وزارة التعليم العالي والبحث العلمي
جامعة ميسان
كلية التربية الاساسية
Ministry of Higher Education and Scientific Research
University of Misan
College of Basic Education

Misan Journal for Academic Studies

Humanities, social and applied sciences

مجلة ميسان للدراسات الاكاديمية

العلوم الانسانية والاجتماعية والتطبيقية

ISSN (Print) 1994-697X
(Online)- 2706-722X

المجلد 24 العدد 54 حزيران 2025

Vol 24 Issue 54 June 2025

مجلة ميسان للدراسات الاكاديمية

العلوم الإنسانية والاجتماعية والتطبيقية

كلية التربية الأساسية/ جامعة ميسان/العراق

Misan Journal for Academic Studies

Humanities, social and applied sciences

College of Basic Education/University of Misan/Iraq

ISSN (Print) 1994-697X (Online) 2706-722X

حزيران 2025	العدد 54	24 المجلد
June 2025	Issue 54	Vol 24

OJS / PKP

www.misan-jas.com

journal.m.academy@uomisan.edu.iq

رقم الايداع في دار الكتب والوثائق بغداد 1326 في 2009

Misan Journal for Academic studies
Vol 24 Issue54 June2025
مجلة ميسان للدراسات الاكاديمية
مجلد 24 العدد 54 حزيران 2025

Early detection and segmentation of asphalt pavement cracks: Iraqi highways as case study

Shemeam T. Muhey [1], Sinan A. Naji [2]
[1] Informatics Institute for Postgraduate Studies, University of Information Technology and Communications, Baghdad, Iraq
[2] Dept. of Postgraduate Studies, University of Information Technology and Communications, Baghdad, Iraq
ms202220727@iips.edu.iq[1] , dr.sinannaji@uoitc.edu.iq[2]
https://orcid.org/0000-0001-9382-0881

Misan jas
Misan Journal of Academic Studies
ISSN (Print) 1994-697X
ISSN (Online) 2706-722X
DOI: https://doi.org/10.54633/2333-024-054-006
Received:15Apr/2025
Accepted: 3/May /2025
Published online:30/June/2025

Abstract:

Ensuring the safety and reliability of roadway infrastructure is a significant task that requires continuous monitoring and analyzing of asphalt cracks. This issue involves automatic detecting and quantifying various types of pavement cracks, which reflect the overall condition and safety of the pavement. Consequently, 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.

This paper proposed an interesting system for detecting asphalt cracks in roads based on deep learning techniques that combines object detection and sematic segmentation, first The YOLOv10 model serves as the object detection framework. Second, crack segmentation model is employed using attention U-Net to carry out pixel-level segmentation.

The system has been trained, evaluated, and tested using two datasets: The SUT-Crack dataset and the IRD-Crack dataset. The proposed system shows excellent performance across different metrics such as Recall, Precision, Accuracy, mAP, Confidence score, and Dice coefficient. The accuracy reached up to 99.01%, demonstrating its ability to be applied in real-world environments.

Keywords: Asphalt crack detection, Semantic segmentation, Object detection, YOLO model, U-NET

1.Introduction:

It is crucial to continuously monitor and analyze asphalt cracks in order to ensure the safety and dependability of road infrastructure (Pauly et al. 2017), (Nguyen et al. 2023). This topic entails automatically recognizing and measuring numerous forms of asphalt cracks, which represent the overall state and safety of the pavement. In order to prevent pavement degradation, lower maintenance costs, and improve safety strategies that aim to shield workers and road users from potential accidents that could result in Serious consequences, early detection of these cracks can significantly support the maintenance planning process (Salman et al. 2013), (Iacono and Levinson 2016). Generally, road networks and other transportation methods positively affected the economic growth of the country, communication, and social opportunities (Iacono and Levinson 2016), (Dhamija and Dhaka 2015). Typically, a number of environmental conditions cause roads to deteriorate in various ways. cracks are a sign for the degradation in road (Mandal et al. 2018). In general, cracks can cause pavement to deteriorate if they are not fixed in a timely manner. Actually, there are a variety of crack kinds, and each one relates to the primary factor causing pavement deterioration (Hamishebahar et al. 2022). We now give a definition of automatic crack detection: Given an arbitrary image, the goal of automatic crack detection is to determine whether or not there are any asphalt cracks in the image and, if present, return the image location and extent of each crack. The challenges associated with crack detection can be attributed to the following factors: lighting conditions, weather conditions, presence or absence of structural components, image orientation, shadows, oil stains, scaling, resolution, etc.

Generally, asphalt crack detection is usually done either by the labor manual visual inspection or by the automatic image-based methods (Liu 2024). The automatic image-based methods are usually divided into two categories: the non-learning-based methods and the learning-based methods (Liu 2024), (Abdellatif et al. 2020). The non-learning-based techniques mostly use different image processing techniques (e.g., edge detection, image segmentation, thresholding, morphological operation, etc.) (Abdel-Qader et al. 2003), (Zhao et al. 2010), (Yang et al. 2022). On the other hand, the deep learning techniques based on deep Convolutional Neural Networks (CNNs) can be considered as one of the most influential innovations in the area of classification and pattern recognition. Generally speaking, certain deep CNNs such as GoogLeNet (Szegedy et al. 2015), VGG-16 (Simonyan 2014), AlexNet (Krizhevsky et al. 2012), ResNet (He et al. 2016), U-Net (Ronneberger et al. 2015), DeepLab (Chen 2017), YOLO (Redmon 2016), RetinaNet (Lin 2017) (Abdulrahman 2024),(Majeed 2023), etc., have become widely known standards and are currently being incorporated within many applications. Many authors have applied these CNN architectures in crack detection and classification models. This study proposes an interesting semantic segmentation model based on deep CNN models for asphalt crack detection and quantification. The proposed model combines the YOLOv10 model and attention U-Net structure with pixel feature extractor networks, which assign a class label for each pixel in the input image to enhance the general performance of the CNNs.

The remainder of this paper is organized as follows. Section 2 gives a brief description of related work in automatic crack detection. Section 3 derives our proposed CNNs model. Section 4 presents the experimental results using two different datasets, and Section 5 concludes this paper with discussions and future work.

2.Related work:

To identify asphalt cracks in a single image of intensity or color, many techniques have been developed; learning algorithm-based techniques have garnered a lot of attention lately and have shown promising results. . Li et al. introduced an interesting version of the road crack detection model called RDD-YOLO (Li et al. 2024). The model integrates a simple attention mechanism (SimAM) into the backbone network to enhance the focus on key information in the input image. The neck structure is optimized by replacing traditional convolution modules with GhostConv. This reduces redundant information, lowers the number of parameters, and decreases computational complexity, and this will achieve more lightweight and efficient performance in the damage detection task. Finally, the upsampling algorithm in the neck is improved by replacing the nearest interpolation with more accurate bilinear interpolation, which better restores the subtle features of the image and improves the accuracy of the detection results through a finer interpolation method.

Deng et al. proposed an integrated framework for automatic detection, segmentation, and measurement of road surface (Deng et al. 2023). In the proposed framework, three separate computer vision algorithms are innovatively combined: Firstly, the real-time object detection algorithm YOLOv5 is utilized for object-level crack detection, which achieves a mean average precision of 91%. Secondly, a modified ResNet is constructed by embedding an attention gate module to more accurately segment the cracks at the pixel level and achieves 87% intersection over union (IoU) when segmenting crack pixels. Finally, a new surface feature quantification algorithm is developed to more accurately calculate the length and width of segmental road cracks with an accuracy of 95% in identifying the cracks.

Shu et al. proposed a pavement crack recognition model that combines the street view image data source, which is a low-cost method, and the YOLOv5 target detection network (Shu et al. 2021). The result shows that this network can effectively detect cracks with mAP of over 70%.

An et al. proposes a concrete surface crack identification and size calculation method combining deep learning identification. The authors named their system the Crack Identification Network (CIN) (An et al. 2021). The accuracy rate reached more than 99%, which can effectively classify concrete cracks/non-cracks and clustering segmentation based on improved K-means and morphological methods.

Zhang Z. et al. proposed the ResU-Net, a semantic segmentation neural network, which combines the strengths of residual learning and U-Net for road area extraction from high-resolution remote sensing images (Zhang et al. 2018). This model has two advantages: first, residual units make deep network training easier. Second, the network's rich skip connections could facilitate information propagation that enables the construction of networks with fewer parameters but higher performance. The result of the proposed method, which are defined as breakeven points, is 0.9187.

Zhang Q. et al. proposed an improved U-net network for pavement crack detection and segmentation with a complex background (Zhang et al. 2024). To improve the recognition accuracy of narrow cracks in the road surface, the VGG16 and novel Up_Conv module are introduced as the backbone network. Furthermore, the Ca (Channel Attention) mechanism was added in U-net's jump connection to distinguish cracks and background noise at the same time. The DG_Conv (Depthwise GSConv Convolution) module and U-NetUp (U-Net Upsampling) module are added in the decoding part to extract richer features through more convolutional layers in the network. The results of the proposed system show a precision reached up to 87.4%.

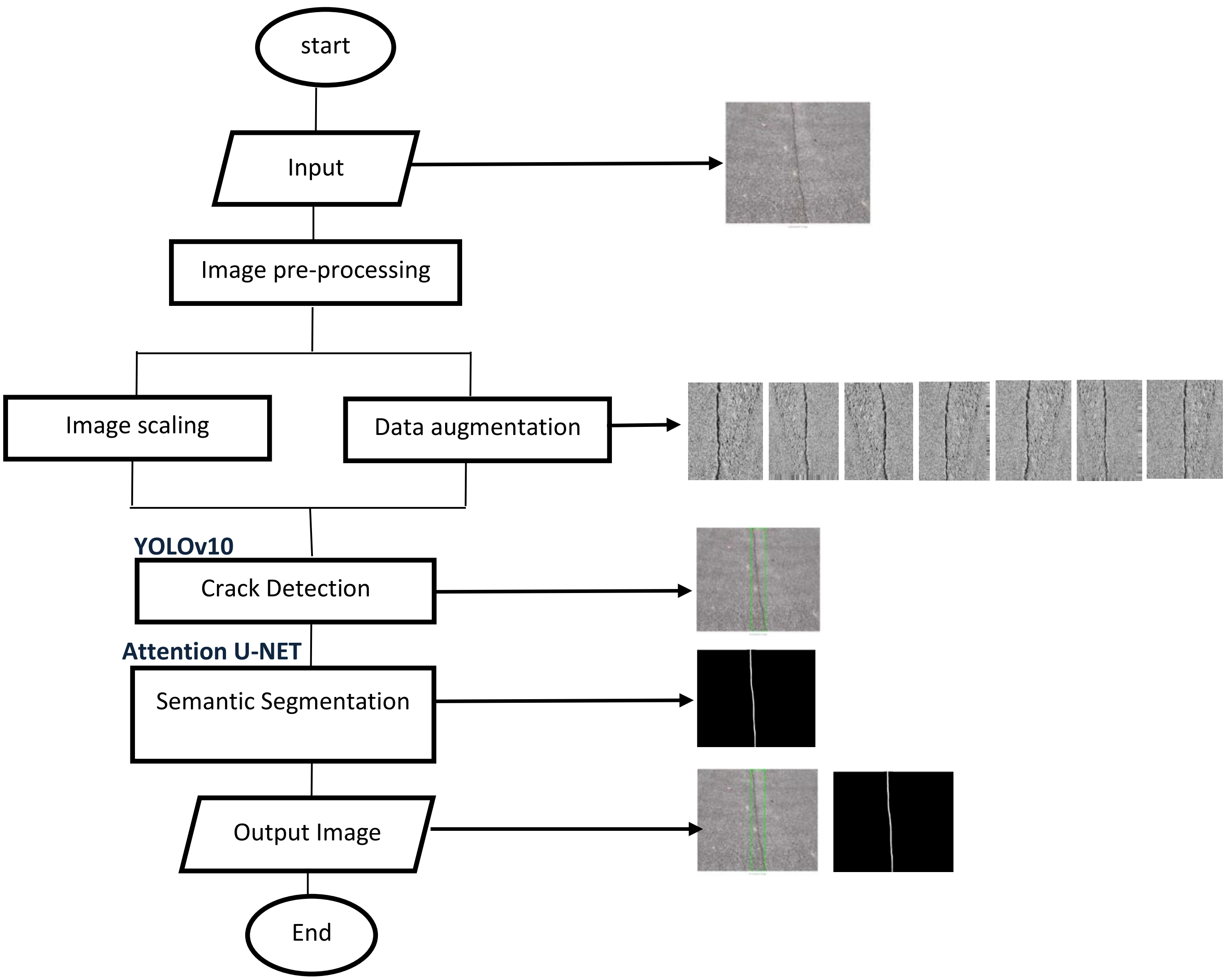
He and Lau proposed an interesting model known as CrackHAM, which is an encoder-decoder network based on the U-Net architecture and a novel network model called the HASP module, which is added to overcome the problem of spatial information degradation (He and Lau, 2024). Furthermore, the channel attention module was used to capture abundant contextual information for high-level features and spatial attention for low-level features to extract rich edge information. The

Multi-Fusion U-Net architecture is proposed to aggregate contextual information from feature maps of varying sizes during the downsampling. The system achieves a precision of 86.41%.

Zhang et al. use a novel approach to recognize multi-type cracks utilizing the ResNet model integrated with a Convolutional Block Attention Module (CBAM) (Zhang et al. 2024). The spatial attention mechanism AM is generated by leveraging the inter-spatial relationship of features. In order to create an effective feature descriptor, average-pooling and max-pooling are used. The proposed model achieves a precision of 92.9%.

3.THE PROPOSED SYSTEM:

Three stages make up the suggested system architecture for detecting asphalt cracks: image pre-processing, object detection using YOLOv10, semantic segmentation using modified U-Net model (attention U-NET), The flowchart of the proposed system design is shown in Figure (1). In the sections that follow, each of these models is briefly explained.



Figure(1): The general architecture of the proposed system of the asphalt crack detetion system

.3.1 image pre-processing:

In deep learning, image pre-processing is a crucial step to improve both the quantity and the quality of the dataset images needed for system training and to obtain a more effective learning model. a range of pre-processing techniques, including cropping, flipping, rotation, contrast improvement, color-space transformation, noise reduction, and color enhancement (Islam et al. 2022). The following sections show the various procedures that were applied on the SUT-Crack Datasets.

3.1.1 Image scaling:

Since the collected images of asphalt cracks in the different datasets have different size, the input images are resized to (640×640×3) and (320×320×3). Images are scaled to standardize all the input to be with the same dimensions. Image resizing may provide a standard input size for the model and ensures computational efficiency.

3.1.2 Image Augmentation:

The dataset images of asphalt cracks are augmented to produce new images, thereby preventing the acquisition of undesired features, mitigating overfitting, and improving overall performance. CNN-based methodologies require larger datasets for training to enhance the model's ability to learn additional image patterns and make precise predictions. The augmentation procedure enhances the training dataset by employing numerous transformations, including rotation, shifting, shearing, zooming, flipping, and reflecting (Shorten and Khoshgoftaar 2019). The corresponding values of these transformations (generators) are detailed in Table (1).

Table (1): Transformation types and their corresponding values

Transformation Type	Facilities
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:

In deep learning, it is standard practice to divide datasets into distinct subsets to facilitate efficient model training, hyperparameter tuning, and performance evaluation. In this study the dataset was divided into three parts: 70% for training to learn model parameters, 20% for validation to fine-tune hyperparameters and monitor performance, and 10% for testing to evaluate the model’s generalization capability.

3.2. object Detection Using YOLOv10:

Object detection techniques aim to locate and classify cracks within an image as objects, typically by drawing bounding boxes around the crack regions. Popular models like YOLO (You Only Look Once) (Diwan et al. 2023) or Faster R-CNN (Liu et al. 2017) are often adapted for this task. These models don't identify the exact pixel boundaries but rather detect the crack as an entity within the image Oliveira and Correia, 2009).

The YOLO (You Only Look Once) series (from v1 to v10), a single-stage object detection method, which was created in recent years, is one of the best real-time object detection algorithms available. It has significant and wide-ranging impact on numerous computer vision research projects (Wang and Liao 2024).

By combining training techniques and architectural innovations, YOLOv10 improves accuracy and efficiency. As following, we introduce some of the special features of YOLOv10:

•Dual Label Assignment: Use the label assignment method, and add stop gradient operation to the one-to-one branch.

•NMS-free Object Detection: The design of one-to-one matching mechanism enables the prediction process without relying on NMS for post-processing.

•Rank-guided Block Design: Proposed to use rank to determine which stages use conventional convolution and which stages use depth-wise convolution.

•Partial Self-attention. YOLOv10 combined CSPNet and Transformer and proposed the self-attention module (Wang and Liao 2024).

3.3 SEMANTIC SEGMENTATION BASED ON attention U-Net:

Semantic Segmentation aims to classify each pixel in an image into a pre-defined set of categories, such as road, building, or vehicle. FCN, U-Net and their variants have been widely applied for semantic segmentation in various domains, since they predict one segmentation map based on pixel-wise classification (Long et al. 2015), (Ronnerberger et al. 2015). Our asphalt crack semantic segmentation is based on U-Net architecture with attention gate, the U-Net has two paths: encoder on the left and decoder on the right. encoder path adheres to the standard convolutional network architecture. In order to perform downsampling, it involves repeatedly applying two 3x3 convolutions (unpadded convolutions), each followed by a rectified linear unit (ReLU) and a 2x2 max pooling operation with stride 2. An upsampling of the feature map, a 2x2 convolution (also known as a "up-convolution") that halves the number of feature channels, a concatenation with the correspondingly cropped feature map from the encoder path, and two 3x3 convolutions, each followed by a ReLU, make up each step of the decoder path (Sabouri and Sepidbar 2023).

In this paper Soft Attention gate is added to improve the performance of CNNs. Instead of focusing on the entire image, this technique focusses on small parts of the input image (i.e. Sub image) with different degrees of attention. Only relevant features will be the focus of training this model. So, it will provide a better generalization to the network and reduce the computational resources wasted on irrelevant features by using a soft attention mechanism that works to add weights to pixels based on relevance. Relevant parts of image get large weights and less relevant parts get small weights.

Attention Gate (AG) implemented at every skip connection between encoder and decoder for U-Net . It takes two inputs: the input (g) gate signal comes from the next lower layer of the network (decoder stage), which has the better features and the input (x) which comes from the 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. Bilinear 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.

The general structure of the attention U-Net model is depicted in Figure (2).

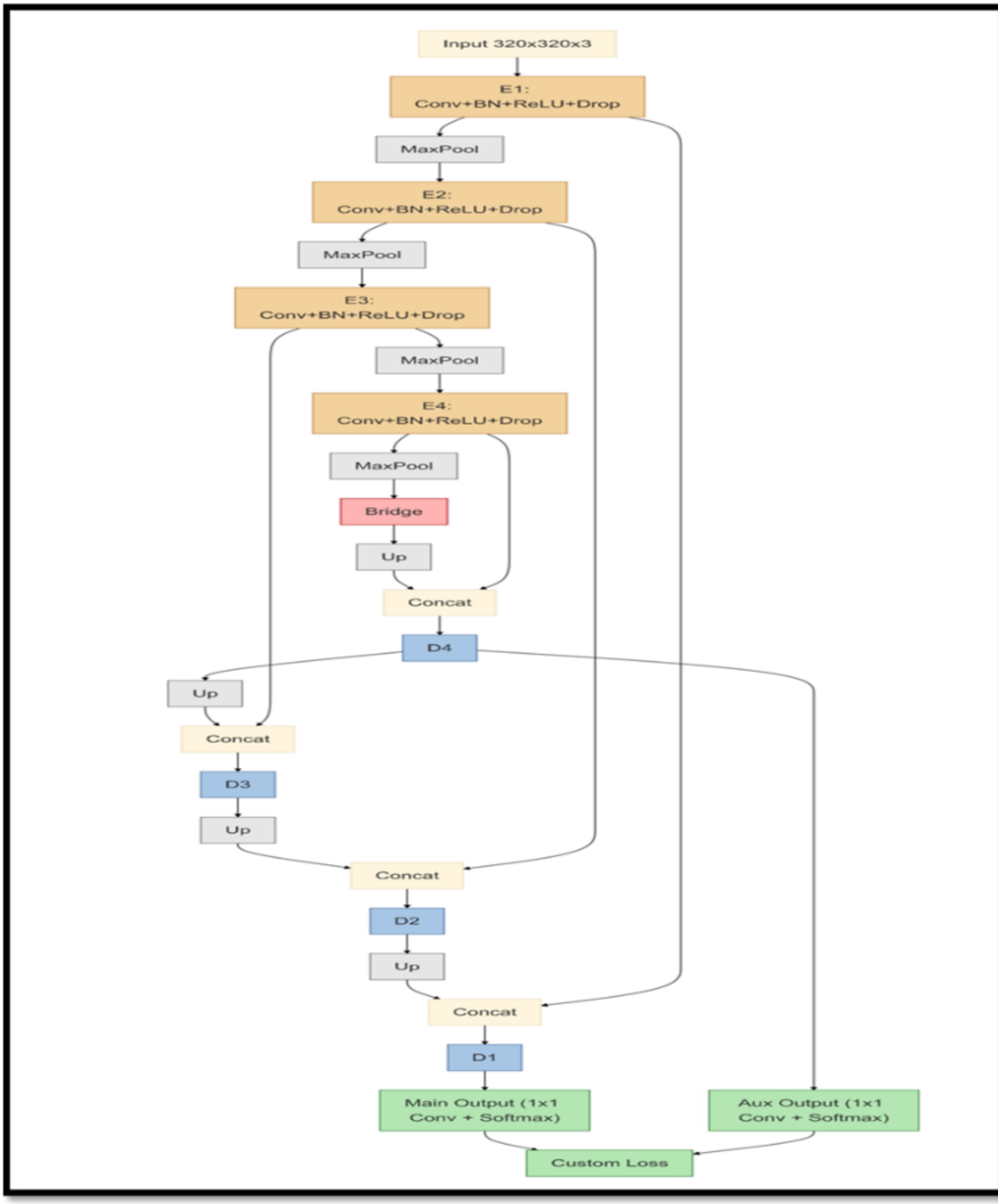


Figure (2) Attention U-Net Model Architecture

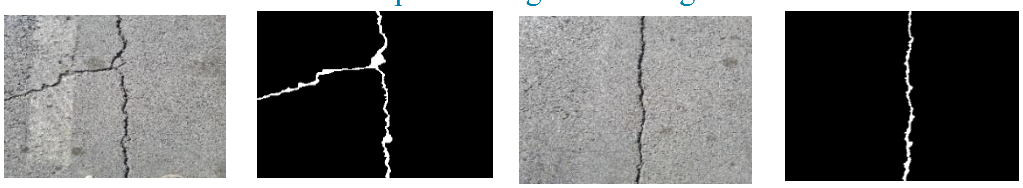
Misan Journal for Academic studies
مجلة ميسان للدراسات الاكاديمية
Vol 24 Issue54 June2025
مجلد 24 العدد 54 حزيران 2025

4. EXPERIMENTAL RESULTS AND DISCUSSION:

4.1 Dataset:

In this research, we conducted experiments using two datasets to get more accurate results these datasets are:

Set 1: The “SUT-Crack Dataset” (Sabouri and Sepidbar 2023) of the “Sharif University of Technology” implies a high-resolution original image depicting asphalt road cracks with dimensions of (3024 × 4032) pixels. The dataset is publicly available to enable crack detection through the use of many deep learning methods. Figure (3) shows Sample of SUT-Crack dataset combined with Ground truth image. SUT-Crack is available at https://doi.org/10.17632/gsbmknrhkv.6 .



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

Figure (3): Sample of SUT-Crack dataset of real cracks; (a) Original image;
(b) Ground truth image:

Set 2: The “IRD-Crack Dataset”, which represents our local dataset. It consists 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 (4) shows a sample of these images with their corresponding masks.

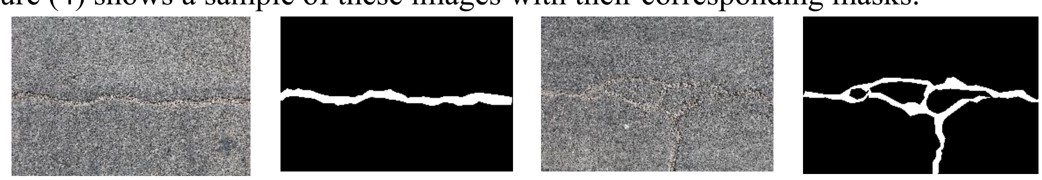


Figure (4): Samples of the IRD-Crack Dataset

4.2 Evaluation metrics:

To statistically assess the experimental results, many performance metrics were analyzed, 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 (1), (2), (3), (4), (5), and (6) respectively [18].

$$ACC = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Pr = \frac{TP}{TP+FP} \quad (2)$$

$$Re = \frac{TP}{TP+FN} \quad (3)$$

$$DC = \frac{2TP}{FP + 2TP + FN} \quad (4)$$

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:

$$mAP = \frac{AP}{N} = \frac{\sum_{1}^{N} \int_{0}^{1} p(r)dr}{N} \quad (5)$$

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:

$$IoU = \frac{A \cap B}{A \cup B} \quad (6)$$

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

4.3 Results:

Crack Detection Results:

Table (2) shows Performance metrics for YOLOv10 model on SUT-Crack dataset using 200 epochs.

Table (2): object Detection Result.

	Precision	Recall	mAP@0.5	mAP@0.5:0.95
YOLOv10(Ours)	100	0.91	0.68901	0.54582

Figure (5) shows the precision-confidence curve of YOLOv10 on SUT-Crack dataset

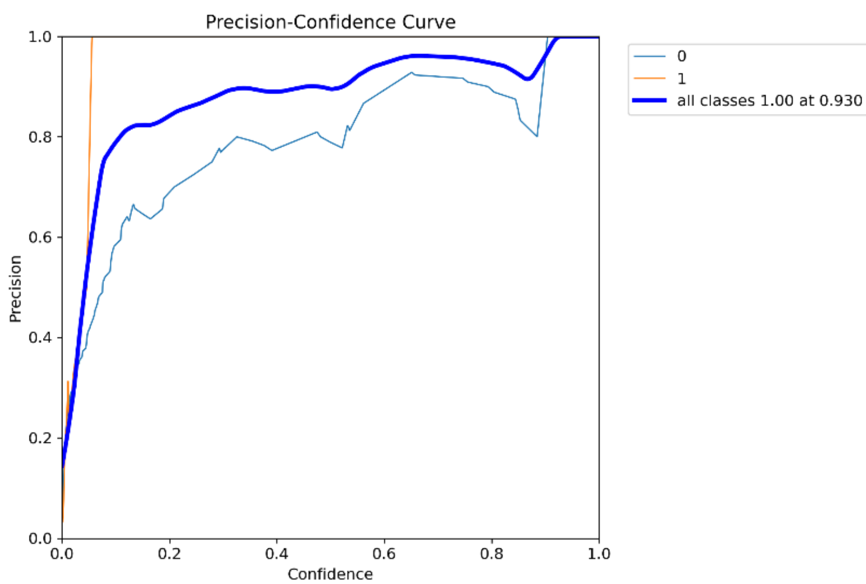


Figure (5) : Precision-confidence curve of YOLOv10 on SUT-Crack dataset

Figure (6) shows the bounding boxes around cracks as outputs of YOLOv10 model on SUT-Crack dataset

Misan Journal for Academic studies
مجلة ميسان للدراسات الاكاديمية
Vol 24 Issue54 June2025
مجلد 24 العدد 54 حزيران 2025

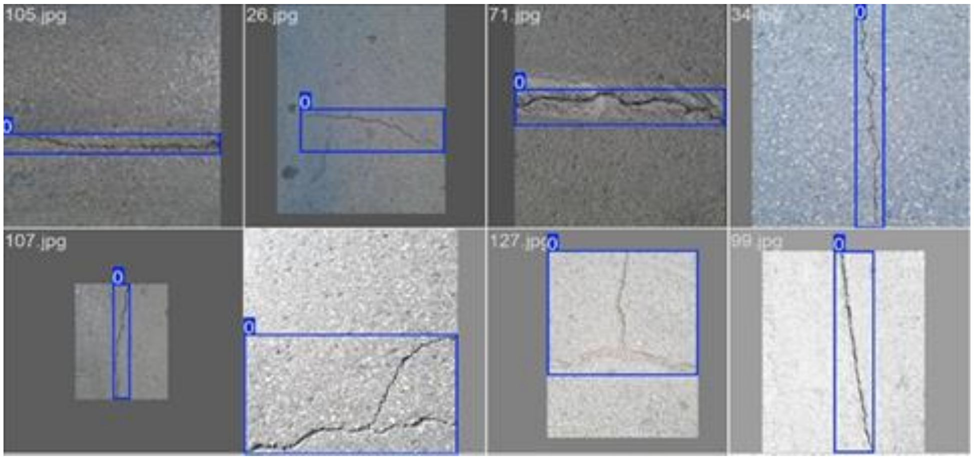
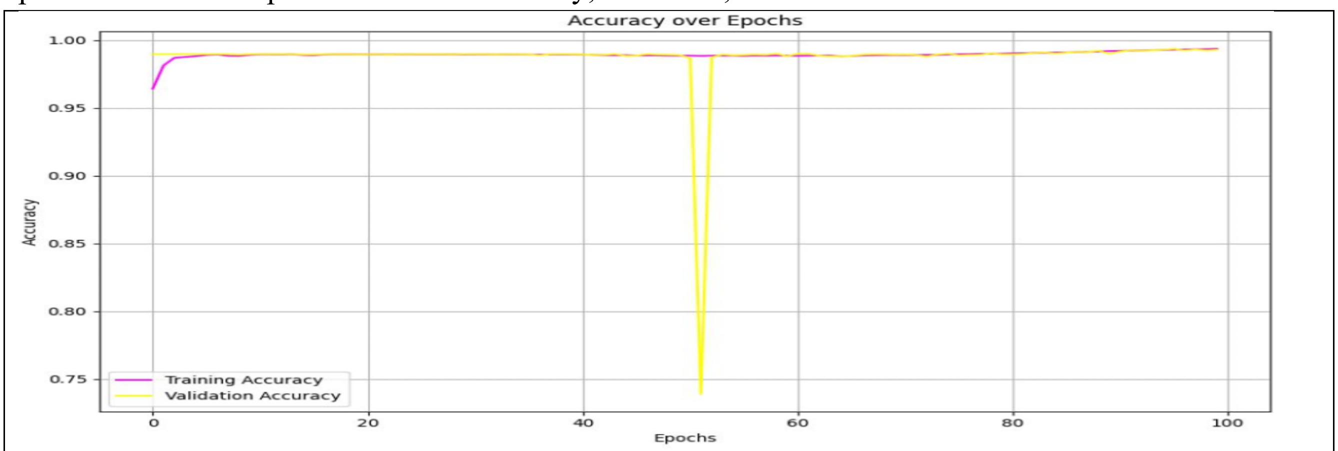


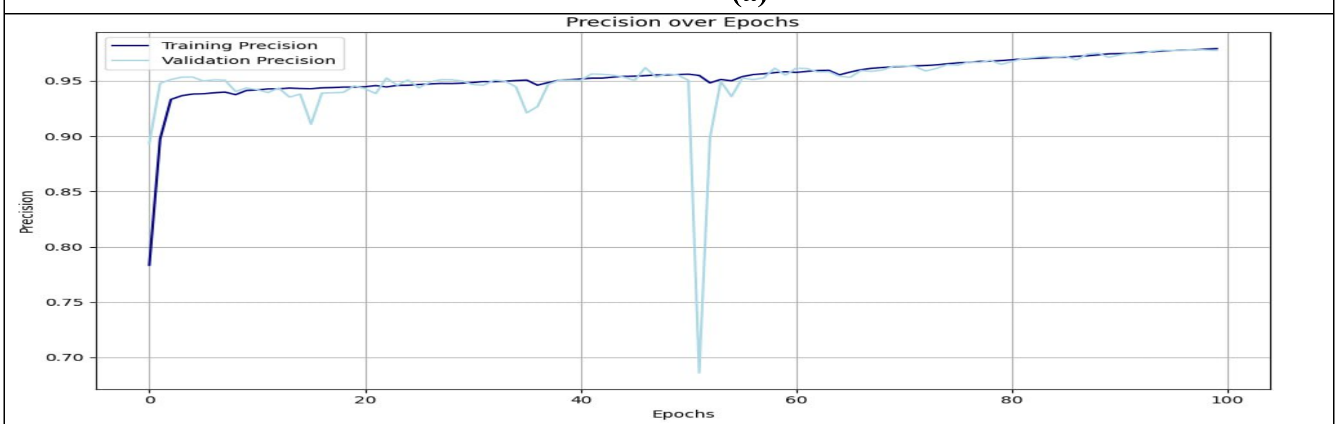
Figure (6): Object detection using YOLOv10 on SUT-Crack dataset

4.3.2 Crack Segmentation Results:

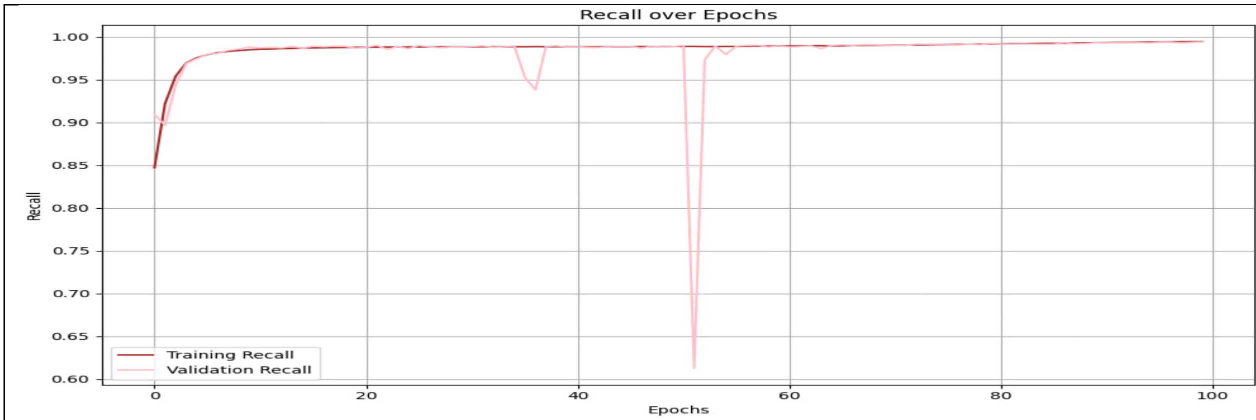
Figure (7) depicts visual representations of the Attention U-Net model's training and validation performance with epoch = 100 for Accuracy, Precision, and Recall metrics.



(a)



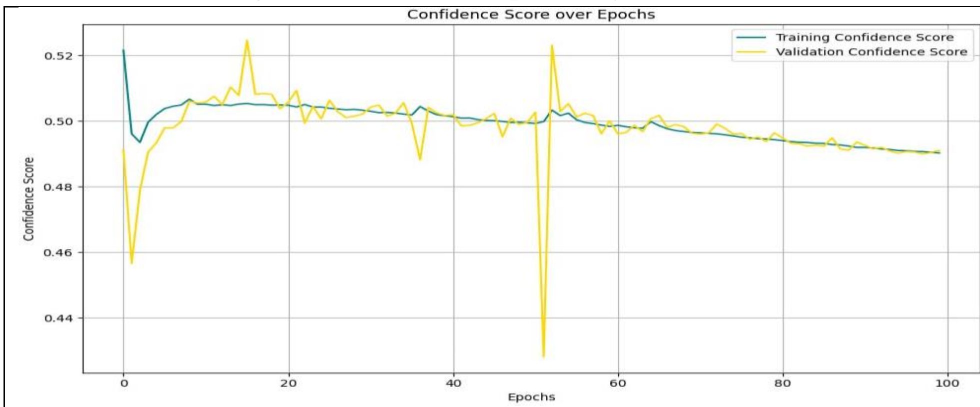
(b)



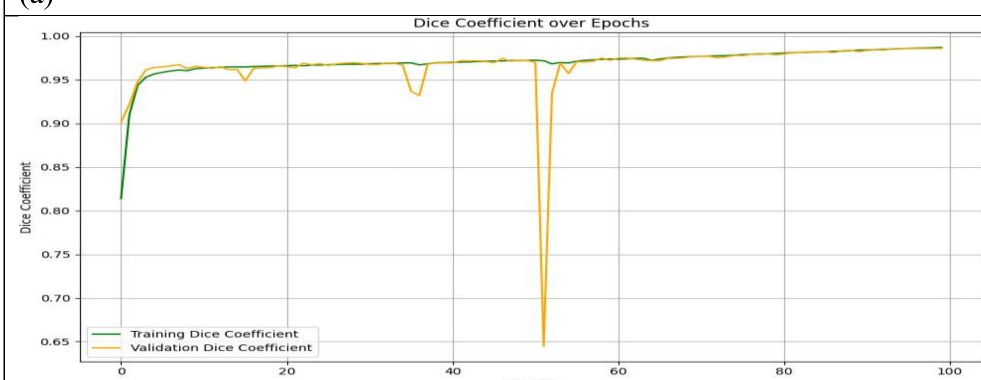
(c)

Figure (7): Training and Validation of the Attention U-Net Model (a) Accuracy (b) Precision and (c) Recall using 100 epochs

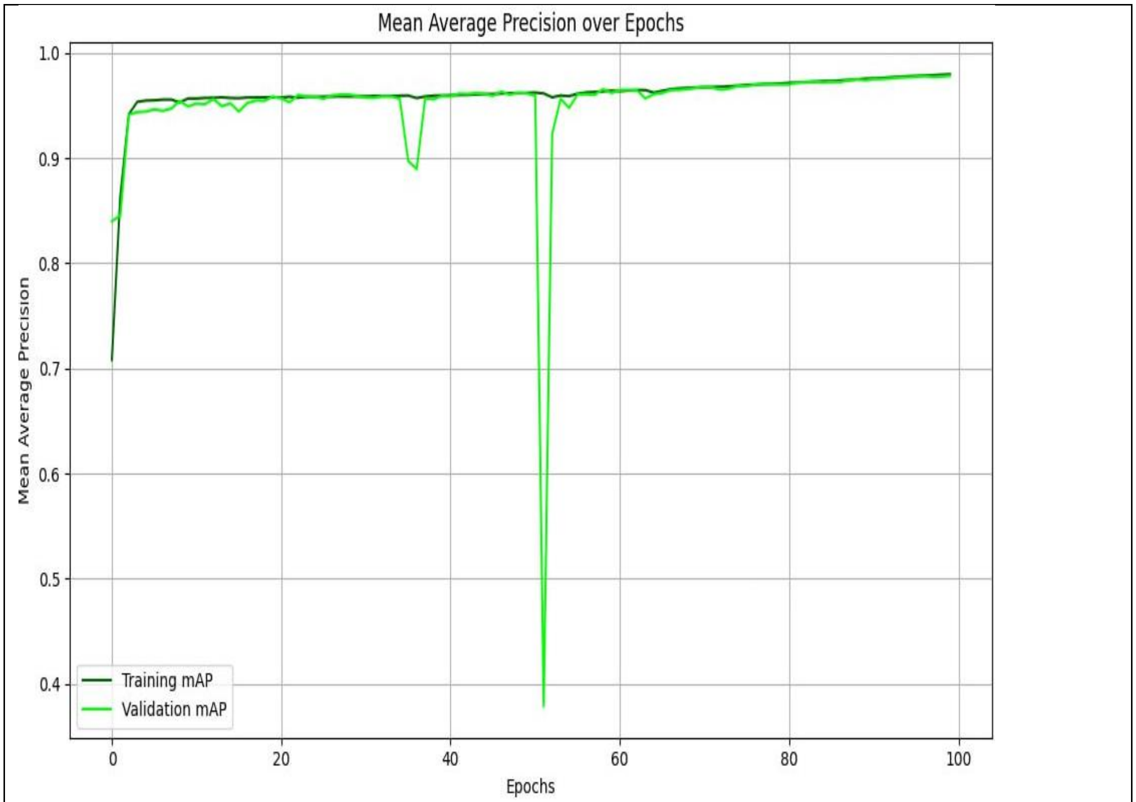
Figure (8) illustrates the trends of the Confidence score, Dice coefficient, and mAP for our proposed model across the epochs of both the training and validation sets. The confidence score was 0.49029 during training and validation, while the Dice coefficient recorded values of 0.9803 and 0.985843479 for training and validation, respectively. The mAP values were 0.97903 and 0.97806566953 for training and validation, respectively.



(a)



(b)



(c)

Figure (8) Training and Validation of the Attention U-Net Model (a) Confidence Score (b) Dice Coefficient and (c) Mean Average Precision using 100 epochs

Table (3) demonstrates the summary performance values attained by the Soft Attention U-Net algorithm for image segmentation.

Table (3): Summary of Attention U-Net Performances Model dependent on various metrics

Training			**Validation**		
Loss		**IoU Loss**	**Loss**		**IoU Loss**
0.06902		0.025829	0.077087		0.027878
Accuracy	Precision	Recall	Accuracy	Precision	Recall
0.9901	0.97902	0.9901	0.993089	0.977408	0.994474
Confidence Score	Dice Coefficient	mAP	Confidence Score	Dice Coefficient	mAP
0.49029	0.9803	0.97903	0.5	0.985843479	0.97806566953

Tabel (4) lists the performance values achieved by our proposed method compared to those reported in previous studies, considering that different datasets were used in these studies.

Misan Journal for Academic studies
مجلة ميسان للدراسات الاكاديمية
Vol 24 Issue54 June2025
مجلد 24 العدد 54 حزيران 2025

Table (4): Comparison between the Proposed model and Previous Studies

Other work	**Architecture MOEDL**	**Dataset**	**Performance Metrics**
(Deng et al. 2023)	YOLOV8	RDD2022	mAP50 = 62.5% mAP50-95 = 36.4%, F1= 69.6%.
(Zhang,Q et al. 2024)	U-net&VGG16	CFD & Deepcrack	precision=0.879, Recall=0.860,F1=0.869 mIoU=0.784 Loss=0.162 on CFD dataset
(Zhang Z. et al. 2024)	ResNet34	Images of CCD cameras	Precision=92.9%, average, recall =92.5%.
(HE and LAU 2024)	U-Net	Crack 500, DeepCrack, FIND	Precision=73.40% Recall= 77.31% F1 =73.13%,IoU =0.5973 ,AP= 0.8264
Our Model	YOLOv10, Attention U-Net	SUT-Crack & IRD dataset	Accuracy=0.9901, Precision=0.97902,Recall=0.9901, Dice Coefficient=0.9803, Map=0.97903

Figure (9) show the results of test image on IRD dataset using yolov10 and attention U-Net models

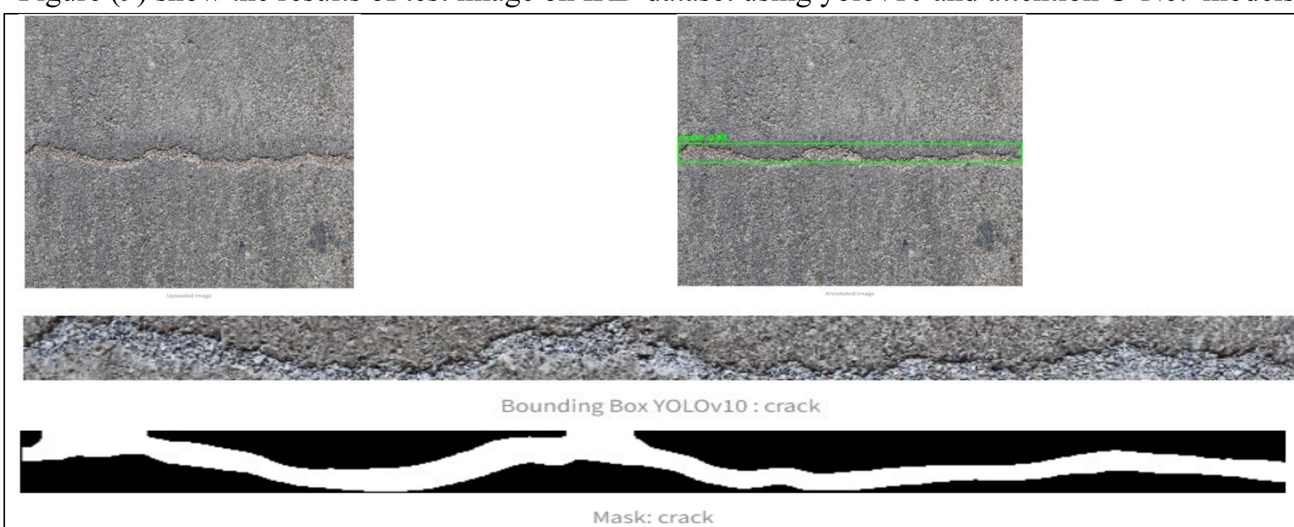


Figure (9) show the result of test image using IRD dataset

Conclusion:

A two-phase method that may enhance asphalt crack detection through the use of deep learning models and semantic recognition is the primary goal of this work. The pavement's crack region is initially defined using the YOLOv10 model, and a pixel-level segmentation procedure is carried out in the second phase using the attention U-Net.

The following conclusions are made in light of the experimental findings pertaining to the topic of asphalt crack detection:

1) A crucial step before beginning model training is image enhancement and augmentation of the existing dataset.

2) The segmentation and localization of the cracks are carried out using the suggested system in combination with the object detection model.

3) According to performance measures, the suggested method performed better than many earlier studies, which may indicate that it can be used to identify and examine local road crack images. Road maintenance and construction planning decision-making can be greatly aided by this.

4) Our original dataset, designated IRD-Crack (Iraqi-Roads Dataset), includes a wide variety of asphalt cracks and real-world image environments with varying lighting conditions.

REFERENCES:

Pauly, L., Hogg, D., Fuentes, R. and Peel, H., 2017, July. Deeper networks for pavement crack detection. In Proceedings of the 34th ISARC (pp. 479-485). IAARC.

Nguyen, S.D., Tran, T.S., Tran, V.P., Lee, H.J., Piran, M.J. and Le, V.P., 2023. Deep learning-based crack detection: A survey. International Journal of Pavement Research and Technology, 16(4), pp.943-967.

Salman, M., Mathavan, S., Kamal, K. and Rahman, M., 2013, October. Pavement crack detection using the Gabor filter. In 16th international IEEE conference on intelligent transportation systems (ITSC 2013) (pp. 2039-2044). IEEE.

Iacono, M. and Levinson, D., 2016. Mutual causality in road network growth and economic development. Transport Policy, 45, pp.209-217.

Dhamija, A. and Dhaka, V., 2015, October. A novel cryptographic and steganographic approach for secure cloud data migration. In 2015 international conference on green computing and internet of things (ICGCIoT) (pp. 346-351). IEEE.

Mandal, V., Uong, L. and Adu-Gyamfi, Y., 2018, December. Automated road crack detection using deep convolutional neural networks. In 2018 IEEE international conference on big data (big data) (pp. 5212-5215). IEEE.

Hamishebahar, Y., Guan, H., So, S. and Jo, J., 2022. A comprehensive review of deep learning-based crack detection approaches. Applied Sciences, 12(3), p.1374.

Liu, Y., 2025, January. Evaluation and optimization of image segmentation models in pavement crack detection. In Fourth International Conference on Computer Vision, Application, and Algorithm (CVAA 2024) (Vol. 13486, pp. 509-514). SPIE.

Abdellatif, M., Peel, H., Cohn, A.G. and Fuentes, R., 2020. Pavement crack detection from hyperspectral images using a novel asphalt crack index. Remote sensing, 12(18), p.3084.

Abdel-Qader, I., Abudayyeh, O. and Kelly, M.E., 2003. Analysis of edge-detection techniques for crack identification in bridges. Journal of computing in civil engineering, 17(4), pp.255-263.

Zhao, H., Qin, G. and Wang, X., 2010, October. Improvement of canny algorithm based on pavement edge detection. In 2010 3rd international congress on image and signal processing (Vol. 2, pp. 964-967). IEEE.

Yang, Z., Ni, C., Li, L., Luo, W. and Qin, Y., 2022. Three-stage pavement crack localization and segmentation algorithm based on digital image processing and deep learning techniques. Sensors, 22(21), p.8459.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V. and Rabinovich, A., 2015. Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1-9).

Simonyan, K. and Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

Krizhevsky, A., Sutskever, I. and Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 25.

He, K., Zhang, X., Ren, S. and Sun, J., 2016. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

Ronneberger, O., Fischer, P. and Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. In Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18 (pp. 234-241). Springer international publishing.

Chen, L.C., Papandreou, G., Schroff, F. and Adam, H., 2017. Rethinking atrous convolution for semantic image segmentation. arXiv preprint arXiv:1706.05587.

Redmon, J., Divvala, S., Girshick, R. and Farhadi, A., 2016. You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).

Lin, T.Y., Goyal, P., Girshick, R., He, K. and Dollár, P., 2017. Focal loss for dense object detection. In Proceedings of the IEEE international conference on computer vision (pp. 2980-2988).

Li, Y., Yin, C., Lei, Y., Zhang, J. and Yan, Y., 2024. RDD-YOLO: Road damage detection algorithm based on improved You Only Look Once version 8. Applied Sciences, 14(8), p.3360.

Deng, L., Zhang, A., Guo, J. and Liu, Y., 2023. An integrated method for road crack segmentation and surface feature quantification under complex backgrounds. Remote Sensing, 15(6), p.1530.

Shu, Z., Yan, Z. and Xu, X., 2021, June. Pavement crack detection method of street view images based on deep learning. In Journal of Physics: Conference Series (Vol. 1952, No. 2, p. 022043). IOP Publishing.

An, Q., Chen, X., Du, X., Yang, J., Wu, S. and Ban, Y., 2021. Semantic Recognition and Location of Cracks by Fusing Cracks Segmentation and Deep Learning. Complexity, 2021(1), p.3159968.

Zhang, Z., Liu, Q. and Wang, Y., 2018. Road extraction by deep residual u-net. IEEE Geoscience and Remote Sensing Letters, 15(5), pp.749-753.

Zhang, Q., Chen, S., Wu, Y., Ji, Z., Yan, F., Huang, S. and Liu, Y., 2024. Improved U-net network asphalt pavement crack detection method. Plos one, 19(5), p.e0300679.

He, M. and Lau, T.L., 2024. Crackham: A novel automatic crack detection network based on u-net for asphalt pavement. IEEe Access, 12, pp.12655-12666.

Zhang, Z., Yan, K., Zhang, X., Rong, X., Feng, D. and Yang, S., 2024. Automated highway pavement crack recognition under complex environment. Heliyon, 10(4).

Islam, M.M., Hossain, M.B., Akhtar, M.N., Moni, M.A. and Hasan, K.F., 2022. CNN based on transfer learning models using data augmentation and transformation for detection of concrete crack. Algorithms, 15(8), p.287.

Shorten, C. and Khoshgoftaar, T.M., 2019. A survey on image data augmentation for deep learning. Journal of big data, 6(1), pp.1-48.

Diwan, T., Anirudh, G. and Tembhurne, J.V., 2023. Object detection using YOLO: Challenges, architectural successors, datasets and applications. multimedia Tools and Applications, 82(6), pp.9243-9275.

Liu, B., Zhao, W. and Sun, Q., 2017, October. Study of object detection based on Faster R-CNN. In 2017 Chinese automation congress (CAC) (pp. 6233-6236). IEEE.

Oliveira, H. and Correia, P.L., 2009, August. Automatic road crack segmentation using entropy and image dynamic thresholding. In 2009 17th European signal processing conference (pp. 622-626). IEEE.

Wang, C.Y. and Liao, H.Y.M., 2024. YOLOv1 to YOLOv10: The fastest and most accurate real-time object detection systems. APSIPA Transactions on Signal and Information Processing, 13(1).

Long, J., Shelhamer, E. and Darrell, T., 2015. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3431-3440).

Ronneberger, O., Fischer, P. and Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. In Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18 (pp. 234-241). Springer international publishing.

Sabouri, M. and Sepidbar, A., 2023. SUT-Crack: A comprehensive dataset for pavement crack detection across all methods. Data in Brief, 51, p.109642.

Abdulrahman, S.A., 2024. Convolutional Neural Networks in Detection of Plant Diseases. (Humanities, social and applied sciences) Misan Journal of Academic Studies, 23(52), pp.30-42.

Majeed, H.L., 2023. Using neural networks to solve image problems through artificial intelligence. (Humanities, social and applied sciences) Misan Journal of Academic Studies, 22(47), pp.409-416.

Conflicts of Interest Statement…………………………………………………

Manuscript title:

Early detection and segmentation of asphalt pavement cracks: Iraqi highways as case study

The authors whose names are listed immediately below certify that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript**.**

Author names:

1. Shemeam T. Muhey

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

2 Sinan A. Naji.

Dept. of Postgraduate Studies, University of Information Technology and Communications, Iraq

The authors whose names are listed immediately below report the following details of affiliation or involvement in an organization or entity with a financial or non-financial interest in the subject matter or materials discussed in this manuscript. Please specify the nature of the conflict on a separate sheet of paper if the space below is inadequate.

Author names:

This statement is signed by all the authors to indicate agreement that the above information is true and correct (a photocopy of this form may be used if there are more than 10 authors):

Author's name (typed)	**Author's signature**	**Date**
Shemeam T. Muhey		13/5/2025
Sinan A. Naji		13/5/2025