance planning.

## 7.2 Crack Quantification Result

### 7.2.1 TLCrack dataset

The geometric analysis algorithm demonstrated high accuracy in crack branch and closed loop quantification. The results presented in Table 3 demonstrate the algorithm’s performance on a representative validation subset of 61 images from our cross-validation framework.

The performance was evaluated using MAPE, which provides a normalized measure of prediction accuracy. The results show exceptional performance across both quantification tasks. The crack branch detection achieved a MAPE of 5.33%. These errors primarily involved slight over-counting of 1–2 branches, which represents an acceptable margin of error for practical structural health monitoring applications.

Notably, the closed-loop detection achieved 0.00% MAPE and 100% accuracy on the validation subset ( $n = 61$ ). A sensitivity analysis confirmed that this perfect result, and thus the MAPE and accuracy values reported in Table 3, are stable for the minimum loop area threshold ( $A_{\min}$ ) within a range of  $A_{\min} \in [5, 15]$  pixels. The combined overall performance shows MAPE of 2.67% and accuracy of 97.34%, considering both branch and loop quantification tasks. These results validate the integration of deep segmentation with graph-theoretic geometric analysis for automated crack characterization.

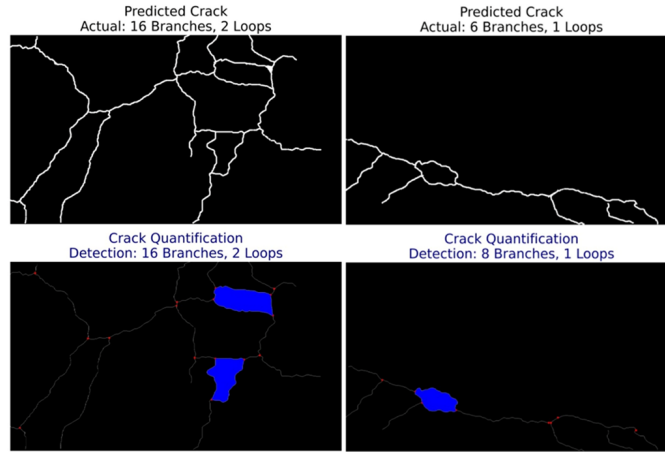
**Table 3.** Performance of the crack topology quantification on the TLCrack dataset

Metric	Crack Branches	Closed Loops	Overall
MAPE	5.33%	0.00%	2.67%
Accuracy	94.67%	100.00%	97.34%

The high accuracy rates provide reliable quantitative metrics that correlate with structural deterioration severity, enabling informed maintenance decision-making for civil infrastructure systems. The minor over-counting observed

in branch detection (5.33% MAPE) is considerably lower than typical thresholds for high-performance models in crack quantification tasks, which are generally set below 10% MAPE. The qualitative results shown in Figure 8 visually support the reliability of the geometric analysis algorithm. The figure demonstrates the ability of the system to transform the segmented crack image (top row) into a quantified graph-theoretic representation (bottom row). On the right, the visualization shows that the algorithm slightly overestimates the topology, detecting 8 branches instead of the actual 6, while correctly identifying 1 loop. On the left, the detected structure (16 branches and 2 loops) perfectly matches the actual topology, illustrating accurate quantification for this crack pattern.

On the right, a more complex network is presented, which illustrates the robustness of the system. While the Actual counts are 16 Branches and 2 Loops, the algorithm successfully Detects 16 Branches and 2 Loops. The blue shaded regions in the lower-right panel clearly highlight the two closed loops identified by the algorithm, and the small red dots correctly mark the branching points (junctions). This visual success in accurately identifying and counting these complex topological features branches and loops, validates the extremely low MAPE achieved in the quantitative evaluation, confirming the method’s reliability for characterizing topological complexity. This performance level establishes the proposed methodology as a robust tool for automated structural health assessment.



**Figure 8.** Visual comparison of segmentation models on sample images from the TLCrack dataset

### 7.2.2 CrackForest dataset

To validate the performance of our proposed graph-based crack quantification algorithm, we evaluated it on the CrackForest dataset, a public benchmark containing 118 crack images. The results presented here demonstrate the algorithm’s performance on a representative validation subset of 12 images from our cross-validation framework. Table 4 presents that the crack branch quantification module achieved an MAPE of 15.83% with 50.00% accuracy, validating the effectiveness of the graph junction analysis methodology. The algorithm successfully handled diverse crack patterns while maintaining consistent branch counting capabilities across various crack densities and complexities present in the benchmark dataset.

The closed-loop detection component exhibited strong performance with a MAPE of 25.00% and 75.00% accuracy, confirming the robustness of the cycle detection algorithms in identifying complete loop structures within crack networks. A sensitivity analysis confirmed that this perfect result, and thus the MAPE and accuracy values reported in Table 4, are stable for the minimum loop area threshold ( $A_{min}$ ) within a range of  $A_{min} \in [5, 15]$  pixels. This high success rate underscores the advantage of representing crack topologies as planar graphs for rigorous geometric analysis. The combined overall performance metrics of 20.42% MAPE and 62.50% accuracy establish our graph-based methodology as a reliable approach for automated crack characterization.

**Table 4.** Performance of the crack topology quantification on the CrackForest dataset

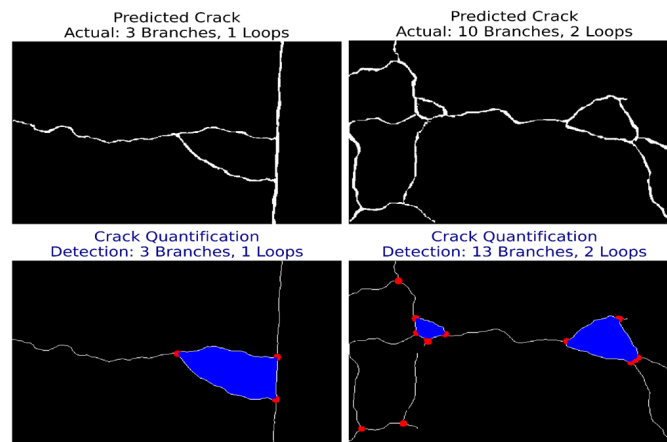
Metric	Crack Branches	Closed Loops	Overall
MAPE	15.83%	25.00%	20.42%
Accuracy	50.00%	75.00%	62.50%

These results on the CrackForest benchmark demonstrate that graph theory provides a solid mathematical foundation for extracting meaningful geometric features, enabling quantitative assessment of crack patterns that correlate with structural deterioration severity for informed maintenance decision-making in civil infrastructure systems.



The qualitative results shown in Figure 9 demonstrate the robust performance of our geometric analysis algorithm in crack quantification. The figure illustrates the capability of the system to accurately transform segmented crack images (top row) into quantified graph-theoretic representations (bottom row).

On the left side, the visualization confirms precise quantification for a moderately complex crack pattern, where the actual counts of 3 branches and 1 loop perfectly match the detection results of 3 Branches and 1 loop. This exact correspondence showcases the reliability of the algorithm in handling standard crack configurations. On the right side, a more complex crack network demonstrates the scalability of the system and robustness. While the Actual counts are 10 branches and 2 loops, the algorithm detects 13 branches and 2 loops. The visualization clearly shows that the two closed loops are perfectly identified and highlighted, confirming the strong performance of the algorithm in loop detection. The minor discrepancy in branch counting (13 detected vs. 10 actual) occurs in areas with dense, interconnected crack patterns where junction identification becomes challenging. This visual analysis validates the quantitative metrics obtained in our evaluation, demonstrating that while branch counting may experience slight over-counting in complex networks, the closed-loop detection maintains high accuracy. The consistent performance in loop identification across both simple and complex patterns underscores the method’s reliability for characterizing topological complexity in crack structures.



**Figure 9.** Results of the graph-based analysis on the CrackForest dataset, showing original cracks and their topological graphs with branch (red dots) and loop (blue regions) counts

## 8 Discussion

The integration of deep segmentation and graph-based analysis establishes a robust two-stage framework for automated crack quantification, validated across both TLCrack and CrackForest datasets. The core advancement of this work lies in the direct, automated extraction of topological features, specifically branch points and closed loops. This feature provides a quantitative and objective link to pavement maintenance decision-making. The transition from a simple linear crack to a complex network is a definitive indicator of structural deterioration.

Our analysis, based on the correlation between topological metrics and established severity levels, allows us to propose practical, quantitative guidance for engineers.

We find that a rising branch count serves as a key indicator of active crack propagation, where the development of more than three to five branches is consistently associated with moderate-severity damage, suggesting a need for preventative sealing to arrest further growth. More critically, the presence of closed loops serves as a definitive metric for advanced failure.

The emergence of even a single closed loop signals interconnected cracking, while the presence of two or more strongly indicates severe alligator or block cracking. This necessitates immediate intervention such as patching or resurfacing to restore structural integrity. By translating these topological features into actionable thresholds, our method provides maintenance crews with a direct, quantifiable basis for prioritizing repair schedules and optimizing resource allocation, moving beyond simple pixel-level segmentation.

Our statistical comparisons used paired t-tests on image-level metrics across all test folds. While this approach provides meaningful comparison between models, we acknowledge that observations within the same cross-validation fold are not fully independent. Future work could employ more sophisticated statistical methods specifically designed for cross-validation results, such as nested cross-validation or corrected repeated k-fold testing, to account for this non-independence while maintaining rigorous statistical validity.

The observed performance variation across datasets also highlights the challenges posed by domain shift and differences in image quality and resolution, which affect the precision of skeleton extraction and subsequent

topological quantification. Addressing these factors through domain adaptation and enhanced preprocessing is essential to make the framework more reliable when applied to new datasets and real-world conditions.

The qualitative results across both the TLCrack and CrackForest datasets consistently demonstrate the practical potential of the proposed graph-based quantification framework. In all examples, the system accurately identifies loop structures, even in complex crack networks, confirming the robustness of cycle detection in the graph representation. For branch detection, the method performs accurately on simpler crack patterns. As shown on the left, the detected topology matches the actual number of branches and loops. In more intricate crack networks, as illustrated on the right, the algorithm tends to slightly overestimate branch counts. This over-counting arises from small spurs and irregularities in the predicted crack skeleton, which are interpreted as branch junctions during graph extraction. Despite these minor deviations, the overall topological structure including major branches and all loop regions remains faithfully captured, demonstrating strong generalization across different datasets and crack complexities.

To address these limitations and enhance the framework's utility for maintenance planning, future work will focus on implementing morphological smoothing and pruning filters before graph skeletonization to eliminate noise-induced branches while preserving genuine topological features. We will also conduct comprehensive robustness tests under varying noise conditions to establish performance boundaries. Furthermore, we will explore advanced segmentation architectures and training strategies to produce cleaner crack outlines, thereby enhancing the input quality for the graph-based quantification stage and further solidifying the reliability of the automated branch and loop counts for infrastructure management systems. Finally, an ablation study on the specific contribution of the skeleton recall loss component will be conducted to further understand and optimize the training objective for topological accuracy.

## 9 Conclusions

Effective road maintenance planning requires moving beyond simple crack detection to robust quantification of structural complexity, particularly topological features like branches and loops. This work successfully developed a two-stage framework integrating deep segmentation and graph-based analysis for automated crack characterization, validated across both proprietary and public benchmark datasets.

Experimental results demonstrated the strong performance of the framework across different data environments, with particularly notable accuracy in closed-loop detection, highlighting the robustness of the graph-theoretic approach. The consistent performance across datasets confirms the generalizability of our methodology for topological crack analysis.

This study validates that combining deep learning with graph-based topological analysis provides direct, objective, and interpretable measures of pavement deterioration severity. The quantitative metrics generated enable reliable structural health assessment and informed maintenance decision-making for civil infrastructure systems.

Looking forward, future work will focus on enhancing the robustness of the framework and real-world applicability. Key directions include implementing morphological filtering to reduce branch over-counting caused by segmentation artifacts, improving generalizability through domain adaptation techniques, employing advanced statistical validation methods, and conducting an ablation study on the skeleton recall loss to optimize topological fidelity. These refinements will further solidify the reliability of the framework for diverse, real-world pavement management applications.

## Author Contributions

Conceptualization, V.P. and H.F.; methodology, V.P. and H.F.; software, V.P.; validation, V.P. and H.F.; formal analysis, V.P.; investigation, V.P.; resources, H.F.; data curation, V.P.; writing—original draft preparation, V.P.; writing—review and editing, V.P. and H.F.; visualization, V.P.; supervision, H.F.; project administration, H.F.; funding acquisition, H.F. All authors have read and agreed to the published version of the manuscript.

## Data Availability

The data used to support the research findings are available from the corresponding author upon request.

## Acknowledgements

The author gratefully acknowledges the support provided by the Japan International Cooperation Center (JICE) through its JDS scholarship program for doctoral studies in Japan.

## Conflicts of Interest

The authors declare no conflict of interest.

## References

- [1] C. P. Ng, T. H. Law, F. M. Jakarni, and S. Kulanthayan, "Road infrastructure development and economic growth," *IOP Conf. Ser.: Mater. Sci. Eng.*, vol. 512, p. 012045, 2018. <https://doi.org/10.1088/1757-899X/512/1/012045>
- [2] A. Mohan and S. Poobal, "Crack detection using image processing: A critical review and analysis," *Alexandria Eng. J.*, vol. 57, no. 2, pp. 787–798, 2018. <https://doi.org/10.1016/j.aej.2017.01.020>
- [3] X. Feng, L. Xiao, W. Li, L. Pei, Z. Sun, Z. Ma, H. Shen, and H. Ju, "Pavement crack detection and segmentation method based on improved deep learning fusion model," *Math. Probl. Eng.*, vol. 2020, no. 1, p. 8515213, 2020. <https://doi.org/10.1155/2020/8515213>
- [4] D. Ai, G. Jiang, S. Lam, P. He, and C. Li, "Computer vision framework for crack detection of civil infrastructure—A review," *Eng. Appl. Artif. Intell.*, vol. 117, p. 105478, 2023. <https://doi.org/10.1016/j.engappai.2022.105478>
- [5] Z. Mao, X. Ma, M. Geng, M. Wang, G. Gao, and Y. Tian, "Development characteristics and quantitative analysis of cracks in root-soil complex during different growth periods under dry-wet cycles," *Biogeotechnics*, vol. 3, no. 1, p. 100121, 2025. <https://doi.org/10.1016/j.bgtech.2024.100121>
- [6] Z. Fan, C. Li, Y. Chen, P. Di Mascio, X. Chen, G. Zhu, and G. Loprencipe, "Ensemble of deep convolutional neural networks for automatic pavement crack detection and measurement," *Coatings*, vol. 10, no. 2, p. 152, 2020. <https://doi.org/10.3390/coatings10020152>
- [7] Z. Li, T. Zhang, Y. Miao, J. Zhang, M. Eskandari Torbaghan, Y. He, and J. Dai, "Automated quantification of crack length and width in asphalt pavements," *Comput. Aided Civ. Infrastruct. Eng.*, vol. 39, no. 21, pp. 3317–3336, 2024. <https://doi.org/10.1111/mice.13344>
- [8] Y. Zhou, Y. Huang, Q. Chen, and D. Yang, "Graph-based change detection of pavement cracks," *Autom. Constr.*, vol. 174, p. 106110, 2025. <https://doi.org/10.1016/j.autcon.2025.106110>
- [9] Y. Wang, Z. He, X. Zeng, J. Zeng, Z. Cen, L. Qiu, X. Xu, and Q. Zhuo, "GGMNet: Pavement-crack detection based on global context awareness and multi-scale fusion," *Remote Sens.*, vol. 16, no. 10, p. 1797, 2024. <https://doi.org/10.3390/rs16101797>
- [10] L. C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 801–818. [https://doi.org/10.1007/978-3-030-01234-2\\_49](https://doi.org/10.1007/978-3-030-01234-2_49)
- [11] O. Oktay, J. Schlemper, L. Le Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, B. Kainz *et al.*, "Attention U-Net: Learning where to look for the pancreas," in *the 1st Conference on Medical Imaging with Deep Learning (MIDL)*, Amsterdam, Holland, 2018.
- [12] E. Xie, W. Wang, Z. Yu, A. Anandkumar, J. M. Alvarez, and P. Luo, "SegFormer: Simple and efficient design for semantic segmentation with transformers," *Adv. Neural Inf. Process. Syst.*, vol. 34, pp. 12 077–12 090, 2021.
- [13] Y. Kirchhoff, M. R. Rokuss, S. Roy, B. Kovacs, C. Ulrich, T. Wald, M. Zenk, P. Vollmuth, J. Kleesiek, F. Isensee, and K. Maier-Hein, "Skeleton Recall Loss for connectivity conserving and resource efficient segmentation of thin tubular structures," in *Computer Vision—ECCV 2024, Lecture Notes in Computer Science*. Cham: Springer, 2024, vol. 15135, p. 218–234. [https://doi.org/10.1007/978-3-031-72980-5\\_13](https://doi.org/10.1007/978-3-031-72980-5_13)
- [14] F. Milletari, N. Navab, and S. A. Ahmadi, "V-Net: Fully convolutional neural networks for volumetric medical image segmentation," in *Proceedings of the 2016 Fourth International Conference on 3D Vision (3DV)*, Stanford, CA, USA, 2016. <https://doi.org/10.1109/3DV.2016.79>
- [15] S. Shit, J. C. Paetzold, A. Sekuboyina, I. Ezhov, A. Unger, A. Zhylyka, J. P. W. Pluim, U. Bauer, and B. H. Menze, "cIDice—A novel topology-preserving loss function for tubular structure segmentation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, TN, USA, 2021, pp. 16 560–16 569. <https://doi.org/10.1109/CVPR46437.2021.01629>
- [16] X. Hu, F. Li, D. Samaras, and C. Chen, "Topology-preserving deep image segmentation," *Adv. Neural Inf. Process. Syst.*, vol. 32, pp. 1–12, 2019.
- [17] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.
- [18] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *Int. J. Forecast.*, vol. 22, no. 4, pp. 679–688, 2006. <https://doi.org/10.1016/j.ijforecast.2006.03.001>
- [19] L. Z. Liu, A. Zhou, X. R. Ran, Y. P. Wu, W. G. Zhao, and H. Zhang, "A crack detection and quantification method using matched filter and photograph reconstruction," *Sci. Rep.*, vol. 15, p. 25266, 2025. <https://doi.org/10.1038/s41598-025-08280-z>
- [20] H. Bae and Y. K. An, "Computer vision-based statistical crack quantification for concrete structures," *Measurement*, vol. 211, p. 112632, 2023. <https://doi.org/10.1016/j.measurement.2023.112632>
- [21] L. A. S. Calderón, "A system for crack pattern detection, characterization and diagnosis in concrete structures by means of image processing and machine learning techniques," 2017, phdthesis, Universitat Politècnica de

Catalunya.

- [22] Y. Liu and J. K. W. Yeoh, "Automated crack pattern recognition from images for condition assessment of concrete structures," *Autom. Constr.*, vol. 128, p. 103765, 2021. <https://doi.org/10.1016/j.autcon.2021.103765>
- [23] J. Yu, Y. Xu, C. Xing, J. Zhou, and P. Pan, "Pixel-level crack detection and quantification of nuclear containment with deep learning," *Struct. Control Health Monit.*, vol. 2023, no. 1, p. 9982080, 2023. <https://doi.org/10.1155/2023/9982080>
- [24] L. Deng, A. Zhang, J. Guo, and Y. Liu, "An integrated method for road crack segmentation and surface feature quantification under complex backgrounds," *Remote Sens.*, vol. 15, no. 6, p. 1530, 2023. <https://doi.org/10.3390/rs15061530>
- [25] A. Br  heret, "Pixel Annotation Tool," 2017. <https://github.com/abreheret/PixelAnnotationTool>
- [26] Y. Shi, L. Cui, Z. Qi, F. Meng, and Z. Chen, "Automatic road crack detection using random structured forests," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 12, pp. 3434–3445, 2016. <https://doi.org/10.1109/TITS.2016.2552248>
- [27] S. Chen, G. Fan, J. Li, and H. Hao, "Automatic complex concrete crack detection and quantification based on point clouds and deep learning," *Eng. Struct.*, vol. 327, no. 15, p. 119635, 2025. <https://doi.org/10.1016/j.engstruct.2025.119635>
- [28] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*, ser. Lecture Notes in Computer Science, vol. 9351. Cham: Springer, 2015. [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)
- [29] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *J. Big Data*, vol. 6, no. 60, 2019. <https://doi.org/10.1186/s40537-019-0197-0>
- [30] I. Loshchilov and F. Hutter, "Decoupled weight decay regularization," in *Proceedings of the International Conference on Learning Representations (ICLR 2019)*, New Orleans, LA, USA, 2019.
- [31] T. Y. Zhang and C. Y. Suen, "A fast parallel algorithm for thinning digital patterns," *Commun. ACM*, vol. 27, no. 3, pp. 236–239, 1984. <https://doi.org/10.1145/357994.358023>