Diyala Journal of Engineering Sciences

Journal homepage: https://djes.info/index.php/djes

ISSN: 1999-8716 (Print); 2616-6909 (Online)

# Deep Learning-Based Detection, Segmentation, and Quantification of Asphalt Pavement Cracks

Shemeam T. Muhey*, Sinan A. Naji

Informatics Institute for Postgraduate Studies, University of Information Technology and Communications. Iraq

**ARTICLE INFO**

***Article history:***
Received March 25, 2025
Revised May 14, 2025
Accepted June 11,2025
Available online September 01,2025

***Keywords:***
YOLOv10
UNet 3+
Attention gate
Residual unit
Cracks identification

**ABSTRACT**

The primary factor influencing road performance is pavement deterioration. Pavement cracking, a prevalent form of road deterioration, is a significant challenge in road maintenance. This paper proposes a method utilizing deep convolutional neural network models for precise crack detection, segmentation, and geometric parameter calculation in pavement crack identification. The system operates through three primary stages: Commencement, crack identification employs YOLOv10, a rapid and efficient object detection model. Secondly, crack segmentation employs a modified Unet 3+ variant known as Residual-Attention UNet 3+, which effectively distinguishes crack pixels from the background by utilizing attention mechanisms and residual connections to enhance accuracy. Finally, crack quantification, wherein the system computes the crack's geometric parameters, including width, length, angle, and orientation. We assessed performance using two datasets: SUT-Crack, a publicly accessible dataset, and IRD-Crack, a new real-world dataset compiled by the authors from roads in Diyala, Iraq, with diverse lighting conditions and surface complexities. The suggested technique attained an accuracy of 98.96% on the SUT-Crack dataset. It showed superior performance on the IRD-Crack dataset under actual situations, therefore validating its efficacy and generalization capability. This method offers a pragmatic and computationally efficient instrument for monitoring pavement cracks and can facilitate road repair choices.

## 1. Introduction

Pavement deterioration is the primary element influencing road performance. The prompt and precise identification of pavement deterioration is essential for pavement upkeep. Cracks are the primary indication of multiple forms of pavement deterioration. Pavement cracks would adversely impact both the aesthetic quality and driving comfort while also potentially escalating to induce structural damage and diminish the overall service performance and lifespan of the pavement [1]. Therefore, early detection of pavement cracks is crucial for preventing pavement degradation, safeguarding the underlying foundation layers, minimizing maintenance efforts and expenses, and ensuring safety for all road users.

Early pavement identification and repair usually depend on manual detection, which is time-consuming and laborious, has weak detection accuracy, and has some associated dangers [2]. From crack classification, it is revealed that they exhibit diverse shapes, extensive coverage areas, varying extension lengths, and irregular widths.

* Corresponding author.
E-mail address: ms202220727@iips.edu.iq
DOI: 10.24237/djes.2025.18305

This work is licensed under a Creative Commons Attribution 4.0 International License.

In recent years, the fast development of computer technology and Artificial intelligence (AI) has led to its integration across numerous fields. In recent years, it has become extensively utilized for road crack identification, resulting in numerous automated detection techniques. Traditional automatic detection techniques encompass the Canny algorithm, which relies on threshold segmentation [3], and the Otsu approach [4]. Nevertheless, owing to the intricate characteristics of the road surface and pavement environment, along with the general applicability and robustness of the traditional Canny and Otsu algorithms, the accuracy of the detection findings is suboptimal. Subsequently, the minimal cost path search algorithm [4], the support vector machine (SVM) detection algorithm [5], the Crack Tree detection technique, and others emerged. These techniques have solved low Precision, however, detecting Precision still takes a long time. Additionally, the complex architecture of detection and Crack Tree techniques using SVM prevents their practical application. Based on these difficulties and the prevalent machine learning technology [6], a deep learning and neural network-based automatic crack identification and detection technique is developed. Crack detection data can also monitor pavement conditions and determine road maintenance strategies. Thus, pavement crack identification would considerably impact road monitoring and maintenance automation, thus its precision and speed must be improved.

Currently, the direction of Road pavement crack recognition research is separated into two sections. The digital image processing method uses artificial feature identification, including frequency, edge, HOG, gray level, texture, and entropy, to construct feature recognition conditions for limited and total recognition. The second type uses deep learning to create a Convolutional Neural Network (CNN) for automatic feature recognition. The network adjusts to meet or exceed label accuracy by following specific rules. This paper establishes a deep learning-based convolutional network to detect pavement cracks automatically.

Considering the aforementioned issues in pavement crack detection, this paper proposes a method for pavement crack identification utilizing a deep convolutional neural network fusion model, which effectively identifies cracks and ensures recognition accuracy through the YOLOv10 model. A detected crack can be segmented using Residual-Attention UNet 3+, and the resulting binary image can be utilized to compute the geometry characteristics of the crack. Consequently, the suggested model holds substantial importance for intelligent pavement detection and can concurrently perform detection and segmentation, thereby markedly enhancing model efficiency.
The main contributions to the suggested model include the following:

1. The suggested system employs YOLOv10 for object detection because it effectively resolves a significant obstacle in organizational development: balancing accuracy and computational efficiency compared to earlier versions.
2. This study presents a novel technique for image segmentation. We introduced Residual-Attention UNet 3+, a composite neural network that amalgamates the advantages of UNet 3+, residual units, and attention mechanisms for the segmentation of crack images. This technique has improved predicted accuracy relative to earlier methods, hence differentiating our approach from prior methodologies. This results in attaining a high level of precision in the identification of pavement cracks.
3. Utilizing another dataset by capturing pavement crack images of local roads to test the proposed system in order to reflect its applicability and generalization in the real environment.

This paper is organized into the following sections: Section 2 presents the related research, Section 3 outlines our methodology, and Section 4 details the experiments and analysis. In conclusion, Section 5 encapsulates the entirety of the work.

## 2. Related works

In recent years, the automated identification of pavement cracks has garnered heightened interest. Authors and maintenance specialists are exploring diverse strategies and methodologies to improve maintenance dependability and efficacy. This section summarizes the literature in this field. Li et al. presented an interesting form of the road crack detection model called RDD-YOLO [7]. The model combined a simple attention mechanism (SimAM) to the backbone network to bring attention to significant details in the input image. By using GhostConv instead of traditional convolution modules, the neck structure is enhanced. As a result, the task of damage recognition will execute more lightweight and effective because there is less redundant data, fewer parameters, and less computing complexity. Lastly, the upsampling algorithm in the neck is improved by replacing the nearest interpolation with more accurate bilinear interpolation. This finer interpolation method more successfully restores the image's delicate details and improves the accuracy of the detection results. The proposed model achieves an mAP50 and mAP50-95 of 62.5% and 36.4% on the validation set respectively on the RDD2022 dataset. This study is constrained by its dependence on a singular dataset (RDD2022), perhaps limiting its applicability to diverse pavement or lighting situations. Deng et al. suggested an integrated framework for automatic detection, segmentation, and measurement of road surface [8]. In the proposed framework, three different computer vision algorithms are effectively combined: First, to identify cracks, the real-time object detection algorithm YOLOv5 is employed, it achieves a mean average precision (mAP) of 91%. Secondly, a modified ResNet is created by adding an attention gate module to increase accuracy of cracks segmentation at the pixel level which achieves 87% intersection over union (IoU) on crack pixels segmentation. Lastly, an innovative surface feature quantification method is created to measure both the width and length of segmental road cracks, achieving a 95% identification accuracy. However, the framework presupposes optimal conditions and fails to include real-world environmental fluctuations, such as shadows, debris, or illumination discrepancies. Shu et al. presented a pavement crack detection model that utilizes the YOLOv5 target detection network with the street view image data source [9], which is a cost-effective method. With a mAP of more over 70%. However, the model has difficulties in identifying small or hairline fractures due to the constrained resolution and noise inherent in street view data. An et al. suggested a system called the Crack Identification Network (CIN) [10] for identifying and calculating the size of concrete surface cracks by integrating deep learning convolutional neural networks, clustering segmentation and morphological techniques. The accuracy rate achieves 99%, although the approach demonstrates great accuracy, its computational complexity and absence of real-time performance may impede its implementation in practical settings. Zhang Z. et al. introduced the ResUnet, a semantic segmentation neural network, which gathers the strength of residual learning and U-Net from high-resolution remote sensing images [11]. The first benefit of this model is that residual units facilitate deep network training. Second, information might spread more easily due to the network's rich skip connections, making it possible to build networks with fewer parameters but better performance. The suggested method's break-even points which defined as the point on the relaxed precision-recall curve, was 0.9187. Nonetheless, it is computationally demanding and inappropriate for implementation on low-resource devices or in real-time applications. Zhang Q. et al. introduced an improved U-net network for crack detection and segmentation with a complicated background [12]. The VGG16 and the novel Up_Conv module are added as the backbone network to increase the recognition accuracy of small cracks in the road surface. Moreover, U-net's skip connection was enhanced using the Ca (Channel Attention) mechanism to distinguish between cracks and background noise. In order to extract richer information through more convolutional layers in the network, the

DG_Conv (Depthwise GSConv Convolution) and UnetUp (Unet Upsampling) modules are introduced in the decoding stage. The suggested system's results show a precision of up to 87.4%. However, the model's efficacy is contingent upon hyperparameter configurations, and the research is deficient in detail on its resilience under diverse environmental circumstances. He and Lau put out an interesting model called CrackHAM. This encoder-decoder network is based on the U-Net design and incorporates a novel model network called the HASP module to address the problem of deteriorating spatial data [13]. Additionally, the channel attention and spatial attention modules were used to capture abundant contextual information for high-level features and extract rich edge information for low-level features respectively. Through downsampling, the Multi-Fusion U-Net architecture is suggested as a way to aggregate contextual information from feature maps of different sizes. The accuracy of the system is 86.41%. Nonetheless, the model's attention methods introduce a considerable computational burden, rendering it less appropriate for real-time crack investigation. Zhang et al. employ an innovative technique that combines a Convolutional Block Attention Module (CBAM) with a ResNet model to identify multi-type cracks [14]. The suggested model achieves a precision of 92.9%. Nevertheless, the study is hindered by its geographically narrow dataset, which may not accurately reflect different road conditions worldwide. Table 1 illustrates a summary of related works.

**Table 1:** Summary of related works

| Study | Technique | Dataset | Performance Metrics |
|---|---|---|---|
| [7] | YOLOv8 and simple attention mechanism | RDD2022 | mAP50: 62.5% and mAP50-95: 36.4% |
| [8] | YOLOv5 and Attention ResNet | (RDD) dataset for training and validation, and Road-Crack-Images-Test from Hunan University | mAP: 91% IoU: 87%, Accuracy: 95% |
| [9] | YOLOv5 | Street view images | mAP > 70% |
| [10] | CNN, clustering segmentation and morphology | Collect 1000 crack original images. | Accuracy: 99% |
| [11] | residual learning and U-Net | Massachusetts roads dataset | Precision-recall:.91.87% |
| [12] | Improved U-Net and VGG16 | CFD and Deepcrack | Precision: 87.4% |
| [13] | Multi-Fusion U-Net | Deepcrack, Crack500, and FIND | Accuracy:86.41%. |
| [14] | Convolutional Block Attention Module and ResNet model | Collect crack images from China streets | Precision: 92.9% |

## 3. Materials and methods

This study provides a fully automated procedure for the detection, segmentation, and measurement of asphalt pavement cracks located in different shapes and forms within the image , as illustrated in Figure 1 two methodologies have been used in our system: first, object detection using YOLOv10 which is a single-stage object detection method, offers rapid detection speed and effective identification of small targets. YOLOv10 enhances both precision and efficiency via a synthesis of training methodologies and architectural developments. This method has been utilized in numerous engineering applications and is particularly effective for crack-detecting tasks that include stringent time limitations and significant safety threats. Consequently, the YOLOv10-based method [15] is initially employed to detect road crack areas utilizing bounding boxes. Second, semantic segmentation using the improved Residual-Attention UNet 3+ algorithm. The outcomes of Stage 1 are input into the improved Residual-Attention UNet 3+ algorithm as Stage 2. To enhance pixel-level crack segmentation, we have refined the UNet 3+ model by developing an integrated neural network that combines the advantages of UNet 3+, residual units, and attention gates (AG) for crack image

segmentation. In Stage 3, a novel approach for estimating surface cracks is introduced to measure the length, width, and orientation of the segmented cracks. The primary benefit of the suggested model is the substantial enhancement in the precision and efficacy of road crack segmentation among intricate backgrounds. Simultaneously, a novel technique for surface feature estimation has been devised to examine surface feature data, emphasizing crack morphology precisely [16]. The specifics of each phase of the planned architecture are explained in the subsequent subsections.

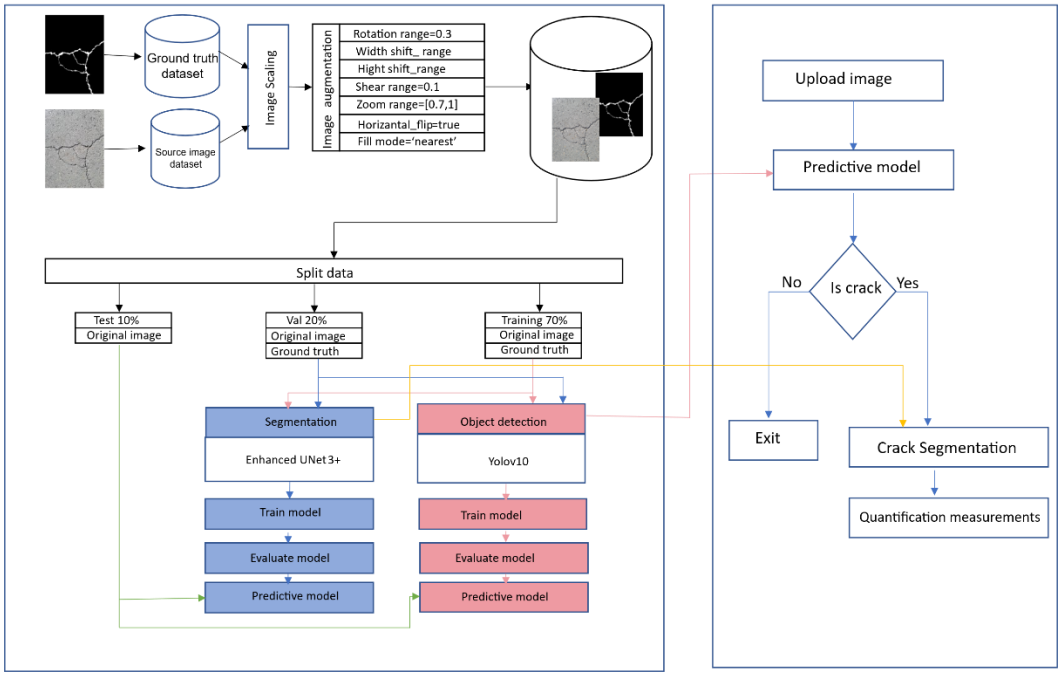



**Figure 1.** The general architecture of system

### *3.1 Image pre-processing*

Image pre-processing is an essential stage in deep learning that increases the quantity and quality of dataset images required for system training and yields a more efficient learning model. cropping, flipping, rotation, improvement of contrast, colour-space transformation, noise reduction, and colour enhancement, are examples of pre-processing methods [17]. Various approaches performed to the SUT-Crack Datasets are shown in the following sections.

#### *3.1.1 Image scaling*

All of the images used in this study were resized to 640 x 640 x 3 and 320 x 320 x 3 in order to be compatible with the inputs utilized by the YOLOv10 and UNet 3+ models, respectively. In addition to ensuring computational efficiency, image scaling provides the model with a standard input size.

#### *3.1.2 Image Augmentation*

Large datasets are typically needed during the training phase of deep learning with CNN-based methodologies in order to improve the ability of the model to learn new image patterns and generate accurate predictions. The augmentation process improves the training dataset by using multiple image transformations. Rotation, shifting, shearing, zooming, flipping, and reflecting are a few examples of these transformations. Through the production of new images from the dataset of asphalt cracks, overfitting is mitigated, undesired feature acquisition is avoided, and overall performance is enhanced [17,18]. The various transformation types and their associated parameters are displayed in Table 2.

**Table 2:** Dataset augmentation with different transformations.

| Transformation Type | Corresponding Values |
|---|---|
| Range of Rotation | 30 degrees |
| Range of Width-Shift | 10% |
| Range of Height-Shift | 10% |
| Range of Shear | 10% |
| Range of Zoom | [70% - 100%] |
| Horizontal-Flip | 'True' |
| Fill Mode Reflection | 'Nearest' |

*3.1.3 Splitting the dataset*

A common technique in machine learning, data mining, pattern recognition, and other fields is splitting the dataset into smaller sub-datasets. SUT-Crack datasets were split into three subsets for this study: 70% for training, 20% for validation, and 10% for testing. comprehensive information is available in Table 3.

**Table 3:** Details of splitting dataset

| Training (70%) | Validation (20%) | Testing (10%) |
|---|---|---|
| 5756 | 1644 | 822 |

*3.2 YOLOv10 for crack detection*

The initial stage of the suggested model is crack detection (object detection) using YOLOv10 which is employed to identify road cracks in images. The purpose of object detection techniques is to locate and classify an object in image by drawing bounding box around object region [11,19]. Popular models like YOLO (You Only Look Once) [20] or Faster R-CNN [21] are often adapted for this task. These models don't detect the exact pixel boundaries but rather identify the crack as an object within the image [22]. One of the greatest real-time object detection algorithms is the YOLO (You Only Look Once) series (from v1 to v10), a single-stage object detection technique developed in recent years. It significantly and broadly affects several computer vision studies [23,24]. Numerous tests demonstrate that YOLOv10 is superior to other advanced detectors by achieving state-of-the-art performance and latency [25], through fusing training techniques and architectural innovations YOLOv10 improves accuracy and efficiency. As following, we introduce some of the special features of YOLOv10: The road crack images are initially processed by the backbone to extract crack features, subsequently, feature fusion is conducted in the neck utilizing an Enhanced Feature Pyramid Network (FPN) with spatial-channel decoupling and Partial Self-Attention (PSA) [15]. ultimately, the outputs consist of predicted values for class probability, item level, and bounding box location of road cracks. YOLOv10's architecture comprises three components: backbone, neck, and head. Below is an explanation of the basic components:

a. The initial component is the backbone, primarily responsible for feature extraction from the input image. YOLOv10's backbone employs an advanced generation of CSPNet (Cross-Stage Partial Network) to boost gradient flow and minimize computational redundancy. CSPDarknet has numerous essential modules, such convolutional layers, batch normalization, activation functions, and residual blocks. A

vital element of CSPDarknet is the Cross Stage Partial (CSP) connections, which partition maps features into two sections and integrate them via a cross-stage hierarchy to enhance learning efficiency. Furthermore, spatial-channel decoupled-down sampling is implemented to improve computing efficiency. Additionally, YOLOv10 integrates large-kernel convolutions and partial self-attention methods during the feature extraction phase, enhancing detection precision while preserving computing efficiency. The enhancements in the backbone architecture enable YOLOv10 to attain enhanced efficacy in object detecting tasks.

b. The architecture's neck integrates a path aggregation network (PAN) module, optimized for efficiency, along with up sampling layers to improve feature map resolution; it comprises an FPN and a PSA positioned between the backbone and head layers. Utilizing an FPN architecture facilitates the transmission of substantial semantic attributes from the highest to the lowest feature maps. This design guarantees the accuracy of minor object details while enabling the abstract representation of big objects. The PSA architecture transmits precise localization data across feature maps of differing granularity. By integrating the FPN and PSA, YOLOV10 improves efficiency through the PSA module and the Compact Inverted Bottleneck (CIB) block, facilitating effective multi-scale feature processing and attention mechanisms. Consequently, the neck attains adequate power for feature fusion.
c. The predictive header eliminates the need for non-maximum suppression (NMS) used by previous versions; a technique used to eliminate duplicate predictions and select the most confidently selected boxes. By introducing a double-assignment strategy into its training process, it thus significantly reduces processing time. Finally, the predicted specified box is generated, and the object is classified and labelled. During post-processing, confidence criteria, typically established at 0.25, and Intersection-over-Union (IoU) thresholds, set at 0.45, are employed to eliminate weak detections. The resulting bounding boxes are then transformed into image-scale coordinates for visual overlay and annotation.

### *3.2 Improved Residual-Attention UNet 3+ for Crack Segmentation*

We have improved UNet 3+ model by constructing an integrated neural network that combines the strengths of UNet 3+, residual unit, and attention gate (AG) to carry out crack semantic segmentation. Semantic Segmentation involves assigning a class label to every pixel in an image into a pre-defined set of categories, such as road, building, or vehicle. Since FCN, U-Net, and their variations predict one segmentation map based on pixel-wise classification, they have been extensively used for semantic segmentation across a variety of applications [26,27]. The core network framework used in the suggested model is UNet 3+, which connects the encoder and decoder networks using deep supervision and full-scale skip connections [28]. Following each encoding step (E1 to E4), the encoder network's feature map was translated to the decoder network using dense convolution blocks and a residual block (Conv+Maxpooling+Dropout(0.2)). We inserted an attention gate between (E4-D4) to help the model focus on the most significant features and disregard the unimportant ones. It takes two inputs g, gate signal comes from the next lowest layer of the network (decoder stage), which has the better features and x, comes from skip connection at early layers (encoder stage). An element-wise sum is performed on the two vectors. Because of this process, aligned weights get bigger and unaligned weights get smaller. A ReLU activation function is applied to the resulting vector. The attention coefficients (weights) are produced by scaling this vector between [0,1] using a sigmoid layer; more relevant features are indicated by coefficients closer to 1. Trilinear interpolation is used to up-sampling the attention coefficients to the x vector's original dimensions. The original x vector is scaled based on significance by

multiplying the attention coefficients element by element. The skip connection then transmits this as usual [29]. The general structure of Improved Residual-Attention UNet 3+is depicted in Figure 2.

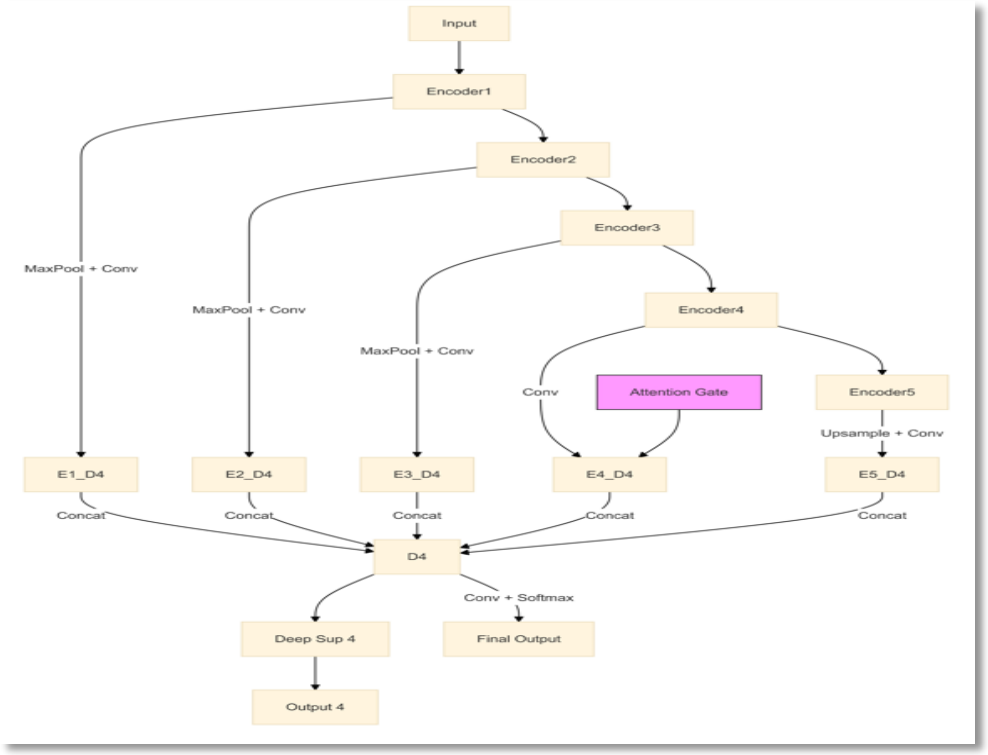



**Figure 2**. Enhancement of UNet 3+ Model Architecture

*3.2.1 UNet 3+*

Which benefits full-scale skip connections and deep supervisions. While the deep supervision learns hierarchical representations from the full-scale aggregated feature maps, the full-scale skip connections combine low-level details with high-level semantics from feature maps on different scales. The main advantage of UNet 3+ is its ability to be efficiently trained on small datasets [28]. The primary architecture of the UNet 3+ is consists of two main parts: Encoder and Decoder. The encoder means a chain of convolutional layers that capture high-level features. Each decoder layer in Unet 3+ includes both smaller- and same-scale feature maps from the encoder and larger-scale feature maps from the decoder, which capture fine-grained features and coarse-grained semantics in complete sizes. The basic architecture of Unet 3+ also contains skip connections, The basic idea of skipping connections is that as the encoder lowers the spatial resolution, which can cause a loss of fine details, the skip connections assist in maintaining spatial details by directly transmitting them to the decoder. For example, Figure 3 shows how to extract the feature map of $X_{De}^{3}$. Like the UNet, the decoder receives the feature map from the same-scale encoder layer $X_{En}^{3}$directly. Unlike to the UNet, a set of inter encoder-decode skip connections transfers the low-level detailed information from the smaller-scale encoder layer $\mathrm{X}_{\mathrm{En}}^{1}$' and $\mathrm{X}_{\mathrm{En}}^{2}$ , by using non-overlapping max pooling operation; while by applying bilinear interpolation a chain of intra decoder skip connections transfers the high-level semantic information from larger-scale

decoder layer $X_{De}^{4}$and $X_{De}^{5}$. As a result, it will be formed five same resolution feature maps. To eliminate unnecessary information and further standardize the number of channels a convolution with 64 filters of size 3 × 3 could be a good option. Furthermore, a feature aggregation process, comprising 320 filters of size 3 × 3, batch normalization, and a ReLU activation function, has been applied on the concatenated feature map from five scales in order to smoothly combine the shallow exquisite information with deep semantic information [28].

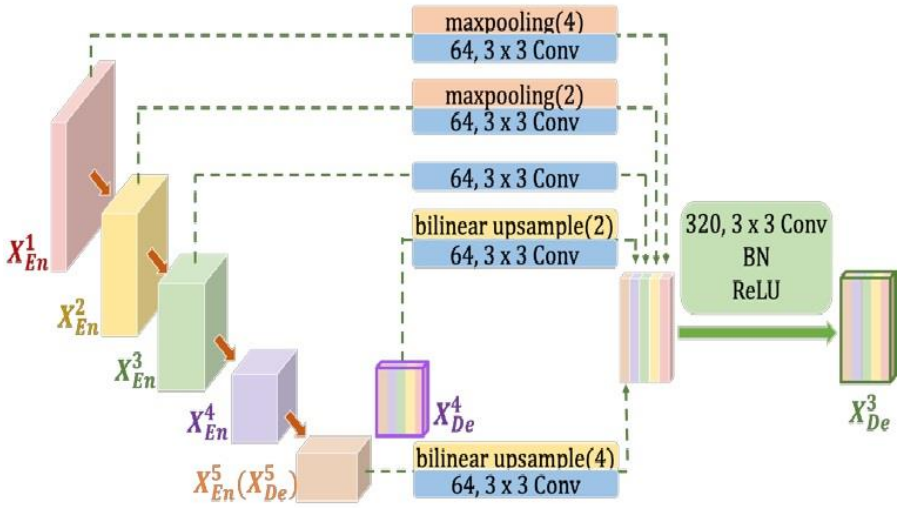



**Figure 3.** Illustration of how to construct the full-scale aggregated feature map of third decoder layer $X_{De}^{3}$ in original Unet 3+

### *3.2.2 Residual Blocks or Units*

A series of stacked layers make up residual blocks or units, in which inputs are added back to their outputs in order generate identity mappings. In practice, identity mappings are implemented using what are known as skip or residual connections. However, there are several possible ways to apply these connections, depending on where they are inserted within the stacked layers that form a residual block [30]. According to learning theory, deeper neural networks should achieve lower training and test error, but in practice, the opposite occurs. Once the error rate reaches a minimum value, the error rate starts increasing again. The exploding and vanishing gradient descent problem is the source of this, as it leads to overfitting of the model and an increase in error, Fortunately, Residual Networks have proved to be quite efficient in solving this problem because they employ a skip connection or a "shortcut" between every two layers along with using direct connections between all the layers [31].

### *3.2.3 Attention mechanisms*

The main idea behind attention mechanisms is to recognize the most important elements of feature maps in convolutional neural networks (CNNs) that the redundancy is removed for machine vision applications. Attention mechanisms generate attention maps that help CNNs focus on important spatial or channel-wise features [32,33].

### *3.3 Crack Quantification*

We utilized deep learning techniques using the improved UNET 3+ algorithm to extract the crack precisely. Nevertheless, the dimensions of the crack remain indeterminate. The pavement crack is often measured in terms of width, length, and depth, all of which are critical indications that assess the severity of the crack and inform the restoration plan. In most studies, crack quantification is performed on the anticipated binary crack map using image

processing techniques and geometric calculations. However, the morphological characteristics of cracks are not thoroughly addressed, which reduces efficiency and accuracy.

At this stage, the suggested model is subjected to a case study to verify its robust and dependable performance in a real-world environment. We used a dataset comprising asphalt crack images from local Iraqi roadways. The segmentation model produces segmented images. Through this rigorous testing, the proposed model can be validated for its effectiveness in solving many safety problems, improving road performance, and reducing maintenance costs.

We provide a region-connected search method based on the linked component of cracks to make the visible cracks more comprehensive, distinct, and consistent with the real trend of cracks. Following the acquisition of the crack binary image's contour coordinates, the crack's length is computed using the coordinates that were obtained, and the average crack width is computed by dividing the crack's length by the area of the linked component. The contour is examined after it is sketched in the image of the crack area. The contour is analysed. Finally, the results of the crack length and width are displayed in the crack image, as shown in Figure 4, which shows the steps of crack quantification.

i. Image Pre-processing: As shown in Figure 5, A series of operations is applied to find and analyze the contour of the crack (Convert image to Gray-scale, to blur the image, apply Gaussian filter [34], to convert the image pixels to a binary image, apply adaptive Thresholding, Morphological Operations, most common morphological operations are erosion and dilation. Erosion removes pixels from image borders, whereas dilation adds pixels. Morphological processing removes tiny cracks and fills gaps in detected cracks, improving crack detection accuracy [35].


**Figure 4.** Crack quantification by real-world data steps


**Figure 5.** Image pre-processing steps

ii. Find Optimal Contour: We relied on geometric analysis of the crack contour to accurately detect real cracks, which helps in filtering real cracks from noise in the image. To achieve this, we applied canny edges filter to find the edges (connected components) and then find the contour using the OpenCV, a “library in Python programming language,” simplifies locating and drawing crack contours through two basic functions: find-Contours () and draw-Contours () [35].

The optimal contour was chosen based on the area. The contour area is then calculated using the function counter area (); small area contours are neglected because they often represent noise or unimportant details, while large areas are considered because they represent mostly cracks.

The optimal crack is calculated by assuming a minimum threshold (Min value), so areas smaller than the previously specified value are neglected, and areas larger than the specified value are mostly considered a contour of cracks.

After that, we verify whether the contour represents a crack or not by determining the smallest rectangle surrounding the contour through (GetMinAreaRect(contour)), and the width and length of the rectangle are calculated. Through (aspect-ratio = length/width) and if the cracks are real, this ratio is greater or equal to (3), but if this ratio is less, it is not considered a crack [36]. This ratio was adopted as a minimum because studies of crack analysis in road engineering and materials science show that actual cracks in substructures have a length-to-width ratio ranging from 3 to 20 or more... If the ratio is less than 3, this means that the shape is square or circular, but if the ratio is greater than 3, this means that the shape is longitudinal and thin, as the cracks are considered longitudinal and thin, which is caused by mechanical stress and thermal changes, which leads to linear cracking, making them much longer than their width.

iii. Crack analysis

The algorithm (1) includes analyzing cracks to determine their dimensions (width, length, angle, and orientation), which will then be translated into actual measurements. This algorithm mixes mathematical geometry, image analysis, and engineering data processing to present an accurate and efficient method for analyzing cracks in infrastructure. It generates interpretable results, making it suitable for practical applications in road maintenance. We employ a specified angle threshold (Angle Threshold = 30°) to consistently classify crack direction into horizontal, vertical, and diagonal orientations. Previous investigations corroborate this threshold, which indicated that an angular tolerance of 25°–35° was efficient in categorizing cracks under diverse situations. The 30° threshold value signifies an effective equilibrium between accuracy and tolerance in practical settings, particularly where fractures may display minor angular variations due to surface defects or perspective distortions.

**Algorithm (1): Analysis Crack**

Input: Valid crack contour , Angl_Threshold=30

Output: crack properties

```
Begin
Step 1: Get rotated rectangle properties:
        rect = GetMinAreaRect(contour)
Step 2: Calculate dimensions:
        width = Min(rect.width, rect.height)
        length = Max(rect.width, rect.height)
        angle = rect.angle
Step 3: Normalize angle
      area = CalculateArea(contour)
      IF width > length:
         angle = angle + 90
      END IF
Step 4: Determine orientation
      IF (Angle < Angl_Threshold (Angle > (180 -
      Angl_Threshold))Then
         orientation ="horizontal"
      IF |Angle - 90| < Angl_Threshold Then
      orientation ="vertical"
      IF Angle < 90 Then orientation ="diagonal-
      right"
      IF Angle > 90 Then orientation = "diagonal-
      left"
Step 5: return{contour: contour,
                    width: width,
                   length: length,
                    angle: angle,
                  center: CenterPoint,
                  orientation: orientation}
End
```

iv. Convert to real measurements

It is necessary to translate the resultant measurements (in pixels) to millimetres (mm) for informed decision-making in road maintenance. We derive the conversion factor for applying it to the real measurements of

length, width, and area, as demonstrated in the subsequent equation [37]:

$$Conversion\ Factor\ (CF) = \frac{Reference\ width\ (mm)}{Reference\ width\ (pixel)} \quad (1)$$

$$length_{mm} = CF * Length_{pixel} \quad (2)$$

$$width_{mm} = CF * width_{pixel} \quad (3)$$

$$Area_{mm^2} = CF^2 * Area_{pixel^2} \quad (4)$$

## 4. Experiments results and analysis

### *4.1 Implementation details and Dataset Collection*

The training utilized the Adam optimizer, including a learning rate of 0.0001 and a batch size of 32 at 100 epochs. All tests were conducted with TensorFlow on a Windows 10 computer with an Intel Core i7 running at 3.60 GHz and 16 GB of RAM. In this study, we conducted experiments using two datasets to get more accurate results:

In this study, we conducted experiments using two datasets to get more accurate results:

- Set 1: The "SUT-Crack Dataset "which contains 130 high-resolution images in jpg format, with dimensions of 3024 by 4032 pixels [38]. The images are organized dually, meaning that each original image is matched by its corresponding ground truth image, as shown in Figure 6. SUT-Crack is available at https://doi.org/10.17632/gsbmknrhkv.6

- Set 2: The "IRD-Crack Dataset": which represents our local dataset. It comprises of asphalt crack images that were collected in cooperation with the directorate of highways and bridges in Diyala governorate. It includes various types of images that present various problems for crack detection, such as shadows and stains of oil. A fixed height of one meter, directly above the pavement, was used to capture the high-quality photos. using a digital camera type (Canon RP + 18-135mm), with a resolution of (6240 × 4160). All pictures were captured during morning hours to ensure clarity and similar lighting conditions. The images in this dataset were annotated by the use of Labelme application. This dataset was prepared specially to reflect the real-world environment on local asphalt roads. Figure 7 shows a sample of these images with their corresponding masks. These images of datasets present various problems for crack detection, such as oil stains, shadows, and varying lighting conditions. This feature improves the reliability of automated pavement crack-detecting methods and simulates real-world circumstances.

The training, validation, and testing images are from the SUT-Crack dataset, whilst the real-time test images are sourced from the IRD-Crack dataset. During pre-processing, images designated for training, validation, and testing are downsized to 640 x 640 x 3 and 320 × 320 pixels. The limited size of the SUT-Crack dataset may not furnish sufficient training data to attain appropriate outcomes. To tackle this difficulty, diverse strategies are utilized to enhance the dataset and increase the quantity of photographs. Augmentation employs rotation, shifting, shearing, zooming, flipping, and reflection as shown in Figure 8.

**(a)** **(b)** (a) (b)

**Figure 6.** Sample of SUT-Crack dataset of real cracks; (a) Original image; (b) Ground truth image.

**Figure 7.** Samples of the LIR-Crack Dataset.

Original Rotation Horizontal Flip Vertical Flip Zoom Horizontal shift Vertical shift

**Figure 8**. Results of the Augmentation Operations for SUT-Crack dataset.

*4.2 Evaluation metrics*

To statistically assess the experimental To statistically assess the experimental results, many performance metrics were analysed, including Accuracy (ACC), Precision (Pr), Recall (Re), Dice Coefficient (DC), mean Average Precision (mAP), and Intersection over Union (IoU). The methods used for metric calculation are delineated in Equations (5), (6), (7), (8), (9), and (10) respectively[39].

$$ACC = \frac{TP+TN}{TP+TN+FP+FN} \quad (5)$$

$$Pr = \frac{TP}{TP+FP} \quad (6)$$

$$Re = \frac{TP}{TP+FN} \quad (7)$$

$$DC = \frac{2TP}{FP+2TP+FN} \quad (8)$$

Where TP = True Positives, TN =True Negatives, FP = False Positive, and FN = False Negatives.

The average precision (AP) denotes the area below the precision-recall curve, whereas mean average precision (mAP) refers to the average of different classes of AP values:

$$mAP = \frac{AP}{N} = \frac{\sum_1^N \int_0^1 p(r)dr}{N} \quad (9)$$

where N is the number of crack classes, p is the percentage of all anticipated positive samples that are successfully detected, and r is the percentage of all actual positive samples that were correctly detected.

The Intersection over Union (IoU) is the ratio of the intersection to the union of the predicted mask and the ground truth data, expressed as:

$$IoU = \frac{A \cap B}{A \cup B} \quad (10)$$

where A and B indicated the predicting image mask and ground truth image mask, respectively.

*4.3 Crack detection results*

We utilized validation data from the SUT-Crack dataset to assess the efficacy of the proposed crack detection algorithm in the object detection phase. The findings are displayed in Table 4, illustrating performance indicators like Precision, Recall, mAP@0.5, and mAP@0.5:0.95. At an IoU threshold of 0.5, the suggested YOLOv10 model attained an

(mAP@0.5) of 68.90%. When the IoU threshold was varied from 0.5 to 0.95, it produced an mAP@0.5:0.95 of 54.58%, as seen in Figure 9. Furthermore, Figure 10 and Figure 11 depict the precision-confidence and recall-confidence curves, affirming the model's resilience across different confidence levels. Figure 12 illustrates that YOLOv10 effectively detects fractures, even under adverse situations like noise, illumination fluctuations, and oil stains. YOLOv10 not only delivered great detection accuracy but also enhanced computational efficiency, rendering it particularly suitable for real-time asphalt crack detection systems. This results from its enhanced design, which diminishes processing time and resource use while preserving detection accuracy.

**Table 4:** Detection Results and compare with previous studies. This mark "N/R" indicates that the results are not available.

| Author | method | dataset | Precision | Recall | mAP@0.5 | mAP@0.5:0.95 | mAP |
|---|---|---|---|---|---|---|---|
| Deng et al,2023 | YOLOv5& Attention ResNet | RDD | N/R | N/R | N/R | N/R | 91 |
| Li et al,2024 | YOLOv8& attention mechanism (SimAM) | RDD 2022 | N/R | N/R | 62.5 | 36.4 | N/R |
| YOLOv10 (ours) | YOLOv10& UNET 3+ | SUT-Crack | 100 | 91 | 68.90 | 54.58 | N/R |

Table 4 demonstrates that the suggested YOLOv10 model attained enhanced accuracy and recall relative to prior studies. Although the study of Li et al,2024 [7] indicated a diminished mAP\@0.5 and lacked other metrics, and Deng et al,2023[8] presented merely an aggregate mAP score without a detailed analysis, our methodology delivers a more thorough and dependable assessment across various thresholds, substantiating the efficacy of YOLOv10 in practical detection contexts.

*4.4 Crack Segmentation Results*

The loss percentages indicate the disparity between the predicted outcomes and the actual ground truth values. Lower loss percentages indicate a higher concordance between the predicted and actual values, so implying that the model successfully absorbed the fundamental patterns present in the training data. Figure 13 Illustrates both training/validation loss. Utilizing 100 epochs. It indicates that the minimal loss score attained during training and validation was 0.16.

The aforementioned values indicate the model's excellent performance and efficacy, demonstrating its generalizability, as the results are consistently low and equivalent in both training and validation scenarios.

Figure 14 depicts visual representations of the model's training and validation performance with epoch = 100 for Accuracy, Precision, and Recall metrics. The maximum accuracy, precision, and recall scores achieved during training and validation were 0.9906, 0.974, and 0.999, respectively.

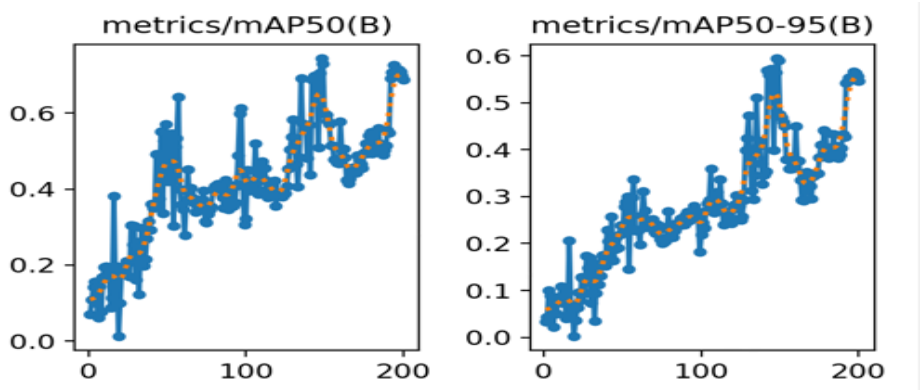



**Figure 9.** The network's performance during the validation process: (a) at an IoU threshold of 0.5, the computed mAP (mAP@0.5), and (b) with the IoU threshold varying from 0.5 to 0.95, the computed mAP (mAP@0.5: 0.95).

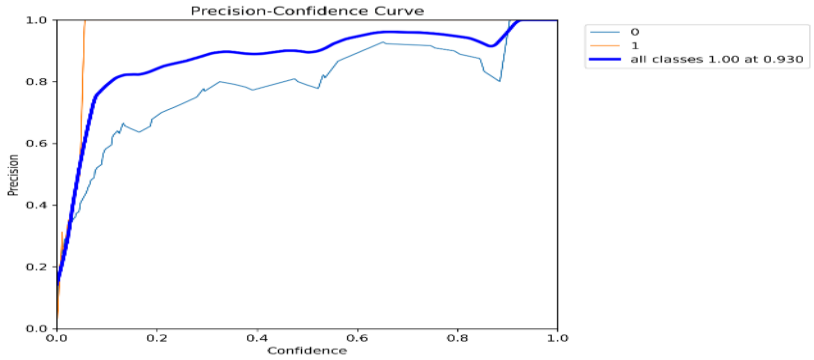



**Figure 10.** Results of the Precision-Confidence Curve.

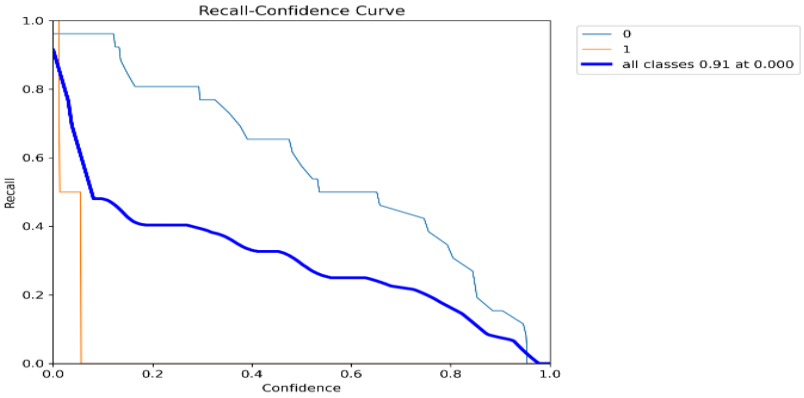



**Figure 11.** Results of the Recall-Confidence Curve

The results indicate that the model demonstrated efficiency in making accurate predictions. In contrast, the equilibrium between recall and precision indicates that the model is accurate in identifying cracks and thorough in encompassing every relevant feature, hence minimizing false negatives. Reflecting effective performance and robust generalization capability.

The proposed system assesses the model utilizing various metrics, including (IoU) and the Dice coefficient. Our model attained an IoU of 0.956 and a Dice coefficient of 0.977, signifying a substantial correspondence between the predicted and actual segmentations.

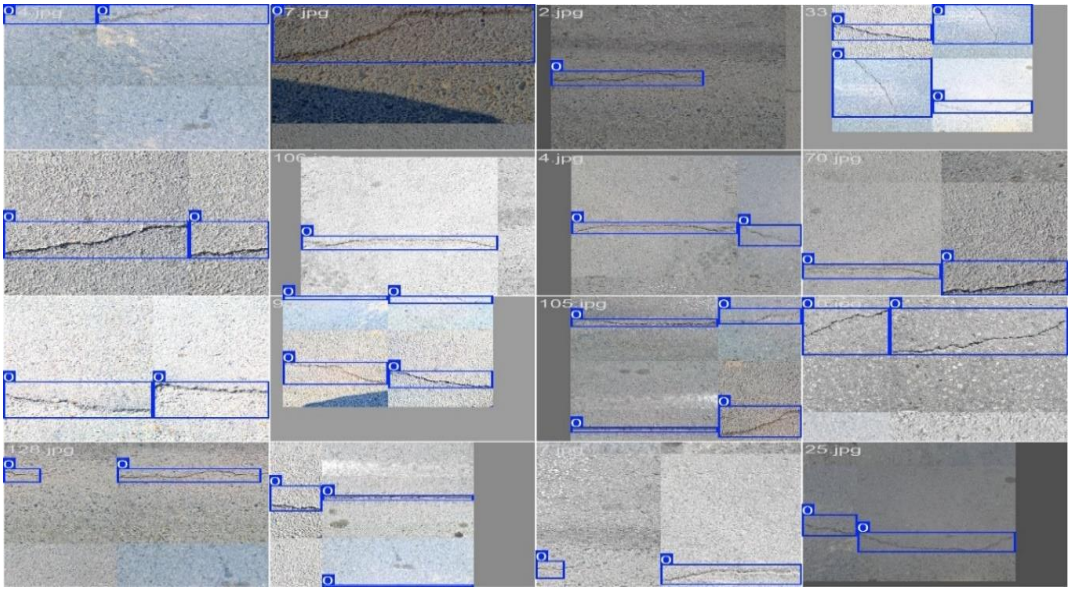


**Figure 12.** Results of the crack detection with YOLOv10

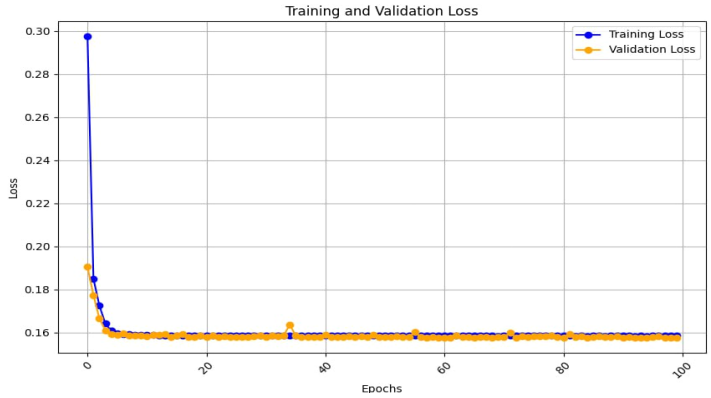
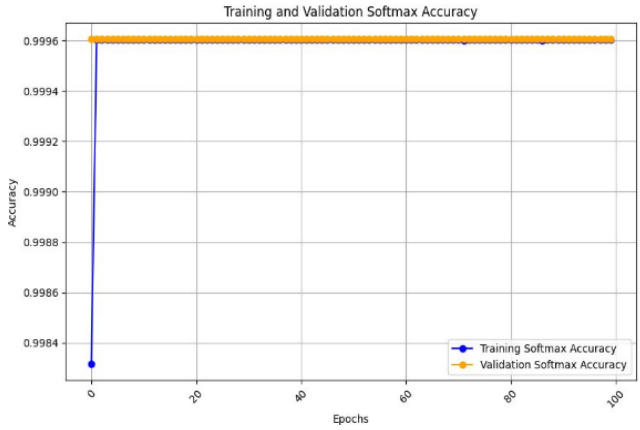
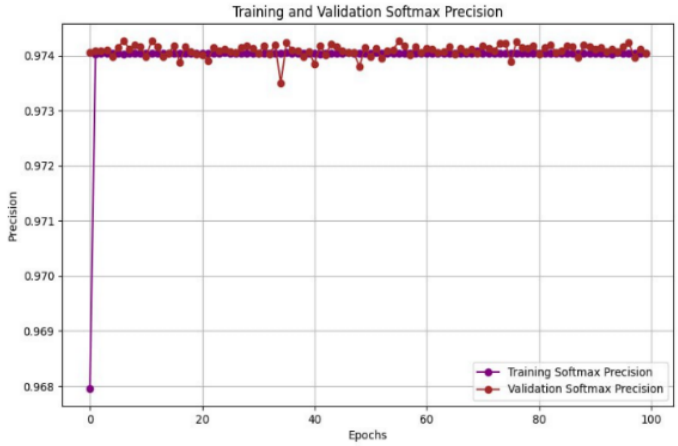


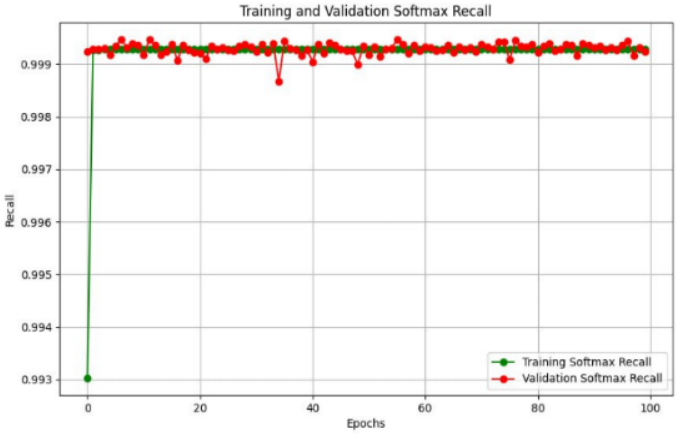

**Figure 13.** Training and validation loss




**(a)**




**(b)**




**(c)**

**Figure 14.** Training and Validation (a) Accuracy (b) Precision (c) Recall.

Table 5 provides a comparative evaluation of our model's segmentation efficacy relative to past studies. Compared with the study in [7], which attained an IoU of 87.00 and a Dice coefficient 93.14, our model markedly enhances both measures, achieving 95.6% IoU and 97.7% Dice. In comparison to [8], which obtained an IoU of 0.7644, our method exhibits a significant enhancement in segmentation accuracy. The results demonstrate the superiority of our model in accuracy and segmentation efficacy, attributable to the integration of the Residual-Attention UNet architecture, which improves feature extraction and segmentation precision. Integrating residual connections and attention mechanisms enhances fracture identification by emphasizing the most related features, resulting in superior segmentation tasks. Figure 14 provides samples of Results of the Crack Segmentation with Residual-Attention UNet. 3+

*4.5 Crack Quantification Results:*

Figure 15 shows the ideal accuracy of the proposed system in analyzing cracks and converting them into actual measurements (mm) with high accuracy. This outstanding accuracy makes the system very suitable for use in the early detection of pavement cracks, where accurate and reliable image analysis plays a crucial role in preventing pavement deterioration and is a valuable tool for road and bridge maintenance officials, enabling them to make informed decisions based on the results of high accuracy image analysis. Figure 16 represents the image testing process and crack analysis (direction, length, width, and angle).

**Table 5:** Segmentation Results and compare with previous studies. This mark "-" indicates that the results are not available.

| models | Accuracy | Precision | Recall | IoU | Dice |
|---|---|---|---|---|---|
| [7] | 98.47 | - | - | 87.00 | 93.14 |
| [8] | - | 86.41 | 84.97 | 0.7644 | - |
| **(Ours)** | 98.96 | 96.76 | 98.74 | 95.60 | 97.74 |

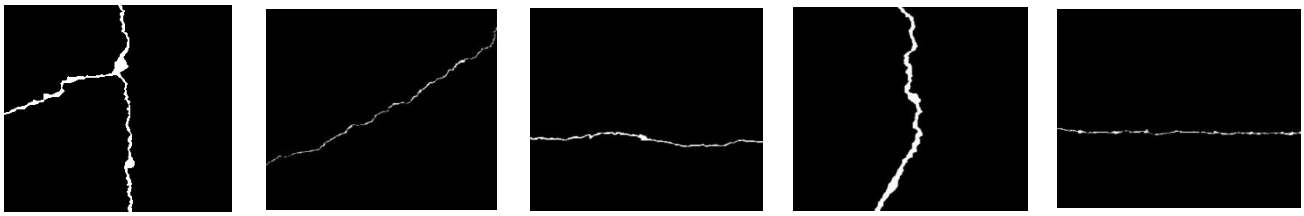


**Figure 15.** Results of the crack segmentation with residual-attention UNet. 3+.

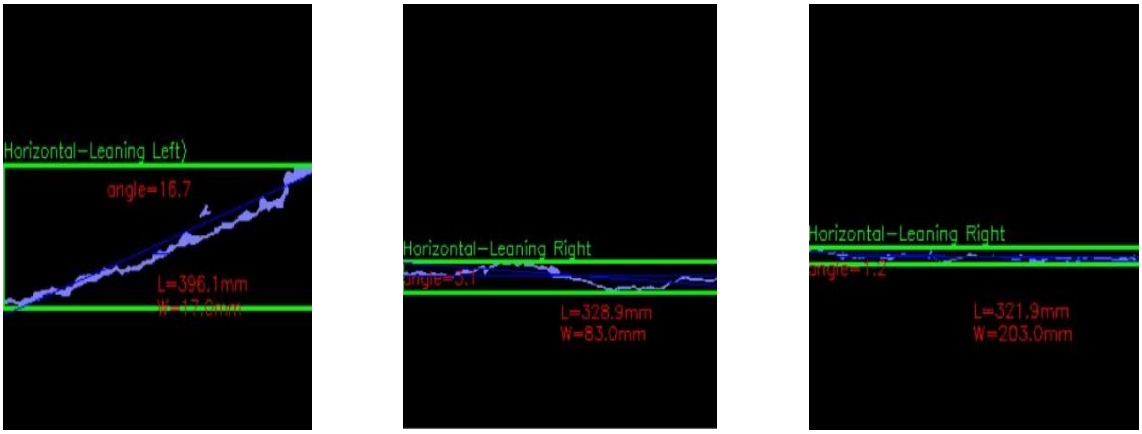



**Figure 16.** Results of the crack quantification results

### *4.6 Limitations and Challenges*

Notwithstanding the promising outcomes attained by the proposed method, several limits must be recognized. The restricted size of the SUT-Crack dataset may impede the model's generalizability, notwithstanding the augmentation procedures employed. Ultimately, fluctuations in illumination, shadows, and oil stains within the IRD-Crack dataset may provide issues for consistent identification, necessitating enhanced resilience strategies in future implementations.

## 5. Conclusion

In conclusion, we proposed a method for highly accurately assessing road cracks under complex backgrounds. An integrated framework is proposed that combines crack detection, segmentation, and quantification based on an image-processing approach with the help of deep learning techniques. Crack detection is first detected using YOLOv10 and then fed into the improved residual-attention UNet 3+ model for crack segmentation. We proposed a new method for quantification by

introducing an algorithm to search for connected components of cracks, find the crack's optimal contour, and analyze it for real measurements.

Consequently, we reached the following conclusions: The suggested technique can accurately detect at the pixel scale. Our method's superiority was assessed using precision, recall, mAP@0.5, and mAP@0.5:0.95 measures, demonstrating greater object detection accuracy than prior studies. The crack detection method attained 100% precision, 91% recall, and 68.90% mAP@0.5. Incorporating an attention gate and residual connection significantly enhances the accuracy of Residual-Attention UNet 3+ for crack segmentation, resulting in an IOU of 95.60% and a dice coefficient of 97.74% for the segmented cracks. The advanced crack quantification technique can significantly mitigate pavement damage by analyzing cracks and translating them into precise measures (mm). These data provide an accurate assessment and characterization of the cracks, hence aiding maintenance teams in executing appropriate maintenance strategies.

In future work, we aim to expand the local dataset to ensure the diversity and severity of pavement defects to detect patching, erosion, and many other defects, not just cracks in road pavements. We also aspire to promote the system through edge computing to detect cracks directly from edge devices such as drones and IoT evaporators. This can provide real-time crack detection and segmentation from photographs or videos.

## References


[1] X. Feng *et al.*, "Pavement crack detection and segmentation method based on improved deep learning fusion model," *Math. Probl. Eng.*, vol. 2020, no. 1, p. 8515213, 2020.

[2] Z. Li, H. Zhang, Z. Li, and Z. Ren, "Residual-attention UNet++: a nested residual-attention U-net for medical image segmentation," *Appl. Sci.*, vol. 12, no. 14, p. 7149, 2022.

[3] K. Malek, A. Mohammadkhorasani, and F. Moreu, "Methodology to integrate augmented reality and pattern recognition for crack detection," *Comput. Civ. Infrastruct. Eng.*, vol. 38, no. 8, pp. 1000–1019, 2023.

[4] M.-V. Pham, Y.-S. Ha, and Y.-T. Kim, "Automatic detection and measurement of ground crack propagation using deep learning networks and an image processing technique," *Measurement*, vol. 215, p. 112832, 2023.

[5] K. Sarkar, A. Shiuly, and K. G. Dhal, "Revolutionizing concrete analysis: An in-depth survey of AI-powered insights with image-centric approaches on comprehensive quality control, advanced crack detection and concrete property exploration," *Constr. Build. Mater.*, vol. 411, p. 134212, 2024.

[6] K. C. Laxman, N. Tabassum, L. Ai, C. Cole, and P. Ziehl, "Automated crack detection and crack depth prediction for reinforced concrete structures using deep learning," *Constr. Build. Mater.*, vol. 370, p. 130709, 2023.

[7] Y. Li, C. Yin, Y. Lei, J. Zhang, and Y. Yan, "RDD-YOLO: Road Damage Detection Algorithm Based on Improved You Only Look Once Version 8," *Appl. Sci.*, vol. 14, no. 8, p. 3360, 2024.

[8] 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.

[9] Z. Shu, Z. Yan, and X. Xu, "Pavement Crack Detection Method of Street View Images Based on Deep Learning," *J. Phys. Conf. Ser.*, vol. 1952, no. 2, 2021, doi: 10.1088/1742-6596/1952/2/022043.

[10] Q. An, X. Chen, X. Du, J. Yang, S. Wu, and Y. Ban, "Semantic Recognition and Location of Cracks by Fusing Cracks Segmentation and Deep Learning," *Complexity*, vol. 2021, 2021, doi: 10.1155/2021/3159968.

[11] Z. Zhang, Q. Liu, and Y. Wang, "Road Extraction by Deep Residual U-Net," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 5, pp. 749–753, 2018, doi: 10.1109/LGRS.2018.2802944.

[12] Q. Zhang *et al.*, "Improved U-net network asphalt pavement crack detection method," *PLoS One*, vol. 19, no. 5 May, pp. 1–21, 2024, doi: 10.1371/journal.pone.0300679.

[13] M. He and T. L. Lau, "CrackHAM: A Novel Automatic Crack Detection Network Based on U-Net for Asphalt Pavement," *IEEE Access*, vol. 12, no. November 2023, pp. 12655–12666, 2024, doi: 10.1109/ACCESS.2024.3353729.

[14] Z. Zhang, K. Yan, X. Zhang, X. Rong, D. Feng, and S. Yang, "Automated highway pavement crack recognition under complex environment," *Heliyon*, vol. 10, no. 4, p. e26142, 2024, doi: 10.1016/j.heliyon.2024.e26142.

[15] M. Hussain, "Yolov5, yolov8 and yolov10: The go-to detectors for real-time vision," *arXiv Prepr. arXiv2407.02988*, 2024.

[16] A. J. Yousif and M. H. Al-Jammas, "Real-time Arabic Video Captioning Using CNN and Transformer Networks Based on Parallel Implementation," *Diyala J. Eng. Sci.*, pp. 84–93, 2024.

[17] M. M. Islam, M. B. Hossain, M. N. Akhtar, M. A. Moni, and K. F. Hasan, "CNN based on transfer

learning models using data augmentation and transformation for detection of concrete crack," *Algorithms*, vol. 15, no. 8, p. 287, 2022.

[18] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *J. big data*, vol. 6, no. 1, pp. 1–48, 2019.

[19] R. J. Kolaib and J. Waleed, "Crime Activity Detection in Surveillance Videos Based on Developed Deep Learning Approach," *Diyala J. Eng. Sci.*, pp. 98–114, 2024.

[20] T. Diwan, G. Anirudh, and J. V Tembhurne, "Object detection using YOLO: Challenges, architectural successors, datasets and applications," *Multimed. Tools Appl.*, vol. 82, no. 6, pp. 9243–9275, 2023.

[21] B. Liu, W. Zhao, and Q. Sun, "Study of object detection based on Faster R-CNN," in *2017 Chinese automation congress (CAC)*, IEEE, 2017, pp. 6233–6236.

[22] H. Oliveira and P. L. Correia, "Automatic road crack segmentation using entropy and image dynamic thresholding," in *2009 17th European Signal Processing Conference*, IEEE, 2009, pp. 622–626.

[23] C.-Y. Wang and H.-Y. M. Liao, "YOLOv1 to YOLOv10: The fastest and most accurate real-time object detection systems," *APSIPA Trans. Signal Inf. Process.*, vol. 13, no. 1, 2024.

[24] M. F. Rashad and Q. I. Ali, "Deploying Android-Based Smart RSUs with YOLOv8 and SAHI for Enhanced Traffic Management," *Diyala J. Eng. Sci.*, pp. 70–90, 2025.

[25] A. Wang, H. Chen, L. Liu, K. Chen, Z. Lin, and J. Han, "Yolov10: Real-time end-to-end object detection," *Adv. Neural Inf. Process. Syst.*, vol. 37, pp. 107984–108011, 2024.

[26] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.

[27] O. Ronnerberger, P. Fischer, and T. Brox, "U-Net: Convolutional Neural Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI*, 2015.

[28] H. Huang *et al.*, "Unet 3+: A full-scale connected unet for medical image segmentation," in *ICASSP 2020-2020 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, IEEE, 2020, pp. 1055–1059.

[29] O. Oktay *et al.*, "Attention u-net: Learning where to look for the pancreas," *arXiv Prepr. arXiv1804.03999*, 2018.

[30] J. Naranjo-Alcazar, S. Perez-Castanos, I. Martin-Morato, P. Zuccarello, and M. Cobos, "On the performance of residual block design alternatives in convolutional neural networks for end-to-end audio classification," *arXiv Prepr. arXiv1906.10891*, 2019.

[31] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[32] A. M. Hafiz, S. A. Parah, and R. U. A. Bhat, "Attention mechanisms and deep learning for machine vision: A survey of the state of the art," *arXiv Prepr. arXiv2106.07550*, 2021.

[33] A. J. Yousif and M. H. Al-Jammas, "A Lightweight Visual Understanding System for Enhanced Assistance to the Visually Impaired Using an Embedded Platform," *Diyala J. Eng. Sci.*, pp. 146–162, 2024.

[34] T. Yun *et al.*, "Individual tree crown segmentation from airborne LiDAR data using a novel Gaussian filter and energy function minimization-based approach," *Remote Sens. Environ.*, vol. 256, p. 112307, 2021.

[35] Z. Azouz, B. Honarvar Shakibaei Asli, and M. Khan, "Evolution of crack analysis in structures using image processing technique: A review," *Electronics*, vol. 12, no. 18, p. 3862, 2023.

[36] J. Toribio, J.-C. Matos, and B. González, "Aspect ratio evolution in embedded, surface, and corner cracks in finite-thickness plates under tensile fatigue loading," *Appl. Sci.*, vol. 7, no. 7, p. 746, 2017.

[37] D. Schlicke, E. M. Dorfmann, E. Fehling, and N. V. Tue, "Calculation of maximum crack width for practical design of reinforced concrete," *Civ. Eng. Des.*, vol. 3, no. 3, pp. 45–61, 2021.

[38] M. Sabouri and A. Sepidbar, "SUT-Crack: A comprehensive dataset for pavement crack detection across all methods," *Data Br.*, vol. 51, p. 109642, 2023.

[39] M. Lan, D. Yang, S. Zhou, and Y. Ding, "Crack detection based on attention mechanism with YOLOv5," *Eng. Reports*, vol. 7, no. 1, p. e12899, 2025.

[40] 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, 2023, doi: 10.3390/rs15061530.